Hopf algebras and non-associative algebras in the study of iterated-integral signatures and rough paths

vorgelegt von
M. Sc.
Rosa Lili Dora Preiß
ORCID:0000-0002-6145-4896

an der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr. rer. nat. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Michael Joswig
Gutachter: Prof. Dr. Peter K. Friz
Gutachter: Prof. Dr. Terry J. Lyons
Gutachter: Prof. Dr. Kurusch Ebrahimi-Fard

Tag der wissenschaftlichen Aussprache: 5. Oktober 2021

Berlin 2022

Hopf algebras and non-associative algebras in the study of iterated-integral signatures and rough paths

submitted by
M. Sc.
Rosa Lili Dora Preiß
ORCID:0000-0002-6145-4896

Faculty II - Mathematics and Natural Sciences
Technische Universität Berlin
for acquiring the academic degree

Doctorate in Natural Sciences
- Dr. rer. nat. -

accepted dissertation

Doctoral committee:

Chair: Prof. Dr. Michael Joswig
Reviewer: Prof. Dr. Peter K. Friz
Reviewer: Prof. Dr. Terry J. Lyons
Reviewer: Prof. Dr. Kurusch Ebrahimi-Fard

Date of defense: October 5 2021

Berlin 2022

to the joy of queer life

Published September 15 2022
at Universitätsbibliothek der Technischen Universität Berlin
stored in the repository `https://depositonce.tu-berlin.de/`
`doi:10.14279/depositonce-16256`

Zusammenfassung in deutscher Sprache

In der Gesamtschau von drei verschiedenen gemeinschaftlichen Projekten sammeln wir Erkenntnisse wie Hopf, Lie und pre-Lie, Zinbiel und dendriform, sowie Tortkara Algebren in der systematischen kombinatorischen Betrachtung von iterierten Integral Signaturen (engl. *iterated-integral signatures*) und rauen Pfaden (engl. *rough paths*) auftreten und diese weiterentwickeln.

Erstens untersuchen wir wie Lie und pre-Lie Strukturen von Lie Polynomen und Bäumen Hopf Algebren Homomorphismen erzeugen die dazu verwendet werden können die höheren Einträge von rauen Pfaden zu verschieben und dadurch auf eine Art zu renormieren. Wir beschreiben ein Zusammenspiel auf der Ebene von rauen Differentialgleichungen (engl. *rough differential equations, RDEs*) angetrieben von dem verschoben rauen Pfad vs dem ursprünglichen rauen Pfad, und betrachten ferner eine Bijektion zwischen diesen Verschiebungen-Renormierungen und einer Renormierungsgruppe einer zugehörigen Regularitätsstruktur (engl. *regularity structure*).

Zweitens beantworten wir eine Frage von Bernd Sturmfels dazu wie sich die Signatur eines Pfades $p(X)$ aus der Signatur des ursprünglichen Pfades X berechnen lässt. Nachdem wir dies mit elementaren Methoden diskutiert haben, erklären wir wie es sich auch als Korollar eines viel allgemeineren Theorems zu Homomorphismen der halfshuffle Zinbiel Algebra vs der iterierten Integral Signatur auffassen lässt, welches unmittelbar equivalent ist zu einer klassischen halfshuffle Beziehung der Signatur.

Als letztes untersuchen wir wie die vorzeichenbehaftete Fläche (engl. *signed area*) eines zweidimensionalen Pfades und der Verbindungslinie zwischen Startpunkt und Endpunkt einer algebraischen antikommutativen area Operation entspricht die die Tortkara Identität erfüllt. Wir behandeln erneut die Arbeiten von Rocha zu Koordinaten der ersten Art (engl. *coordinates of the first kind*), die ihn veranlasst haben, eine solche area Operation zum ersten Mal einzuführen, eine Theorie die man mittels einer dendriform Algebra und ihrer kanonischen pre-Lie, symmetrisierten pre-Lie, Lie und assoziativen Operationen beschreiben kann. Während unser Hauptresultat dieses Projektes aus der Aussage besteht, dass die ganze shuffle Algebra von shuffle Polynomen von area Polynomen linear erzeugt wird, eine Erkenntnis die eine Vermutung von Lyons positiv beantwortet, dass die Kenntnis aller areas of areas ausreicht um die Signatur eines Pfades zu berechnen, erhalten wir zusätzlich Anwendungen im Bereich stückweise linearer Pfade, wobei areas of areas denjenigen Signatur Komponenten entsprechen, die garantiert wieder stückweise linear sind, und im Bereich der stochastischen Analysis, wo wir Martingaloide (engl. *martingaloids*), raue Pfade deren erwartete Signatur (engl. *expected signature*) auf den areas of areas verschwindet, als Verallgemeinerung von kontinuierlichen Martingalen und stückweise linearen Interpolationen von zeitdiskreten Martingalen mit endlicher erwarteter Signatur einführen.

Abstract

Over the course of three different collaborative projects, we gather evidence of how Hopf, Lie and pre-Lie, Zinbiel and dendriform, as well as Tortkara algebras appear in and influence the systematic combinatorial treatment of iterated-integral signatures of paths and rough paths.

First, we investigate how Lie and pre-Lie structures of Lie polynomials and trees give rise to Hopf algebra homomorphisms which one can use to translate the higher orders of rough paths, and thus in a sense renormalize them. We obtain an interplay at the level of the rough differential equations (RDEs) driven by the translated rough path vs the original rough path, and furthermore explore how this translation-renormalization is in bijection with a renormalization group of a corresponding regularity structure.

Secondly, we answer a question by Bernd Sturmfels of how the signature of a path under a polynomial map $p(X)$ can be retrieved from the signature of the original path X. After we discussed this with elementary means, we explain how this can be seen as a corollary of a much more general statement on homomorphisms on the halfshuffle Zinbiel algebra vs the iterated-integral signature, which can be seen as being immediately equivalent to the classic halfshuffle relation of the signature.

Finally, we study how the signed area enclosed by a two-dimensional path and the connection line between starting point and end point corresponds to an algebraic anticommutative area operation satisfying the Tortkara identity. We revisit the work of Rocha on coordinates of the first kind which led him to introduce such an area operation for the first time, work which can be formulated in terms of a dendriform algebra and the pre-Lie, symmetrized pre-Lie, Lie and associative operations it canonically induces. With our main result in this project being the fact that the whole shuffle algebra can be expressed in terms of shuffle polynomials of area polynomials, which answers a conjecture by Lyons that the knowledge of all areas of areas suffices to compute any arbitrary signature component, we furthermore obtain applications in terms of piecewise linear paths, where areas of areas correspond to those signature components that are guaranteed to be piecewise linear again, and in terms of stochastic analysis, where we introduce martingaloids, rough paths whose expected signature vanishes on the areas of areas, as a generalization of continuous martingales and piecewise linear interpolations of time-discrete martingales with finite expected signature.

Pre-Publications

The following parts of this thesis have already been published before:

Section 1.5.3 is based on: [Pre21] Rosa Preiß. Areas of areas generate the shuffle algebra. Extended abstract, based on joint work with Joscha Diehl, Terry Lyons and Jeremy Reizenstein. In *Report No. 40/2020. New Directions in Rough Path Theory (online meeting). Organized by Thomas Cass, London; Dan Crisan, London; Peter Friz, Berlin; Massimiliano Gubinelli, Bonn. 6 December – 12 December 2020.* Mathematisches Forschungsinstitut Oberwolfach, April 2021. doi:10.14760/OWR-2020-40.

Chapter 2 is based on the joint publication [BCFP19]. Yvain Bruned, Ilya Chevyrev, Peter K. Friz, and Rosa Preiß. A rough path perspective on renormalisation. *Journal of Functional Analysis*, 277(11), December 2019.

Chapter 3 is based on the joint publication [CP20] Laura Colmenarejo and Rosa Preiß. Signatures of paths transformed by polynomial maps. *Beiträge zur Algebra und Geometrie/Contributions to Algebra and Geometry*, 61(4):695–717, 2020.

Chapter 4 is based on the joint pre-print [DLPR21]. Joscha Diehl, Terry Lyons, Rosa Preiß, and Jeremy Reizenstein. Areas of areas generate the shuffle algebra, 2021. arXiv:2002.02338v2.

The submitted version of this thesis from August 12, 2021 can be found under https://www.rosapreiss.net/files/RosaPreissSubmittedDissertation.pdf.

License statements:

Contents

About the author

Rosa Lili Dora Preiß started her academic education in 2010 with Bachelor's studies in *Mathematical Physics* at the Julius-Maximilians-Universität Würzburg, which she concluded 2014 with a Bachelor's thesis on *The hydrodynamic limit of a symmetric simple exclusion process* under the supervision of Christian Klingenberg. She then proceeded with Master's studies in Mathematics at the Technische Universität Berlin, with a focus on functional analysis and stochastics, finishing 2016 with a Master's thesis on *From Hopf algebras to rough paths and regularity structures* [Pre16] supervised by Peter K. Friz and Sylvie Paycha.

ORCID: `https://orcid.org/0000-0002-6145-4896`
Website: `https://www.rosapreiss.net`
CV: `https://www.rosapreiss.net/files/RosaPreissCV.pdf`

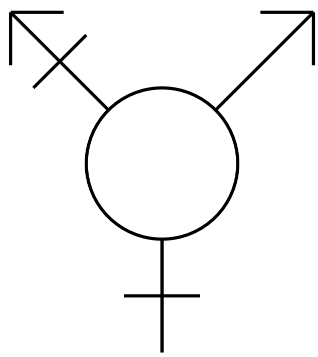


Figure 1: Trans symbol, `tikz` drawn, modern form similar to the Unicode symbol `U+26A7`, based on the original design by Holly Boswell, Wendy Parker and Nancy R. Nangeroni from the 1990s, see `http://www.gendertalk.com/tg-symbol/`

Acknowledgements

First of all, I want to thank my master's and PhD thesis supervisor Peter K. Friz for supporting me like no other in establishing my scientific career and my network of academic contacts, for introducing me to the topic of rough paths and regularity structures, and for all the encouragement in times when I unnecessarily questioned myself.

I am furthermore very grateful to my master's thesis co-supervisor Sylvie Paycha for many inspiring discussions and giving me the opportunity to speak at her research seminar multiple times, and furthermore thank her for highlighting typos in the submitted version of this thesis. I also want to thank the supervisor of my Bachelor's thesis Christian Klingenberg as well as the theoretical physicist Haye Hinrichsen for guidance and many encouraging and motivating discussions during my Bachelor's studies. I thank my fellow PhD student Tom Klose for being a wonderful office mate and for working together so well as teaching assistants in 2018. I want to thank Bernd Sturmfels for giving the motivating question for the project [CP20], for connecting me with my coauthor Laura Colmenarejo and many other researchers, and for establishing my affiliation with MPI MiS Leipzig in autumn 2018. I am grateful to Francesco Galuppi for helping me understand some basics about algebraic geometry, for our discussions on further intersections with iterated integrals and rough paths, and for his helpful suggestions regarding [CP20]. I thank Laura Colmenarejo for really pushing our joint project [CP20] towards publication. I am thankful to Lorenzo Zambotti for a short research stay in September 2016 in Paris at the beginning of my PhD studies, where I had the great opportunity to have a look at a near finished still unpublished pre-print version of [BHZ19] and also got to know Henri Elad Altman and Carlo Bellingeri. I am grateful to Kurusch Ebrahimi-Fard, Dominique Manchon and Hans Munthe-Kaas for organizing the *Workshop on Algebraic Renormalisation of Regularity Structures* discussing [BHZ19] at the University of Bergen in 2017, a wonderful exchange which opened up several important scientific contacts for me. I especially want to thank my coauthor Joscha Diehl not only for an incredibly productive ongoing collaboration, but also support regarding job applications and lots of guidance in general. I furthermore want to thank my coauthor Jeremy Reizenstein for our ongoing collaboration and for having a look at an early version of this thesis. I am also grateful to my coauthors Terry J. Lyons, Yvain Bruned and Ilya Chevyrev as well as my current collaborators Cristopher Salvi, Nikolas Tapia and Michael Ruddy for great work together and many interesting discussions and insights. I thank Mateusz Michałek for the arguments for the proof of Lemma 2.2.1 and Alexander Schmeding for many interesting discussions on topology, infinite dimensional Lie groups and generalized rough paths. I want to thank Yizheng Yuan for giving me a chance to have a look at his thesis [Yiz22] after his handing in. I am incredibly grateful to Jean Downes for helping so much with all the bureaucracy any academic position comes with. I want to thank the reviewers Peter K. Friz, Terry J. Lyons and Kurusch Ebrahimi-Fard for their feedback, and them as well as the chair Michael Joswig for taking the time to be part of the doctoral committee.

Furthermore, I am very happy to have been given the chance to meet, get to know and discuss with so many more scientists during my PhD studies, including but not limited to my first office mate Mario Maurelli, Carlos Améndola, Christian Bayer, Luigi Amedeo Bianchi, Youness Boutaib, Oleg Butkovsky, Giuseppe Cannizzaro, Pierre Clavier, Michele Coghi, Charles Curry, Jean-Dominique Deuschel, Ana Djurdjevac, Paul Hager, Antoine Hocquet, Helena Kremp, Khoa Lê, Torstein Nilssen, Tal Ohrenshtein, Gregor Pasemann, Frédéric Patras, Nicolas Perkowski, Tommaso Rosati, Michael Scheutzow, Anna Seigal, Martin Slowik and Lukas Wessels.

Last but definitely not least, I want to thank my partner Adriane Pfeifer for supporting and encouraging me throughout all the ups and downs of the last four years.

The author of this thesis was throughout much of her doctorial studies supported by European Research Council through CoG-683164. She was affiliated to Max Planck Institute for Mathematics in the Sciences, Leipzig, in autumn 2018.

Chapter 1

Introduction

1.1 From operations on path space to the iterated-integral signature

Just as one understands vectors as objects which can be added and scaled, with addition and scalar multiplication satisfying certain identities, we want to understand paths as objects which can be combined and modified by certain operations. If we look at the space of continuous rectifiable paths $\mathsf{BV}(\mathbb{R}^d)$, also called bounded variation paths, i.e. continuous paths $X : [0,T_X] \to \mathbb{R}^d$ with a finite length, we can come up with the following natural operations:

1. Concatenation, $\sqcup : \mathsf{BV}(\mathbb{R}^d) \times \mathsf{BV}(\mathbb{R}^d) \to \mathsf{BV}(\mathbb{R}^d)$,
   $(X \sqcup Y)(t) := \begin{cases} X(t), & t \leq T_X, \\ Y(t-T_X) - Y(0) + X(T_X), & T_X < t \leq T_X + T_Y, \end{cases}$
   (cf. [Che54, Section 1, page 503]), i.e. putting two paths X and Y after each other by translating Y in such a way that the starting point of Y lands on the endpoint of X,
2. time inversion, $\overleftarrow{\mathrm{D}} : \mathsf{BV}(\mathbb{R}^d) \to \mathsf{BV}(\mathbb{R}^d)$, $\overleftarrow{X} := (\overleftarrow{\mathrm{D}}\, X)(t) := X(T_X - t)$,
3. reparametrization, $\mathfrak{p}_\alpha : \mathsf{BV}(\mathbb{R}^d) \to \mathsf{BV}(\mathbb{R}^d)$, $(\mathfrak{p}_\alpha X)(t) := X(\alpha(t/T_X)T_X)$, where $\alpha : [0,s] \to [0,1]$ is a continuous bijection,
4. translation, $\mathfrak{t}_z : \mathsf{BV}(\mathbb{R}^d) \to \mathsf{BV}(\mathbb{R}^d)$, $(\mathfrak{t}_z X)(t) := z + X(t)$, where $z \in \mathbb{R}^d$,
5. rescaling, $\delta_\lambda : \mathsf{BV}(\mathbb{R}^d) \to \mathsf{BV}(\mathbb{R}^d)$, $(\delta_\lambda X)(t) := \lambda X(t)$, where $\lambda \in \mathbb{R}$,
6. linear transformation, $A : \mathsf{BV}(\mathbb{R}^d) \to \mathsf{BV}(\mathbb{R}^d)$, $(AX)(t) := AX(t)$, where $A \in \mathsf{GL}(\mathbb{R}^d)$.

Now while a vector in a separable Banach space can be stored digitally up to arbitrary precision, in a way that respects both addition and scalar multiplication, by choosing a Schauder basis, one may ask for a way to store a path as a series of numbers which is likewise compatible with our operations 1-5. One may of course interpret paths just as vectors in a separable Banach function space again, but this in no way reflects concatenation and reparametrization. Another strategy, which is indeed quite important, is to approximate through piecewise linear paths, but the author of this thesis can't think of a way to do this which at the same time respects concatenation, time inversion and reparametrization, though she doesn't have a proof for this no-go statement. The best reparametrization invariant way to piecewise linearly approximate a rectifiable path that comes to mind would be to parametrize by path length, take values equidistantly at a fixed small time step Δt and then reduce the time step e.g. dyadically $\frac{1}{2^n}\Delta_t$, with the endpoint as the last timestep, but concatenation or time inversion would not respect the structure of the equidistant time increments and there is no way to generalize this to lower regular paths with infinite variation.

In fact, the solution to this question being central to this thesis will be invariant to both translation and reparametrization, while it respects concatenation and time inversion in the following way: We are going to introduce a map $\sigma : \mathsf{BV}(\mathbb{R}^d) \to G$, where G is a group with product $\bullet$, such that, as said,

$$\sigma(\mathfrak{t}_z X) = \sigma(X) = \sigma(\mathfrak{p}_\alpha X)$$

and such that it is a (semigroup) homomorphism of concatenation, with the group inverse corresponding to time inversion:

$$\sigma(X \sqcup Y) = \sigma(X) \bullet \sigma(Y), \quad \sigma(\overleftarrow{\mathrm{D}}\, X) = \sigma(X)^{-1}. \tag{1.1}$$

From this, one can see that e.g. $\sigma(X \sqcup \overleftarrow{X})$ will be equal to the group unit element for all paths X, and thus the map σ will also not be injective up to translation and reparametrization, i.e. the equivalence classes given by $\sigma(X) = \sigma(Y)$ will be even larger. However, it will turn out to still be sufficiently small for σ to provide a powerful representation of paths as group elements.

To come back to a series of real numbers, the group G will turn out to be a subset of the dual space of an infinite dimensional vector space. The numbers which store $\sigma(X)$ will then be given as $\langle \sigma(X), w_i \rangle$, where $(w_i)_i$ is a basis.

The final remark which leads to the actual definition of σ is that each component $\langle \sigma(\cdot), w \rangle$ can be seen as a function of time again, i.e. $t \mapsto \langle \sigma(X\restriction_{[0,t]}), w \rangle$. Now kind of the fundamental construction principle of the homomorphism σ is that each Stieltjes-Integral of one of its components against another is again a component of σ, i.e. for each a, b there is a c such that

$$\int_0^T \langle \sigma(X\restriction_{[0,t]}), a \rangle \,\mathrm{d}\langle \sigma(X\restriction_{[0,t]}), b \rangle = \langle \sigma(X), c \rangle, \tag{1.2}$$

for all bounded variation X, where t is the integration variable.

We consider the *iterated-integral signature* $\sigma(X)$ of a path X as an element of $T((\mathbb{R}^d))$, the dual space of the space of words $T(\mathbb{R}^d)$, i.e.

$$\langle \sigma(X), \mathtt{i}_1 \cdots \mathtt{i}_n \rangle = \int_0^T \int_0^{r_n} \cdots \int_0^{r_2} \mathrm{d}X^{i_1}_{r_1} \dots \mathrm{d}X^{i_n}_{r_n}.$$

This means that the signature is an infinite series of numbers, assigning the real value of an iterated integral to any word in $T(\mathbb{R}^d)$. This construction was first introduced by Kuo-Tsai Chen in 1954 [Che54, Theorem 3.1] under the name *exponential homomorphism*[Che54, Section 3][1].

The definition of the signature is recursive, as the following equivalent formulation demonstrates:

$$\langle \sigma(X), \mathbf{e} \rangle = 1, \quad \langle \sigma(X), w\mathtt{i} \rangle = \int_0^T \langle \sigma(X\restriction_{[0,t]}), w \rangle \,\mathrm{d}X^i_t$$

Now, let us check how the signature does exhibit the properties we discussed before.

Translation invariance is clear due to the fact that Stieltjes integrals are translation invariant in the function which one integrates against.

[1] Chen's exponential homomorphism ([Che54, Theorem 3.1] was technically not defined on path space directly though, but on equivalence classes of paths under the coarsest equivalence relation $\equiv$ such that $X \sqcup Y \sqcup \overleftarrow{Y} \sqcup Z \equiv X \sqcup Z$ for all paths X, Y, Z.

Reparametrization invariance follows from the change of variables formula for Stieltjes integration (e.g. [Apo74, Theorem 7.7]), which states that if for some continuous bijection h, f is Stieltjes integrable against g on $[h(s), h(t)]$, then $f \circ h$ is Stieltjes integrable against $g \circ h$ on $[s,t]$ with

$$\int_s^t f(h(r))\,\mathrm{d}g(h(r)) = \int_{h(s)}^{h(t)} f(r)\,\mathrm{d}g(r).$$

The compatibility with rescaling is a graded one, i.e. for all words w we have

$$\langle \sigma(\delta_\lambda X), w\rangle = \lambda^{|w|}\langle \sigma(X), w\rangle,$$

were $|w|$ denotes the length of the word w, i.e. the number of letters in w. This is a simple consequence of the fact that Stieltjes-Integrals are bilinear, and thus iterated integrals are multilinear.

As for linear transformation, we have [Che57, Theorem 3.1]

$$\langle \sigma(AX), a\rangle = \langle \sigma(X), M_{A^\top} a\rangle,$$

where $M_{A^\top}$ is recursively given as $M_{A^\top} w\mathtt{i} = (M_{A^\top} w)(A^\top \mathtt{i})$ with $M_{A^\top}\mathtt{i} = A^\top\mathtt{i}$. This also follows directly from the bilinearity of Stieltjes integrals. See [Die13], [DR19] and [DPRT22] for invariants of this map, i.e. linear combinations of iterated integrals which are invariant under the action of certain subgroups of GL on the path, something that Chen already gave examples for in [Che57, Example 2].

The first main algebraic property of the signature is *Chen's identity*, which states how the signature of a concatenation of paths can be computed from the signature of the individual paths, [Che54, Lemma 1.1, Lemma 1.2]

$$\langle \sigma(X \sqcup Y), a\rangle = \sum_{(a)}^{\bullet} \langle \sigma(X), a_1\rangle\langle \sigma(Y), a_2\rangle.$$

Here, $\sum_{(a)}^{\bullet} a_1 \otimes a_2 := \Delta_\bullet\, a$ is *Sweedler's notation* of the *coproduct* $\Delta_\bullet$ dual to the concatenation product $\bullet$ of words given by $w_1 \bullet w_2 = w_1 w_2$. I.e., the *deconcatenation* coproduct $\Delta_\bullet$ is characterized by

$$\langle \Delta_\bullet\, a, b \otimes c\rangle = \langle a, b \bullet c\rangle$$

Chen's identity can thus also be written as $\sigma(X \sqcup Y) = \sigma(X) \bullet \sigma(Y)$ [Che54, Theorem 3.1][Che57, Theorem 2.1], showing that the signature is indeed a semigroup homomorphism respecting the concatenation of paths.

Furthermore, the signature satisfies the *shuffle relation*

$$\langle \sigma(X), a\rangle\, \langle \sigma(X), b\rangle = \langle \sigma(X), a \sqcup\!\sqcup\, b\rangle\,,$$

as proven in Ree's 1958 publication [Ree58, Equation (2.5.2)][2], which is nothing but integration by parts for iterated integrals. While the formal, recursive definition of the shuffle product $\sqcup\!\sqcup$ will be given in Definition 3.2.2 and in Section 4.1.1, more intuitively it can be explained as the bilinear product which maps two words w_1, w_2 to the sum all shuffles of the letters of the two words which preserve the orders of the letters in the two individual original words, just like a riffle shuffle of two decks of cards preserves the orders of the cards in the two individual decks.

[2] Chen already showed before that the *log signature* $\log_\bullet \sigma(X)$ of a C^1 path is a Lie series [Che57, Theorem 4.2], a fact that Ree showed in [Ree58, Theorem 2.5] is equivalent to the shuffle relation, but Ree also showed the shuffle relations for iterated integrals of bounded variation paths [Ree58, Equation (2.5.2)] directly through an induction using integration by parts.

Now, it turns out that whenever two tensor series $g_1, g_2 \in T((\mathbb{R}^d))$ satisfy the shuffle relation, their concatenation $g_1 \bullet g_2$ also does so. This makes the set G of all non-zero tensor series satisfying the shuffle relation, i.e. the set of *characters* for the shuffle product, a semigroup. Since $\mathbf{e} \in G$ with $\mathbf{e} \bullet g = g \bullet \mathbf{e} = g$, it is even a monoid.

The compatibility of $\sqcup\!\sqcup$ and $\Delta_\bullet$ turn $(T(\mathbb{R}^d), \sqcup\!\sqcup, \Delta_\bullet)$ into a bialgebra. The existence of a linear map $S : T(\mathbb{R}^d) \to T(\mathbb{R}^d)$ with the property

$$\sum_{(a)}^{\bullet} S(a_1) \sqcup\!\sqcup a_2 = \sum_{(a)}^{\bullet} a_1 \sqcup\!\sqcup S(a_2) = \langle a, \mathbf{e} \rangle \, \mathbf{e}$$

for all $a \in T(\mathbb{R}^d)$, a so-called *antipode*, makes $(T(\mathbb{R}^d, \sqcup\!\sqcup, \Delta_\bullet)$ a Hopf algebra, where S is uniquely given by

$$S(\mathbf{e}) = \mathbf{e}, \quad S(\mathtt{i}_1 \cdots \mathtt{i}_n) = (-1)^n \mathtt{i}_n \cdots \mathtt{i}_1.$$

In this case, the antipode S is self-dual, and thus, using the shuffle identity for $g \in G$,

$$\langle S(g) \bullet g, a \rangle = \sum_{(a)}^{\bullet} \langle S(g), a_1 \rangle \langle g, a_2 \rangle = \sum_{(a)}^{\bullet} \langle g, S(a_1) \rangle \langle g, a_2 \rangle = \sum_{(a)}^{\bullet} \langle g, S(a_1) \sqcup\!\sqcup a_2 \rangle = \langle \mathbf{e}, a \rangle$$

for all $a \in T(\mathbb{R}^d)$, since $\langle g, \mathbf{e} \rangle = 1$ for all $g \in G$, showing that each $g \in G$ has an inverse $g^{-1} = Sg$. Thus, we indeed have that the image of σ lives inside a group. One then shows $\sigma(\overleftarrow{X}) = S(\sigma(X))$, and with that, we finally arrive at (1.1).

What remains is equation (1.2), and here, it turns out that the shuffle product can be naturally split into two parts,

$$a \sqcup\!\sqcup b = a \succ b + b \succ a,$$

where the bilinear $\succ$ is called the *halfshuffle product*, which is non-associative, but satisfies the (left) *Zinbiel identity*

$$a \succ (b \succ c) = (a \succ b + b \succ a) \succ c.$$

Now, one can actually see from the original proof of Ree's shuffle identity in [Ree58, Section 2, page 216] that we indeed have

$$\int_0^T \langle \sigma(X\restriction_{[0,t]}), a \rangle \, \mathrm{d} \langle \sigma(X\restriction_{[0,t]}), b \rangle = \langle \sigma(X), a \succ b \rangle.$$

For the formal definition of the halfshuffle we refer to Definition and Section, but the halfshuffle $w_1 \succ w_2$ of two words can be understood as those summands of the full shuffle $w_1 \sqcup\!\sqcup w_2$ that have the last letter of w_2 as their last letter. In the analogy to shuffling decks of cards, this corresponds to all the possible decks obtained from two original decks by first putting aside the lowest card of the second deck, then riffle shuffling the two decks once, and then putting back the card from the side as the lowest card of the resulting deck.

The signature uniquely determines the path up to tree-like equivalence. This was shown for piecewise C^1 paths in [Che58, Theorem 4.1], for bounded variation paths in [HL10, Theorem 4], and [BGLY16, Theorem 1.1] implies it for paths of Hölder regularity greater than 1/2. By [BGLY16, Definiton 1.1], a *tree-like path* is a path X for which there is an $\mathbb{R}$-tree τ, a continuous map $\phi : [0, T_X] \to \tau$ with $\phi(0) = \phi(T_X)$ and a map $\psi : \tau \to \mathbb{R}^d$ such that $X = \psi \circ \phi$. An $\mathbb{R}$-tree (τ, d) is a metric space such that for any two distinct points $x, y \in \tau$ there is a unique isometry $Y : [0, d(x, y)] \to \operatorname{Im} Y \subseteq \tau$ with $Y(0) = x$ and $Y(d(x, y)) = y$ (cf. [HL08, Definition 2.1]). We then say that X and Y are *tree-like equivalent*, $X \overset{\text{tree}}{\sim} Y$, iff $X \sqcup \overleftarrow{Y}$ is tree-like ([BGLY16, Section 2]). As tree-like paths form the kernel of the semigroup homomorphism

σ, the signature forms group monomorphism $\sigma : \big(\mathsf{BV}(\mathbb{R}^d)/\overset{\text{tree}}{\sim}\big) \to \mathcal{G}_d$, where elements of the quotient can be represented by the unique element of an equivalence class with the shortest length, parametrization according to length and zero as starting point. These representatives are called *reduced paths with standard parametrization*, and we call $\big(\mathsf{BV}(\mathbb{R}^d)/\overset{\text{tree}}{\sim}\big)$ the *reduced path group* (of bounded variation paths). Note that in particular $(\mathfrak{p}_\alpha X) \overset{\text{tree}}{\sim} X \overset{\text{tree}}{\sim} (\mathfrak{t}_z X)$. It is furthermore interesting to note that the reduced path group of bounded variation paths (or equivalently its image inside $\mathcal{G}_d$), together with the metric given by the distance of two reduced paths X and Y being the length of the reduced path tree-like equivalent to $\overleftarrow{X} \sqcup Y$, is itself an $\mathbb{R}$-tree (cf. the similar statement for finite p-variation rough paths [BGLY16, Proposition 4.1]). This metric does *not* give the reduced path group a topological group structure however, as the group inverse $[X] \mapsto [\overleftarrow{X}]$ is incontinuous under the R-tree metric.

Through the theory of *Young integration*, all iterated integrals of paths with finite p-variation for some $p < 2$ are well-defined as Stieltjes-integrals [You36, Theorem on Stieltjes integrability, pages 264-265]. Thus, all of the above holds true completely analogously for p-variation paths with $p < 2$, which include Hölder continuous paths of Hölder regularity greater than 1/2, only the problem of defining a representative for any tree-like equivalence class and an $\mathbb{R}$-tree-metric on the reduced path group is now much more involved.

Signatures have proven a very valuable tool in machine learning. I.e., one uses a truncated signature to store information about the path or to classify data sets consisting of several multi-dimensional time series (see e.g. [**?**, chevyrevkormilitzin2016]. Beyond using a truncated signature, also signature kernel methods have been recently proposed ([KO19, Section 3],[SCF$^+$21, Section 2 and Section 4] and the "normalized signature kernel" in [CO22, Section 6]), where the untruncated kernel defined by $k_{s,t}(X,Y) := \big\langle \sigma(X\restriction_{[0,s]}), \sigma(Y\restriction_{[0,t]})\big\rangle$ [SCF$^+$21, Definition 2.4] can be obtained numerically by numerical schemes for the partial differential equation [SCF$^+$21, Theorem 2.5]

$$\frac{\partial^2 k_{s,t}(X,Y)}{\partial s\, \partial t} = \langle \dot{X}_s, \dot{Y}_t\rangle\, k_{s,t}(X,Y),$$

with initial conditions $k_{s,0}(X,Y) = k_{0,t}(X,Y) = 1$. While this Goursat PDE is only well-defined for X, Y differentiable, a more general integral equation determining $k_{s,t}(X,Y)$ is given by [SCF$^+$21, Theorem 4.11].

Fascinating about signature theory is how it brings together so many different branches of mathematics, including rough path analysis, stochastic analysis and classical analysis, abstract and universal algebra, representation theory, invariant theory ([Die13],[DR19],[DPRT22]), (co)homology and homotopy ([Che77],[Hai02]), matrix theory, noncommutative geometry [Kap09], algebraic geometry ([AFS19],[Gal19],[CGM20]), ODE solution theory (see e.g. [BCE20, Section 3]), dynamical systems, machine learning ([CK16],[Rei19]).

1.2 Rough Paths and RDEs

The theory of rough paths carries over signature theory to lower regular paths. Moreover, it provides a solution theory for differential equations

$$dY = f\,(Y)\, dX$$

driven by low regular paths as they arise naturally in stochastic analysis, the most prominent example being Brownian motion, which is γ-Hölder continuous for any $\gamma < 1/2$. Even lower regular is fractional Brownian motion for Hurst parameter smaller than 1/2.

While stochastic integration theories due to Itô and Stratonovich provide a probabilistic solution theory for large classes of stochastic processes, rough path theory gives a deterministic solution theory once all iterated integrals of the driving path are specified. Furthermore, it can be shown that the solution map, the Itô-Lyons map, which for a fixed collection of vector fields takes the driving rough path (i.e. the path together with all its iterated integrals over all subintervals) as an input and gives the solution path as an output, is a Lipschitz continuous map provided that vector fields are in C_b^3 (see [FH14, Theorem 8.5]). In contrast, the solution map in stochastic analysis, the Itô map, is only measurable.

The modern formulation of rough paths distinguishes two classes, weakly geometric rough paths on the one hand and branched rough paths on the other hand. Weakly geometric rough paths (introduced by Lyons in [Lyo98, Definition 2.1.2, Lemma 2.1.1 and Definition 2.3.1] as *geometric p-muliplicative functionals*, i.e. character valued paths of finite p-variation) are defined analogously to signatures of paths of higher regularity, though now the shuffle identity and Chen's identity become part of the definition instead of being theorems.

Branched rough paths (introduced by Gubinelli in [Gub10, Defintion 7.2]) are designed to allow for integration theories which do not respect the usual rules of calculus, and thus violate the integration by parts condition encoded in the shuffle identity. Therefore, integrals such as

$$\int_s^T X_{sr}^1 X_{sr}^2 \, \mathrm{d}X_r^3,$$

need to be specified additionally, while in the weakly geometric setting they can be decomposed into usual iterated integrals via the shuffle relation.

Instead of words, for branched rough paths one now uses trees and forests to encode iterated integrals including products. Chen's identity, however, is part of the definition of branched rough paths, in the formulation

$$\langle \mathbf{X}_{st}, \tau \rangle = \langle \mathbf{X}_{su} \otimes \mathbf{X}_{ut}, \Delta_\star \tau \rangle,$$

where $\Delta_\star$ is the so-called *Connes-Kreimer* coproduct, which cuts down trees into all combinations of a lower part (trunk) and an upper part (crown, possibly a forest).

Uniqueness up to tree-like equivalence has been shown for the signature of weakly geometric rough paths in [BGLY16, Theorem 1.1] and for branched rough paths in [BC19, Theorem 5.1], which means that indeed signature theory extends seamlessly to the rough regime.

While geometric and branched rough paths are by now the two well-established classical notions of rough paths, in fact, one can define a class of rough paths for any commutative connected graded Hopf algebra. Important examples include *quasi-geometric rough paths* [KH13][Bel20, Definition 3.7], which are based on the Hopf algebra of a commutative quasishuffle product with the deconcatenation coproduct [Hof00, Equation (1), Theorem 2.1 and Section 3], and *planarly branched rough paths* [CEMM20, Definition 6.7] which are based on the *Munthe-Kaas-Wright Hopf algebra of Lie group integrators* [MW08, Definition 5 and Theorem 1].

1.3 Regularity structures and renormalization

While Rough Path Analysis provides a solution theory for SDEs, the theory of regularity structures, which in some sense generalizes the concept of rough paths, does so for a class of *stochastic partial differential equations* (SPDEs), most notably those of the form

$$(\partial_t - \Delta)u = F(u, \nabla u) + G(u, \nabla u)\xi, \tag{1.3}$$

where ξ is some stochastic noise, generically often white noise.

A *regularity structure*, an object which encodes the general form of an SPDE, consists of a *structure group* G continuously acting on a normed vector space $T = (T_\alpha)_{\alpha \in A}$ graded by a set of real numbers A, such that $\{\alpha \in A : \alpha < n\}$ is finite for all n, such that each T_α is a Banach space and such that for any $\Gamma \in G$ and $x \in T_\alpha$, one has $\Gamma\tau - \tau \in T_{<\alpha}$ ([Hai14, Definition 2.1]).

Furthermore, it often makes sense, for example for parabolic SPDEs like 1.3, to scale time differently than space. We thus introduce a *scaling* $\mathfrak{s} = (\mathfrak{s}_1, \ldots, \mathfrak{s}_d) \in \mathbb{N}^d$ ([Hai14, Section 2.2][3]) and, even though it is not a norm, we write

$$\|x\|_{\mathfrak{s}} = \sum_{i=1}^{d} |x|^{1/\mathfrak{s}_i},$$

where $\|x - y\|_{\mathfrak{s}}$ does define a proper metric ([Hai14, Equation (2.10)]). We also scale the test functions we use in the theory according to $\mathfrak{s}$, namely we define

$$\varphi_x^\lambda(y) := \lambda^{-|\mathfrak{s}|}\phi(\lambda^{-\mathfrak{s}_1}(x_1 - y_1), \ldots, \lambda^{-\mathfrak{s}_d}(x_d - y_d)),$$

where $|\mathfrak{s}| := \sum_{i=1}^{d} \mathfrak{s}_i$ ([Hai14, Equations (2.12) and (2.13)]).

A *model* for a regularity structure (A, T, G) over $\mathbb{R}^d$ consists then of two maps $\Pi : \mathbb{R}^d \to \mathcal{L}(T, \mathcal{S}'(\mathbb{R}^d))$ and $\Gamma : \mathbb{R}^d \times \mathbb{R}^d \to G$ such that

$$\Gamma_{xy}\Gamma_{yz} = \Gamma_{xz}, \quad \Pi_x\Gamma_{xy} = \Pi_y,$$

satisfying the analytical properties

$$\|\Gamma_{xy}\tau\|_\ell \leq C_{\mathfrak{K},m}\|x - y\|_{\mathfrak{s}}^{m-\ell}, \quad |(\Pi_x\tau)(\varphi_x^\lambda)| \leq C_{\mathfrak{K},m}\lambda^m,$$

for all compact $\mathfrak{K} \subseteq \mathbb{R}^d$ and all $m \in A$, uniformly over all $x, y \in \mathfrak{K}$, all $l < m$, all test functions φ supported on the $\|\cdot\|_{\mathfrak{s}}$ unit ball around zero with $\|\varphi\|_{C^r} \leq 1$, and all $\lambda \in (0,1]$, where r is the smallest integer being strictly greater than $|\min A|$ ([Hai14, Definition 2.17]). Each model gives then rise to linear spaces of *modelled distributions* $\mathscr{D}^\gamma$, which consist of all locally bounded maps $f : \mathbb{R}^d \to T_{<\gamma}$ such that

$$\|f_x - \Gamma_{xy}f_y\|_\ell \leq C_{\mathfrak{K}}\|x - y\|_{\mathfrak{s}}^{\gamma-\ell}$$

for all compact sets $\mathfrak{K} \subseteq \mathbb{R}^d$, uniformly over all $\ell < \gamma$ and all $x, y \in \mathfrak{K}$ ([**?**, []efinition 3.1]Hairer14). An SPDE like (1.3) can then be turned into an equation on $\mathscr{D}^\gamma$, and any solution $U \in \mathscr{D}^\gamma$ will be mapped to a distribution again via the continuous linear *reconstruction operator* $\mathcal{R} : \mathscr{D}^\gamma \to \mathcal{D}(\mathbb{R}^d)$ corresponding to the model (Π, Γ), which is characterized by

$$|(\mathcal{R}f - \Pi_x f_x)(\varphi_x^\lambda)| \leq C_{\Pi,f,\mathfrak{K}}\lambda^\gamma$$

for all compact $\mathfrak{K} \subseteq \mathbb{R}^d$, uniformly over all $x \in \mathfrak{K}$, all test functions φ supported on the $\|\cdot\|_{\mathfrak{s}}$ unit ball around zero with $\|\phi\|_{C^r} \leq 1$, and all $\lambda \in (0,1]$, where r is as before. This is (part of) the famous *reconstruction theorem*, Theorem 3.10 in [Hai14]. We then consider $u := \mathcal{R}U$ as a solution to (1.3).

These SPDEs however often turn out to be ill-posed a priori. For example, consider the famous KPZ equation (e.g. [Hai14, Equation (1.3)])

$$(\partial_t - \partial_x^2)u = (\partial_x u)^2 + \xi,$$

[3]In [Hai14], there is the additional assumption of the $\mathfrak{s}_i$ being relatively prime, but even though this seems very reasonable, the author of this thesis doesn't see a necessity to impose that, since everything works in the more general case.

where we have a problem when ξ is a distribution of Hölder regularity smaller than -1, like e.g. space-time white noise in one space dimension is of Hölder regularity just under $-3/2$, since the derivative $\partial_x u$ is then expected to be of negative regularity as well (just under $-1/2$ in the white noise case), but the square $(\partial_x u)^2$ isn't well defined for proper distributions of negative regularity. If we now approximate ξ with a sequence of smooth functions ξ^ϵ, for example via a mollifier, then the corresponding models $(\Pi^\epsilon, \Gamma^\epsilon)_{1\geq\epsilon>0}$ appropriate for solving the equation will fail to converge for $\epsilon \to 0$, for the exact same reason: because the regularity structure contains an abstract symbol for e.g. $(\partial_x K * \xi)^2$ and the evaluation with Π_x will then diverge, where K is the *heat kernel* for the operator $\partial_t - \partial_x^2$.

The way to deal with this in the framework of regularity structures is by adding another group R continuously acting on T and another grading $T = (\tilde{T}_\beta)_{\beta\in B}$, where $\{\beta \in B : \beta < n\}$ is finite for all n, such that, at least in the framework developed by Bruned-Hairer-Zambotti [BHZ19, Sections 4 to 6], R leaves G invariant in the sense that $\Upsilon^{-1}\Gamma\Upsilon \in G$ for all $\Gamma \in G$ and $\Upsilon \in R$ (cf. [BHZ19, Theorem 6.16]), and now satisfies $\Upsilon\tau \in T_{\geq\alpha} \cap T_{\leq m(\alpha)}$ and $\Upsilon\varrho - \varrho \in \tilde{T}_{>\beta}$ for all $\tau \in T_\alpha, \varrho \in \tilde{T}_\beta$ and some function $m : A \to A$. [4]

This group R is called *renormalization group*, and one chooses an appropriate $(\Upsilon^\epsilon)_{1\geq\epsilon>0}$ such that the renormalized models $(\bar{\Pi}^\epsilon, \bar{\Gamma}^\epsilon)$ converge for $\epsilon \to 0$, where (cf. [**?**, []heorem 6.16 and Equation (6.13)]BHZ16)

$$\bar{\Pi}^\epsilon_x := \Pi^\epsilon_x \Upsilon_\epsilon, \quad \bar{\Gamma}^\epsilon_{xy} := \Upsilon_\epsilon^{-1} \Gamma^\epsilon_{xy} \Upsilon_\epsilon.$$

The limiting model $(\bar{\Pi}, \bar{\Gamma})$ is then used to solve the equation.

To make this work with a renormalization group in the above sense, one might in many cases first need to extend the regularity structure (A, T, G) to $(A, T^{\mathrm{ex}}, G^{\mathrm{ex}})$, where each T_α is a subspace of T^{ex}_α and G is a subgroup of G^{ex}, just as it was done in great generality in [BHZ19], see Section 6.4 there for the relation of "reduced" and "extended" regularity structure.

1.4 Universality of words and trees as algebraic combinatorial objects

1.4.1 The tensor algebra

The universality of the tensor algebra $T(V) = \mathbb{R} \oplus \bigoplus_{n=1}^\infty V^{\bullet n}$, written as the space of linear combinations of words (non-commutative polynomials) built from the alphabet which represents a linear basis of V (existing due to Zorn's Lemma, assuming the axiom of choice) together with the empty word $\mathbf{e}$, is mostly grounded on its non-unital part $T^{\geq 1}(V) = \bigoplus_{n=1}^\infty V^{\bullet n}$ being both the *free associative algebra* with the word concatenation product $\bullet$ and the *free Zinbiel algebra* with the halfshuffle product $\succ$. It furthermore constitutes the *free Leibniz algebra* with the Leibniz bracket given recursively by [LP93, Lemma (1.3)]

$$[\![\mathtt{i}, x]\!] = \mathtt{i}x, \quad [\![\mathtt{i}w, x]\!] = \mathtt{i} \bullet [\![w, x]\!] - [\![w, \mathtt{i}x]\!],$$

which can be explicitly written as

$$[\![x, y]\!] = r(x) \bullet y,$$

where r is the Dynkin map discussed in Section 4.2, however we will not be concerned with Leibniz algebras in this thesis, we only mention that the Leibniz operad is the *Koszul dual* (see

[4] If one follows exactly the construction of [BHZ19, Sections 4 to 6], then one gets $RT_\alpha = T_\alpha$ However, it is also possible and seems reasonable to tweak their grading $|\cdot|_+$ [BHZ19, Definition 5.3] a little, such that all canonical models and their renormalizations are still models, but such that we have $(\Upsilon - \mathrm{id})T_\alpha \subseteq T_{>\alpha}$ for all $\Upsilon \in R$. This makes the space of all models a little smaller, but it encodes the demand that renormalization of terms shall happen via adding terms of higher degree.

e.g. [HP11, Definition 10.5]) of the Zinbiel operad, which is the reason Zinbiel algebras have been named like that in the first place (Leibniz spelled backwards).

From the free associative algebra $(T^{\geq 1}(V), \bullet)$ we may now of course derive by antisymmetrization a Lie algebra on $T^{\geq 1}(V)$ given by the Lie bracket $[x, y] := x \bullet y - y \bullet x$, and it turns out that the smallest sub-Lie-Algebra $(\mathfrak{g}(V), [\cdot, \cdot])$, $\mathfrak{g}(V) \subsetneq T^{\geq 1}(V)$ containing the letters $\mathtt{i}$ constitutes the *free Lie algebra* over V (see e.g. [Reu93, Sections 0.2 and 1.2]). If we look at the algebra $(T^{\geq 1}(V), ⧢)$, where ⧢ is the symmetrization of the halfshuffle $\succ$, it turns out to be free as a commutative associative algebra over any homogeneous minimal generating subspace (see Section 4.4 and cf. [Reu93, Corollary 5.5, Theorem 6.1 and Section 6.5.1]). We may furthermore extend both ⧢ and $\bullet$ to $T(V)$ by setting

$$x ⧢ \mathbf{e} = \mathbf{e} ⧢ x = x, \quad x \bullet \mathbf{e} = \mathbf{e} \bullet x = x$$

to obtain a free unital commutative associative algebra $(T(V), ⧢)$ and the free unital associative algebra $(T(V), \bullet)$ over V.

We introduce a duality pairing of $T(V)$ with itself by setting

$$\langle w, v \rangle = \delta_{w,v}$$

for all words w, v. What is now the impressing interplay of the free Zinbiel algebra and the free associative algebra is that if we form the dual unshuffle coproduct $\Delta_{⧢}$ of the shuffle product ⧢, we obtain a Hopf algebra $(T(V), \bullet, \Delta_{⧢}, \alpha)$ which is connectedly graded by word length and is nothing but the universal enveloping algebra of the free Lie algebra $\mathfrak{g}(V)$. This in particular means that $\Delta_{⧢}$ is the due to freeness of the tensor algebra unique homomorphism from $T(V)$ to $T(V) \otimes T(V)$ satisfying $\Delta_{⧢}\mathtt{i} = \mathtt{i} \otimes \mathbf{e} + \mathbf{e} \otimes \mathtt{i}$, i.e. such that the letters are primitive elements. It furthermore means that the Lie algebra $\mathfrak{g}(V)$ is characterized as the set of primitive elements in $T(V)$, i.e. the set of all linear combinations of words x such that

$$\Delta_{⧢} x = x \otimes \mathbf{e} + \mathbf{e} \otimes x.$$

This unique and fascinating interplay of the two universal algebraic structures $(T^{\geq 1}(V), \succ)$ and $(T^{\geq 1}(V), \bullet)$ is more than enough motivation from an algebraic combinatorial viewpoint alone to look at the tensor product of $\succ$ and $\bullet$, which is described by the two operations $\succeq$ and $\preceq$ (see Section 4.1.2.1 and Remark 4.2.5) forming a dendriform algebra, and at the antisymmetrization of $\succ$, which we will denote by area.

For the latter, thanks to [DIM19, Theorem 2.5], we have another universal property, namely that for any two letters $\mathtt{i}, \mathtt{j}$, the smallest subspace containing both of these letters and being closed under the area operation does in fact form the free Tortkara algebra over the two-dimensional space spanned by $\mathtt{i}, \mathtt{j}$. It remains a very important to solve and challenging open problem (Conjecture 4.6.5)whether $(\mathscr{A}(V), \mathsf{area})$, the smallest subspace containing all the letters and being closed under area, is actually always the free Tortkara algebra over V, for any dimension of V.

Finally, just for completeness, we mention that the symmetrization of the free associative algebra yields a Jordan algebra $(T(V), \mathsf{jor})$ (as the anticommutator of an associative algebra is always a Jordan algebra) and the smallest subspace $\mathscr{SJ}(V)$ containing the letters and being closed under jor constitutes the free Jordan algebra over V for V two-dimensional, but not for higher dimension (see e.g. [McC78], cf. also [KM21] and the remarks in [Dzh07]).

1.4.1.1 A note on different naming conventions: Zinbiel/Dendriform vs shuffle

We want to highlight that in the algebraic literature relevant for this thesis, two different naming conventions have evolved for the categories we, together with some part of the scholars, denote as

Zinbiel algebras, dendriform algebras and (commutative) tridendriform algebras (cf. the almost identical "dictionary" given in [FP20, Section 1, Terminology]):

Loday et al, e.g. [Lod01] [Lod07] [Dzh07] [LV12] [BBM17]	Foissy-Patras et al, e.g. [EP15] [FP20] [EP21] [DET20a]
Zinbiel algebra / commutative dendriform algebra	commutative shuffle algebra
dendriform algebra	(noncommutative) shuffle algebra
commutative tridendriform algebra	commutative quasishuffle algebra
tridendriform algebra	(noncommutative) quasishuffle algebra

While (commutative) shuffle algebra and (commutative) quasishuffle algebra denote categories of algebras coming from certain quadratic binary operads for scholars following the Foissy-Patras et al convention, for us a (half)shuffle algebra denotes the free Zinbiel algebra $(T^{\geq 1}(V), \succ)$ over some vector space V and a *quasishuffle algebra* (following [Lod07, Section 1.5][5]) denotes a tridendriform algebra of the form $(T^{\geq 1}(A), \preccurlyeq, \succcurlyeq, \bowtie)$ built from an associative algebra $(A, \{\cdot,\cdot\})$ with

$$x\mathtt{i} \preccurlyeq y\mathtt{j} = (x \mathbin{\widehat{\sqcup\!\sqcup}} y\mathtt{j}) \bullet \mathtt{i}, \quad x\mathtt{i} \succcurlyeq y\mathtt{j} = (x\mathtt{i} \mathbin{\widehat{\sqcup\!\sqcup}} y) \bullet \mathtt{j}, \quad x\mathtt{i} \bowtie y\mathtt{j} = (x \mathbin{\widehat{\sqcup\!\sqcup}} y) \bullet \{\mathtt{i}, \mathtt{j}\},$$
$$x\mathtt{i} \mathbin{\widehat{\sqcup\!\sqcup}} y\mathtt{j} = x\mathtt{i} \preccurlyeq y\mathtt{j} + x\mathtt{i} \succcurlyeq y\mathtt{j} + x\mathtt{i} \bowtie y\mathtt{j}, \quad x \mathbin{\widehat{\sqcup\!\sqcup}} \mathbf{e} = \mathbf{e} \mathbin{\widehat{\sqcup\!\sqcup}} x = x$$

for $x, y \in T(A)$ and $(\mathtt{i})_{\mathtt{i}\in I}$ denoting a linear basis of A. If $\{\cdot,\cdot\}$ is commutative, then $x \succcurlyeq y = y \preccurlyeq x$ and $\widehat{\sqcup\!\sqcup}$ is commutative, thus $(T^{\geq 1}(A), \succcurlyeq, \bowtie)$ is a commutative tridendriform algebra. Thus, we denote by *commutative quasishuffle algebras* all commutative tridendriform algebras of the form $(T^{\geq 1}(C), \succcurlyeq, \bowtie)$ for an arbitrary commutative associative algebra $(C, \{\cdot,\cdot\})$.

This constitutes our main argument for sticking to the Loday et al naming convention for the binary quadratic operadic algebra categories: To reserve the names (half)shuffle algebra, quasishuffle algebra and commutative quasishuffle algebra for the objects in the image of the functors[6]

$$V \mapsto (T^{\geq 1}(V), \succ), \quad (A, \{\cdot,\cdot\}) \mapsto (T^{\geq 1}(A), \preccurlyeq, \succcurlyeq, \bowtie), \quad (C, \{\cdot,\cdot\}) \mapsto (T^{\geq 1}(C), \succcurlyeq, \bowtie)$$

since those functorial images form strict subcategories of what we call Zinbiel algebras, tridendriform algebras and commutative tridendriform algebras.

1.4.2 Trees and Forests

Labeled non-planar rooted trees appear naturally as a basis of the free pre-Lie algebra $(\mathcal{B}(V), \curvearrowright)$ over the vector space V (see [CL01]), where once again a basis $(v_i)_{i\in I}$ of V is chosen to serve now as the set of possible labels $i \in I$ of the trees, i.e. V embeds into $\mathcal{B}(V)$ as the span of singleton nodes $(\bullet_i)_{i\in I}$. As any pre-Lie algebra induces a Lie algebra by antisymmetrization, we obtain in this case the Lie algebra $(\mathcal{B}(V), [\cdot,\cdot]_\star)$.

Finally, we form the free unital commutative associative algebra over $\mathcal{B}(V)$, linearly spanned by commutative polynomials formed from the trees in $\mathcal{B}(V)$, which we call *forests*, and an *empty forest* $\mathfrak{e}$, and we denote this algebra by $\mathcal{H}(V)$ with *forest product* $\odot$. We furthermore introduce a linear duality pairing of $\mathcal{H}(V)$ with itself characterized by

$$\langle \zeta_1, \zeta_2 \rangle = \delta_{\zeta_1, \zeta_2}$$

[5]though we 'mirror' Loday's construction and for simplicity define the product with $\mathbf{e}$ only for $\widehat{\sqcup\!\sqcup}$

[6]to fully define those functors, one must of course furthermore specify how they lift linear maps, algebra homomorphisms and commutative algebra homomorphisms to Zinbiel homomorphisms, tridendriform homomorphisms and commutative tridendriform homomorphisms

for all forests ζ_1, ζ_2. Then, we obtain that there is an associative non-commutative product $\star$ on $\mathcal{H}(V)$ graded by the number of nodes a tree has such that the Hopf algebra $(\mathcal{H}(V), \star, \Delta_\odot)$, where the coproduct $\Delta_\odot$ is the dual of $\odot$, forms the universal enveloping algebra of $(\mathcal{B}(V), [\cdot, \cdot]_\star)$, which in particular means that

$$[\tau_1, \tau_2]_\star = \tau_1 \curvearrowright \tau_2 - \tau_2 \curvearrowright \tau_1 = \tau_1 \star \tau_2 - \tau_2 \star \tau_1$$

for all $\tau_1, \tau_2 \in \mathcal{B}(V)$ and that we have that $\mathcal{B}(V)$ consists of all $x \in \mathcal{H}(V)$ such that

$$\Delta_\odot x = x \otimes \mathfrak{e} + \mathfrak{e} \otimes x.$$

This unique product $\star$ is called the *Grossman-Larson product* [GL89]. The free pre-Lie product $\curvearrowright$ then admits a simple representation in terms of the Grossman-Larson product:

$$\tau_1 \curvearrowright \tau_2 = \mathsf{proj}_{\mathcal{B}(V)}(\tau_1 \star \tau_2).$$

It can be shown that there exists a subspace $\mathfrak{T} \subsetneq \mathcal{B}(V)$, whose choice is highly non-unique, such that $\mathcal{B}(V)$ is the free Lie algebra over $\mathfrak{T}$, and equivalently such that $\mathcal{H}(V)$ is the free associative algebra over $\mathfrak{T}$.

Now, of course we can look at the graded dual Hopf algebra $(\mathcal{H}(V), \odot, \Delta_\star)$. Here, the *Butcher-Connes-Kreimer coproduct* $\Delta_\star$, the dual of the universal enveloping algebra product $\star$, is characterized by multiplicativity with respect to $\odot$ and by turning the crafting of a forest ζ onto a new root $\bullet_i$, which we denote by $B_i^+(\zeta) = \lfloor \zeta \rfloor_i$, into a 1-cocycle, i.e.

$$\Delta_\star B_i^+ = (\mathrm{id} \otimes B_i^+)\Delta_\star + (B_i^+ \otimes \mathfrak{e}).$$

Now, we obtain the following universality (see [CK98, Theorem 2 of Section 3], [Foi13, Theorem 3], [HK15, Remark 4.7]): If A is any Hopf algebra with a V-collection of 1-cocycles $L_v : A \to A$, $L_{\beta_1 x + \beta_2 y} = \beta_1 L_x + \beta_2 L_y$, then the unique associative algebra homomorphism $\psi : \mathcal{H}(V) \to A$ with $\psi B_v^+ = L_v \psi$, where $B_{v_i}^+ := B_i^+$, is in fact a Hopf algebra homomorphism. I.e. $(\mathcal{H}(V), \odot, \Delta_\star, B^+)$ is an initial object in the category of Hopf algebras with a V-collection of 1-cocycles.

1.5 Outline

First, a brief disclaimer: The index and the frequently used notation section at the end of this thesis do not at all claim completeness, they are just there in the hope that they may help the reader a little with the overview over the zoo of vocabulary and symbols in this thesis.

1.5.1 A rough path perspective on renormalization

While renormalization theory for SPDEs has been well established over the last couple of years, for SDEs, such a formalism seems to be obsolete at first sight. However, we show that the Itô-Stratonovich correction can precisely be interpreted as a renormalization of both the branched rough path as well as the equation in question. The action of the renormalization group on the rough path is algebraically described as a translation operator on the space of trees and forests. The effect of the translation on the RDE can then be understood in terms of the pre-Lie product on vector fields $g \triangleright h$, which is given by the derivation of h in the direction of g.

In order to uniquely define the translation operator T_v for branched rough paths, we introduce the free pre-Lie algebra structure on trees discovered in [CL01] through Equations (2.7) and (2.8) in Subsection 2.3.2.2. The pre-Lie product $\tau_1 \curvearrowright \tau_2$ is calculated combinatorially by summing up all possibilities of crafting the tree τ_1 onto the tree τ_2. Any collection of smooth vector fields $(f_1, \dots, f_n)$ then induces a canonical map $v \mapsto f_v$ from the space of trees to the space $\mathrm{Vect}(\mathbb{R}^d)$ of smooth vector fields on $\mathbb{R}^d$ via

$$f_{\tau_1 \curvearrowright \tau_2} := f_{\tau_1} \triangleright f_{\tau_2}, \quad f_{\bullet_i} := f_i,$$

where $\triangleright$ is the pre-Lie product on $\mathrm{Vect}(\mathbb{R}^d)$ given in coordinates as

$$(g^i \partial_i) \triangleright (h^j \partial_j) := (g^i \partial_i h^j) \partial_j.$$

As pointed out by Manchon in [Man11, Section 3.2], this map from trees to vector fields has already been studied in 1857 by Cayley [Cay57]. As one of our main results (Theorem 2.5.10), we show how the pre-Lie constructions make it possible to switch between the RDE with respect to the translated/renormalized rough path and the translated/renormalized RDE with respect to the original rough path:

$$dY = f(Y)\, d(T_v \mathbf{X}) \quad \Longleftrightarrow \quad dY = f(Y)\, d\mathbf{X} + f_v(Y) dt.$$

We furthermore investigate in Section 2.6 how our formalism stands in a one-to-one correspondence with the algebraic renormalization theory for regularity structures due to Bruned–Hairer–Zambotti [BHZ19] applied to a suitable regularity structure for (renormalization of) branched RDEs and a suitable class of models encoding branched rough paths.

1.5.2 Signatures of paths transformed by polynomial maps

The objective in starting this project was the following question asked by Bernd Sturmfels towards the author of this thesis:

Can one express the signature of the path $p(X)$, where p is a polynomial map, in terms of the signature of the original path X, and if yes, how?

The answer is indeed positive, and we introduce an algebra homomorphism M_p that only depends on p with $p(0) = 0$ such that the relation

$$\langle \sigma(X), M_p w \rangle = \langle \sigma(p(X)), w \rangle$$

holds (Theorem 3.1.2). We only show this for piecewise C^1-paths, but there is of course a generalisation to all paths for which the signature is uniquely well defined if one uses Young integral calculus.

We show several properties of the map M_p in Proposition 3.3.2, e.g. that it is compatible with composition of polynomial maps, i.e.

$$M_{q \circ p} = M_p M_q.$$

Furthermore, we discuss in Section 3.3.1 how M_p can be nicely introduced as a homomorphism of *Zinbiel algebras*: M_p is indeed characterized by

$$M_p(\mathtt{i}) = \phi(p_i), \quad M_p(w\mathtt{i}) = M_p(w) \succ \phi(p_i),$$

where ϕ is the canonical algebra homomorphism that turns polynomials on $\mathbb{R}^d$ into shuffle polynomials in $T(\mathbb{R}^d)$.

1.5.3 Areas of areas generate the shuffle algebra

This subsection is based on an extended abstract [Pre21] based on a talk given at the Oberwolfach workshop *New directions in Rough Path Theory*.

We are concerned with the signed area, or Levy area, between two components of a path, which comes with very interesting properties: It is a rotation invariant in two dimensions, it is iterable, meaning all kinds of bracketings like

$$\begin{aligned}&\mathsf{Area}(\mathsf{Area}(\mathsf{X}^1,\mathsf{X}^2),\mathsf{X}^1),\mathsf{Area}(\mathsf{Area}(\mathsf{Area}(\mathsf{X}^1,\mathsf{X}^2),\mathsf{X}^1),\mathsf{X}^1)),\\&\mathsf{Area}(\mathsf{Area}(\mathsf{X}^1,\mathsf{X}^2),\mathsf{Area}(\mathsf{X}^2,\mathsf{Area}(\mathsf{X}^1,\mathsf{X}^2))),\dots\end{aligned}$$

can be computed. Area behaves well with respect to discretization and in the stochastic setting, it preserves the martingale property. Its analytic computation is also very simple:

Let $\mathsf{Area}(\mathsf{X}^1,\mathsf{X}^2)_t$ denote *two times* the signed area enclosed by the two-dimensional path $(\mathsf{X}^1,\mathsf{X}^2)$ up to time t and the straight line connecting $(\mathsf{X}^1(t),\mathsf{X}^2(t))$ with $(\mathsf{X}^1(0),\mathsf{X}^2(0))$. We have

$$\mathsf{Area}(\mathsf{X}^1,\mathsf{X}^2)_t=\int_0^t(\mathsf{X}^1_s-\mathsf{X}^1_0)\,\mathrm{d}\mathsf{X}^2_s-\int_0^t(\mathsf{X}^2_s-\mathsf{X}^2_0)\,\mathrm{d}\mathsf{X}^1_s.$$

Thus, we get

$$\mathsf{Area}(\mathbf{X}^a,\mathbf{X}^b)_t=\mathbf{X}_t^{\mathsf{area}(a,b)},$$

where we define the purely algebraic area operation as $\mathsf{area}(a,b):=a\succ b-b\succ a$, which will be the object of our study.

area is obviously anticommutative, but neither associative,

$$\begin{aligned}&\mathsf{area}(\mathsf{area}(1,2),3)=123-132+213-231-312+321\\&\neq 123-132-213+231-312+321=\mathsf{area}(1,\mathsf{area}(2,3))\end{aligned}$$

nor does it satisfy the Jacobi identity,

$$\begin{aligned}&\mathsf{area}(1,\mathsf{area}(2,3))+\mathsf{area}(2,\mathsf{area}(3,1))+\mathsf{area}(3,\mathsf{area}(1,2))\\&=-123+132+213-231-312+321\neq 0.\end{aligned}$$

It does however satisfy the so-called *Tortkara identity* introduced in [Dzh07]:

$$\begin{aligned}&\mathsf{area}(\mathsf{area}(a,b),\mathsf{area}(c,d))+\mathsf{area}(\mathsf{area}(a,d),\mathsf{area}(c,b))\\&=\mathsf{area}(\mathsf{vol}(a,b,c),d)+\mathsf{area}(\mathsf{vol}(a,d,c),b),\end{aligned}$$

where

$$\mathsf{vol}(a,b,c):=\mathsf{area}(\mathsf{area}(a,b),c)+\mathsf{area}(\mathsf{area}(b,c),a)+\mathsf{area}(\mathsf{area}(c,a),b).$$

vol corresponds to the *signed volume*: $\mathsf{Volume}(\mathbf{X}^a,\mathbf{X}^b,\mathbf{X}^c)=\langle\sigma(X),\mathsf{vol}(a,b,c)\rangle$ is six times the signed volume enclosed by the three-dimensional path $(\mathbf{X}^a,\mathbf{X}^b,\mathbf{X}^c)$, where $\mathbf{X}^z_t=\langle\sigma(\mathsf{X}\restriction_{[0,t]}),z\rangle$ ([DR19, Section 3.1, Equation (4), Definition 3.27, Theorem 3.28]).

As our main result, Corollary 4.4.8, we show that the area algebra $A(\mathbb{R}^d)$ spanned by area bracketings ("areas of areas") starting from the letters $\{1,\dots,d\}$ is itself a generator of the shuffle algebra $(T(\mathbb{R}^d),\sqcup\!\sqcup)$ This is a corollary of the following more general fact (Lemma 4.4.2):

Theorem. Let $X_n\subseteq \mathrm{T}_n(\mathbb{R}^d)$ and $X=\bigcup_n X_n$. Then, For all $n \geq 1$, for all nonzero $L \in \mathfrak{g}_n(\mathbb{R}^d)$ there is an $x \in X_n$ such that $\langle x, L \rangle \neq 0$

if and only if

X shuffle generates the shuffle algebra $T(\mathbb{R}^d)$.

Referring back to signatures, our result means that knowledge of the values of the full increments and of all areas of areas computed from a given path is equivalent to the full signature.

We furthermore look at a left bracketing of area, i.e.

$$\overleftarrow{\mathsf{area}}(\mathtt{i}_1 \dots \mathtt{i}_n) := \mathsf{area}(\dots \mathsf{area}(\mathsf{area}(\mathsf{area}(\mathtt{i}_1, \mathtt{i}_2), \mathtt{i}_3), \mathtt{i}_4), \dots, \mathtt{i}_n)$$

Conjecture. In [DIM19, Section 5] and in [Rei19, Section 3.2, Theorem 31], it was proven for $d = 2$ that a linear basis of A is given by the union of the letters and $(\overleftarrow{\mathsf{area}}\mathtt{i}\mathtt{j}w)_{(\mathtt{i},\mathtt{j},w)}$, where w runs over all words in d letters and $(\mathtt{i}, \mathtt{j})$ over all letters such that $\mathtt{i} < \mathtt{j}$.

The question whether this holds true for $d \geq 3$, too, remains an open problem.

While we didn't manage so far to make progress on this conjecture, through trying to prove it we came along properties of the area left-bracketing which highlight how little is understood about the algebraic operation so far, e.g. the following result, Proposition 4.6.10, which seems quite unexpected at first: We have

$$\overleftarrow{\mathsf{area}}(vx) = \overleftarrow{\mathsf{area}}(v)\overleftarrow{\mathsf{area}}(x)$$

for any $\mathrm{T}^{\geq 1}(\mathbb{R}^d)$ and any Lie polynomial x without a first-order term.

Further open problems include understanding the Tortkara identity in geometric terms as well as finding *free*, i.e. polynomially independent, shuffle generating set consisting of areas of areas, which would allow for a non-redundant storing of the information of the signature in area terms, and, due to the fact that areas of areas of piecewise linear paths are again piecewise linear, also for a very efficient numerical computation of the truncated signature.

Chapter 2

A rough path perspective on renormalization

This chapter is based on [BCFP19], which was published under the Creative Commons License CC-BY-4.0, `https://creativecommons.org/licenses/by/4.0/`. Indication of changes, as required by the license: Changes and additions to the material have been made by the author of this thesis towards the version presented in this chapter.

2.1 Introduction

2.1.1 Rough paths and regularity structures

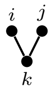
The theory of rough paths, initiated by Lyons through the series of papers [Lyo94], [Lyo95], [Lyo98], deals with controlled differential equations of the form

$$dY_t = f_0(Y_t)dt + \sum_{i=1}^{d} f_i\left(Y_t\right) dX_t^i \quad Y_0 = y_0 \in \mathbb{R}^e \,.$$

with $(X^1, ..., X^d) : [0,T] \to \mathbb{R}^d$, of of low, say α-Hölder, regularity for $0 < \alpha \leq 1$. As may be seen by formal Picard iteration, given a collection $f_0, f_1, ..., f_d$ of nice vector fields on $\mathbb{R}^e$, the solution can be expanded in terms of certain integrals. Assuming validity of the chain-rule, and writing $X^0(t) \equiv t$ for notational convenience, these are just iterated integrals of the form $\int dX^{i_1} \cdots dX^{i_n}$ with integration over n-dimensional simplex. In geometric rough path theory one postulates the existence of such integrals, for sufficiently many words $w = (i_1, \ldots, i_n)$, namely $|w| = n \leq [1/\alpha]$, such as to regain analytic control: the collection of resulting objects

$$\langle \mathbf{X}, w \rangle = \int \cdots \int dX^{i_1} \ldots dX^{i_n} \text{ (integration over } s < t_1 < \cdots < t_n < t \text{, for all } 0 \leq s < t \leq T)$$

subject to suitable analytic and algebraic constraints (in particular, *Chen's relation*, which describes the recentering $s \to \tilde{s}$) is then known as a (level-n) *weakly geometric rough path*, introduced in [Lyo98, Section 2.1 and Definition 2.3.1] (as *geometric p-multiplicative functionals*). For the reader's convenience, we give some precise recalls, along the lines of Hairer–Kelly [HK15, Section 1], in Section 2.1.2 below. Without assuming a chain-rule (think: Itô), iterated integrals of the form $\int X^i X^j dX^k$ appear in the expansion, the resulting objects are then naturally indexed by trees, for example

$$\langle \mathbf{X}, \tau \rangle = \int X^i X^j dX^k \text{ with } \tau = [\bullet_i \bullet_j]_{\bullet_k} \equiv$$  .

The collection of all such objects, again for sufficiently many trees, $|\tau| = \#\text{nodes} \leq [1/\alpha]$ and subject to algebraic and analytic constraints, form what is known as a *branched rough path* ([Gub10, Definition 7.2] [HK15, Definition 1.6 and Definition 2.13]). Here again, we refer to Section 2.1.2 for a precise definition and further recalls.

A basic result - known as the *extension theorem* ([Lyo98, Theorem 2.2.1][Gub10, Theorem 7.3]) - asserts that all "higher" iterated integrals, n-fold with $n > [1/\alpha]$, are automatically well-defined, with validity of all algebraic and analytic constraints in the extended setting.[1] Solving differential equations driven by such rough paths can then be achieved, following [Gub04], see also [FH14, Sections 4, 7, 8], by formulating a fixed point problem in a space of controlled rough paths (first introduced for weakly geometric rough paths in [Gub04, Definition 1]) , essentially a (linear) space of good integrands for rough integration (mind that rough path spaces are, in contrast, fundamentally non-linear due to the aforementioned algebraic constraints). Given a rough differential equation (RDE) of the form

$$dY = f_0(Y)dt + f\,(Y)\,d\mathbf{X}$$

it is interesting to see the effect on Y induced by higher-order perturbations ("translations") of the driving rough path $\mathbf{X}$. For instance, one can use Itô integration to lift a d-dimensional Brownian motion $(B^1, ..., B^d)$ to a (level-2) random rough path, $\mathbf{X} = \mathbf{B}^{\text{Itô}}(\omega)$ of regularity $\alpha \in (1/3, 1/2)$, in which case the above RDE corresponds to the classical Itô SDE

$$dY_t = f_0(Y_t)dt + \sum_{i=1}^{d} f_i\,(Y_t)\,dB_t^i\,, \quad Y_0 = y_0 \in \mathbb{R}^e\,.$$

However, we may perturb $\mathbf{B}^{\text{Itô}} = \left(B, \mathbb{B}^{\text{Itô}}\right)$ via $\mathbb{B}_{s,t}^{\text{Itô}} \mapsto \mathbb{B}_{s,t}^{\text{Itô}} + \frac{1}{2} I\,(t-s) =: \mathbb{B}_{s,t}^{\text{Strat}}$, without touching the underlying Brownian path B. The above RDE then becomes a Stratonovich SDE. On the level of the original (Itô)-equation, the effect of this perturbation is a modified drift vector field,

$$f_0 \rightsquigarrow f_0 + \frac{1}{2} \sum_{i=1}^{d} \nabla_{f_i} f_i\ ,$$

famously known as Itô-Stratonovich correction. Examples from physics (e.g. Brownian motion in a magnetic field) suggest second-order perturbation of the form $\mathbb{B}_{s,t}^{\text{Strat}} \mapsto \mathbb{B}_{s,t}^{\text{Strat}} + a\,(t-s)$, for some $a \in \mathfrak{so}\,(d)$, the SDE is then affected by a drift correction of the form

$$f_0 \rightsquigarrow f_0 + \sum_{i,j} a^{ij}\,[f_i, f_j]\ .$$

In the context of classical SDEs, area corrections are also discussed in [IW89, Chapter VI, Section 7], and carefully designed twisted Wong–Zakai type approximations led Sussmann in[Sus91, Sections 5 and 6] to drift corrections involving higher iterated Lie brackets. This was reconciled

[1] This entire ensemble of iterated integrals is called the *signature* or the *fully lifted rough path*.

with geometric rough path theory in [FO09], and provides a nice example where (Brownian) rough paths (with $\gamma = \frac{1}{2}-$ regularity) need to be studied in the *entire* scale of different rough path topologies indexed by $\gamma \in (0, 1/2)$.

As we shall see, all these examples are but the tip of an iceberg. It will also be seen that there is a substantial difference between the weakly geometric rough path case and the generality aimed for in this chapter.

We finally note that tampering with "higher levels" of the lifted noise also affects the structure of stochastic partial differential equations: this is not only omnipresent in the case of singular SPDEs, see e.g. [Hai14], [GIP15], but very much in every SPDE with rough path noise as remarked e.g. in [CFO11, Theorem 2].

From rough paths to regularity structures. The theory of regularity structures introduced by Hairer in [Hai14] extends rough path theory and then provides a remarkable framework to analyse (singular) semi-linear stochastic partial differential equations, e.g. of the form

$$(\partial_t - \Delta)\, u = f\,(u, Du) + g\,(u)\,\xi\,(t, x, \omega)$$

with $(d+1)$-dimensional space-time white noise. The demarche is similar as above: noise is lifted to a *model*, whose algebraic properties (especially with regard to recentering) are formulated with the aid of the *structure group*. Given an (abstract) model, a fixed point problem is solved and gives a solution in a (linear) space of *modelled distributions*. The abstract solution can then be mapped ("*reconstructed*") into an actual distribution (a.k.a generalized function). In fact, one has the rather precise correspondences as follows (see [FH14, Sections 13.2.2 and 13.3.2] for explicit details in the level-2 setting):

rough path	$\longleftrightarrow$	model
Chen's relation	$\longleftrightarrow$	structure group
controlled rough path	$\longleftrightarrow$	modelled distribution
rough integration	$\longleftrightarrow$	reconstruction map

Table 2.1: Basic correspondences: rough paths $\longleftrightarrow$ regularity structures

Furthermore, one has similar results concerning continuity properties of the solution map as a function of the enhanced noise:

continuity of solution in (rough path $\longleftrightarrow$ model) topology

Unfortunately mollified lifted noise - in infinite dimensions - in general does not converge (as a model), hence *renormalization* plays a fundamental role in regularity structures. The algebraic formalism of how to conduct this renormalization then relies on heavy Hopf algebraic considerations (originally [Hai14, Sections 8.1, 8.3, 9.1 and 9.2]), pushed to great generality in [BHZ19], see also [Hai16, Section 2] for a summary. Our investigation was driven by two questions:

(1) Are there meaningful (finite-dimensional) examples from stochastics which require (infinite) renormalization?

(2) Do we have algebraic structures in rough path theory comparable with those seen in regularity structures?

To be more specific, with regard to (1), consider the situation of a differential equation driven by some finite-dimensional Brownian (or more general Gaussian) noise, mollified at scale ε, followed by the question if the resulting (random) ODE solutions converge as $\varepsilon \to 0$. As remarked explicitly in [FH14, at the end of Section 15.4], this is very often the case (with concrete results given in [FH14, Section 10]), with the potential caveat of area (and higher order) anomalies (see [FO09], [FGL15], ...), leading to a more involved description (sometimes called *finite renormalization*) of the limit. We emphasize, however, that this is not always the case; there are perfectly meaningful (finite-dimensional) situations which require (infinite) renormalization, which we sketch in Section 2.4.3 below together with precise references. We further highlight that a natural example of geometric rough path (over $\mathbb{R}^2$) with a logarithmically diverging area term requiring (infinite) renormalization appears in Hairer's solution of the KPZ equation, see [Hai13, Section 7]. This situation is characteristic of singular SPDEs, in which the procedure described above typically leads to plain divergence, cured by "subtracting infinities", a.k.a. *infinite renormalization*.

Much effort in this work is then devoted to question (2): In [BHZ19], the algebraic formalism in regularity structures relies crucially on two Hopf algebras, $\mathcal{T}_+$ and $\mathcal{T}_-$ [BHZ19, Section 5.5] (which are furthermore in "cointeraction", see [BHZ19, Theorem 5.37]). The first one helps to construct the structure group which, in turn, provides the recentering (as it is called in [BHZ19, Section 1], while in [Hai16, Section 3] or in the talk [Zam16] the term "positive renormalization" appears) and hence constitutes an abstract form of Chen's relation in rough path theory. In this sense, $\mathcal{T}_+$ was always present in rough path theory, the point being enforced in the model case of branched rough paths (see [Gub10],[HK15]) where $\mathcal{T}_+$ is effectively given by the Connes-Kreimer Hopf algebra.

Question (2) is then reduced to the question if $\mathcal{T}_-$, built to carry out the actual renormalization of models, and subsequently SPDEs, ("negative renormalization" in the language of [Hai16, Section 3] and [Zam16]), has any correspondence in rough path theory. Our answer is again affirmative and we establish the precise relation:

$$\boxed{\text{translation of rough paths} \longleftrightarrow \text{renormalization of models}}$$

At last, during the course of writing the paper behind this chapter, the authors of [BCFP19] realized that we have been touching on a third important point, whose importance seems to transcend the rough path setting in which it is discussed.

(3) How does one obtain from the renormalized model, in some algebraic and automated fashion, the renormalized equation?

It is indeed the algebraic approach to "translation of rough paths" (i.e. renormalization of a branched rough path model) that indicated an important role played by pre-Lie structures, which first appear in Section 2.3.2 to construct the translation operator (on forest series) and then to characterize its dual. These considerations help answer the (not very precise) question of what pre-Lie structures (after all, a well-known tool in the renormalization theory, see e.g. [Man11, Section 2.4] and the references therein) have to do with rough paths and regularity structures. From a regularity structures perspective, a remarkable consequence is that this allows to understand *directly* the action of the renormalization group on the (to-be-renormalized) equation at hand, thus providing an answer to question (3). Indeed, by exploiting the pre-Lie structure of the space of trees, we obtain a direct conversion formula for the RDE driven by a translated branched rough path; see Section 2.5.2, Theorem 2.5.10, and Remark 2.5.14. The analogous

statement in regularity structures is an explicit form of an arbitrary renormalised SPDE, a result which was recently established as [BCCH21, Theorem 3.25].
Several remarks are in order.

- We first develop the algebraic renormalization theory for rough paths in its own right, analytic considerations then take place in Section 2.5. The link to regularity structures and its renormalization theory is only made in Section 2.6.
- While pre-Lie homomorphisms play a crucial role in the construction of translation maps, we point out that in certain situations the fine-details of pre-Lie structures are not visible; see the final point of Theorem 2.5.1, as well as Remarks 2.5.5 and 2.6.11.
- In Section 2.4.3 we present several examples, based on finite- (and even one-) dimensional Brownian motion, which do require genuine renormalization. Another interesting type of rough path renormalization, aiming at fractional Brownian (recall divergence of Lévy area for Hurst parameter $H \leq 1/4$ from [CQ02, Theorem 2]) based on Fourier normal ordering, was proposed by Unterberger in [Unt13]. That said, his methods and aims are quite different from those considered in this chapter and/or those from Hairer's regularity structures.
- The existence of a finite-dimensional renormalization group is much related to the stationarity of the (lifted) noise, see [Hai14], [CH18], and the stationarity assumption in [BHZ19, Theorem 6.18]. In the rough path context, this amounts to saying that a random (branched) rough path $\mathbf{X} = \mathbf{X}(\omega)$, with values in a (truncated) Butcher (hence finite-dimensional Lie) group $\mathcal{G}$, actually has independent increments with respect to the Grossmann-Larson product $\star$ (dual to the Connes-Kreimer coproduct $\Delta_\star$). In other words, $\mathbf{X}$ is a (continuous) $\mathcal{G}$-valued Lévy process. This yields a close connection to the works [FS17, Section 3] and [Che18] devoted to the study of Lévy rough paths; in Section 2.4.2 we shall see how Lévy triplets (or rather tuples in the absence of jumps) behave under renormalization.

2.1.2 Weakly geometric and non-geometric rough paths

In this subsection, we briefly recall the notions of weakly geometric and branched rough paths; see [FV10], [Gub10] and [HK15] for further details. See also Sections 2.2.1 and 2.3.1 for further details on the algebraic structures involved.

Weakly geometric rough paths. We follow [HK15, Section 1.1]. Consider a path $X : [0,T] \to \mathbb{R}^d$. A (weakly geometric) rough path over X is a map $\mathbf{X} : [0,T]^2 \to T((\mathbb{R}^d))$, where $T((R^d)) = \prod_{k=0}^{\infty}(\mathbb{R}^d)^{\otimes k}$ is the space of "tensor series" over $\mathbb{R}^d$, which should be thought of as the iterated integrals of X. Equipping $\mathbb{R}^d$ with an inner product, we can identify $T((\mathbb{R}^d))$ with the algebraic dual of the tensor algebra

$$T(\mathbb{R}^d) = \mathbb{R} \oplus \mathbb{R}^d \oplus (\mathbb{R}^d)^{\otimes 2} \oplus \dots \, .$$

One should think of the components of $\mathbf{X}$ as formally being given by [HK15, Equation (1.3)]

$$\langle \mathbf{X}_{s,t}, \mathtt{i}_1 \cdots \mathtt{i}_n \rangle := \int_s^t \cdots \int_s^{t_2} dX_{t_1}^{i_1} \dots dX_{t_n}^{i_n} \ , \tag{2.1}$$

for $\mathtt{i}_1, \dots, \mathtt{i}_n \in \{\mathtt{1}, \dots, \mathtt{d}\}$, where $X_t^i - X_s^i = \langle X_{st}, \mathtt{i} \rangle$ and where we use the shorthand $\mathtt{i}_1 \cdots \mathtt{i}_n = \mathtt{i}_1 \bullet \cdots \bullet \mathtt{i}_n$ with $\bullet$ denoting the tensor product in $T(\mathbb{R}^d)$. We emphasize that, as it is also noted in [HK15, Section 1.1], unless $n = 1$, the definition (2.1) is, in general, only formal and one should think of the rough path $\mathbf{X}$ as defining the RHS.

Observe that if X is smooth and (2.1) is used to define $\mathbf{X}$, then the so-called shuffle identity holds [HK15, Equation (1.4)]

$$\langle \mathbf{X}_t, \mathtt{i}_1 \cdots \mathtt{i}_n \rangle \langle \mathbf{X}_t, \mathtt{j}_1 \cdots \mathtt{j}_m \rangle = \langle \mathbf{X}_t, \mathtt{i}_1 \cdots \mathtt{i}_n \sqcup\!\sqcup \mathtt{j}_1 \cdots \mathtt{j}_m \rangle \;, \quad \text{for all } \mathtt{i}_1 \cdots \mathtt{i}_n, \mathtt{j}_1 \cdots \mathtt{j}_m \in T(\mathbb{R}^d) \;, \tag{2.2}$$

where $\sqcup\!\sqcup$ denotes the commutative associative shuffle product [Reu93, Section 1.4]. While we do not give the definition of $\sqcup\!\sqcup$ here or prove this identity, we remark that it is a direct consequence of integration by parts. Another important algebraic identity which holds in this case is Chen's relation

$$\mathbf{X}_{s,t} = \mathbf{X}_{s,u} \bullet \mathbf{X}_{u,t} \;, \quad \text{for all } s,t,u \in [0,T] \;,$$

which can be shown by an application of Fubini's theorem.

The concept of a (weakly geometric) rough path should be thought of as a generalisation of these identities to paths of lower regularity.

Definition 2.1.1 ([HK15, Definition 1.2])**.** *Let* $\gamma \in (0,1]$*. A* γ-Hölder weakly geometric rough path *is a map* $\mathbf{X} : [0,T]^2 \to T((\mathbb{R}^d))$ *satisfying*

i) $\langle \mathbf{X}_{st}, x \sqcup\!\sqcup y \rangle = \langle \mathbf{X}_{st}, x \rangle \langle \mathbf{X}_{st}, y \rangle$*, for all* $x, y \in T(\mathbb{R}^d)$*,*

ii) $\mathbf{X}_{st} = \mathbf{X}_{su} \bullet \mathbf{X}_{ut}$ *for all* $s,t,u \in [0,T]$*,*

iii) $\sup_{s\neq t} \frac{|\langle \mathbf{X}_{st}, \mathtt{i}_1 \cdots \mathtt{i}_n \rangle|}{|t-s|^{\gamma n}} < \infty$*, for all* $n \geq 1$ *and* $\mathtt{i}_1, \ldots, \mathtt{i}_n \in \{\mathtt{1}, \ldots, \mathtt{d}\}$*.*

Branched rough paths. One is often interested in paths X for which natural definitions of "iterated integrals" do not satisfy classical integration by parts and thus do not constitute weakly geometric rough paths, e.g., integrals defined in the sense of Itô for a semi-martingale X. Branched rough paths are a generalisation of weakly geometric rough paths which allows for violation of the shuffle identity (2.2) and thus of the usual rules of calculus. This is achieved by substituting the space $T((\mathbb{R}^d))$ with a larger (Hopf) algebra $\mathcal{H}^*$ in which natural generalisations of properties i), ii), and iii) are required to hold. The Hopf algebra $\mathcal{H}^*$ is known as the Grossman–Larson algebra of series of forests, and is the algebraic dual of the Connes–Kreimer Hopf algebra [Kre98, Section 3][CK98, Section 2 and Section 3] consisting of polynomials of rooted trees with nodes decorated by the set $\{1,\ldots,d\}$.

Denoting by $\odot$ the (commutative) polynomial product on $\mathcal{H}$ and by $\star$ the (non-commutative) Grossman–Larson product on $\mathcal{H}^*$, we have the following analogue of Definition 2.1.1.

Definition 2.1.2 ([HK15, Definition 1.6])**.** *Let* $\gamma \in (0,1]$*. A* γ-Hölder branched rough path *is a map* $\mathbf{X} : [0,T]^2 \to \mathcal{H}^*$ *satisfying*

a) $\langle \mathbf{X}_{st}, \tau_1 \odot \tau_2 \rangle = \langle \mathbf{X}_{st}, \tau_1 \rangle \langle \mathbf{X}_{st}, \tau_2 \rangle$ *for all* $\tau_1, \tau_2 \in \mathcal{H}$*,*

b) $\mathbf{X}_{st} = \mathbf{X}_{su} \star \mathbf{X}_{ut}$ *for all* $s,t,u \in [0,T]$*,*

c) $\sup_{s\neq t} \frac{|\langle \mathbf{X}_{st}, \tau \rangle|}{|t-s|^{\gamma|\tau|}} < \infty$ *for every forest* $\tau \in \mathcal{H}$*, where* $|\tau|$ *denotes the number of nodes in* τ*.*

Here we set $\langle \mathbf{X}_{s,t}, \bullet_i \rangle := X^i_{s,t}$ and then think of the components of $\mathbf{X}$ given by the formal recursion

$$\langle \mathbf{X}_{s,t}, [\tau_1 \odot \ldots \odot \tau_n]_{\bullet_i} \rangle = \int_s^t \langle \mathbf{X}_{s,u}, \tau_1 \rangle \ldots \langle \mathbf{X}_{s,u}, \tau_n \rangle dX^i_u \tag{2.3}$$

for trees $\tau_1, \ldots, \tau_n \in \mathcal{H}$ and $i \in \{1,\ldots,d\}$, where $[\tau_1 \odot \ldots \odot \tau_n]_{\bullet_i}$ denotes the tree formed by grafting the trees $\tau_1, \ldots, \tau_n$ onto a single root with label i. If X is smooth and one uses (2.3) to define $\mathbf{X}$, then, as before, points a) and b) are direct consequences of integration by parts and Fubini's theorem respectively.

Equipping $T((\mathbb{R}^d))$ with the tensor Hopf algebra structure, there is a canonical graded embedding of Hopf algebras $T((\mathbb{R}^d)) \hookrightarrow \mathcal{H}^*$. Points a), b), and c) are therefore generalisations of points i), ii), and iii), hence every weakly geometric rough path constitutes a branched rough path. We emphasize however that this embedding is strict and a) is more general than i), which allows a general branched rough path $\mathbf{X}$ to violate classical integration by parts. For example, if $\mathbf{X}$ is defined via (2.3) using Itô integrals for a semi-martingale X, then $\mathbf{X}$ is an example of a γ-Hölder branched (but in general not weakly geometric!) rough path for any $\gamma \in (0, \frac{1}{2})$.

2.1.3 Translation of paths

Consider a d-dimensional path X_t, written with respect to an orthonormal basis $e_1, \dots, e_d$ of $\mathbb{R}^d$,

$$X_t = \sum_{i=1}^{d} X_t^i e_i.$$

We are interested in constant speed perturbations, of the form

$$T_v X_t := X_t + tv, \quad \text{with } v = \sum_{i=1}^{d} v^i e_i \in \mathbb{R}^d.$$

In coordinates, $(T_v X_t)^i = X_t^i + tv^i$ for $i = 1, \dots, d$, i.e,

$$\langle T_v X, e_i \rangle = \langle X_t, e_i \rangle + \langle tv, e_i \rangle .$$

Consider now an orthonormal basis $e_0, e_1, \dots, e_d$ of $\mathbb{R}^{1+d}$, and consider the $\mathbb{R}^{1+d}$-valued "time-space" path

$$\bar{X}_t = X_t + X_t^0 e_0 = \sum_{i=0}^{d} X_t^i e_i$$

with scalar-valued $X_t^0 \equiv t$. We can now write

$$T_v \bar{X}_t = \bar{X}_t + tv = X_t + X_t^0 (e_0 + v)$$

which identifies T_v as linear map on $\mathbb{R}^{1+d}$, which maps $e_0 \mapsto e_0 + v$, and $e_i \mapsto e_i$ for $i = 1, \dots, d$. We then can (and will) also look at general endomorphisms of the vector space $\mathbb{R}^{1+d}$, which we still write in the form

$$\begin{aligned} e_j &\mapsto e_j + v_j, \ j = 0, \dots, d \\ v_j &= \sum_{i=0}^{d} v_j^i e_i \in \mathbb{R}^{1+d}. \end{aligned}$$

(The initially discussed case corresponds to $(v_0, v_1, \dots, v_d) = (v_0, 0, \dots, 0)$, with $\langle v_0, e_0 \rangle = 0$, and much of the sequel, will take advantage of this additional structure.)

We shall be interested to understand how such perturbations propagate to higher-level iterated integrals, whenever X has sufficient structure to make this meaningful. For instance, if

$X = B(\omega)$, a d-dimensional Brownian motion, an object of interest would be, with repeated (Stratonovich) integration over $\{(r,s,t) : 0 \le r \le s \le t \le T\}$,

$$(T_v B)^{ijk}_{0,T} := \int \circ (dB^i + v^i\, dr) \circ (dB^j + v^j\, ds) \circ (dB^k + v^k\, dt) = B^{ijk}_{0,T} + ...$$

where the omitted terms (dots) can be spelled out (algebraically) in terms of contractions of v (resp. tensor-powers of v) and iterated integrals of $(1+d)$-dimensional time-space Brownian motion "(t,B)". (Observe that we just gave a dual description of this perturbation, as seen on the third level, while the initial perturbation took place at the first level: v is a vector here.)

There is interest in higher-level perturbations. In particular, given a 2-tensor $v = \sum_{i,j=1}^d v^{ij} e_{i,j}$, we can consider the level-2-perturbation with no effect on the first level, i.e., $(T_v B)^i_t \equiv B^i_t$, while for all $i, j = 1, ..., d$,

$$(T_v B)^{ij}_t = B^{ij}_t + v^{ij}\, t$$

For instance, writing $B^{I;w}$ for iterated Itô integrals, in contrast to B^w defined by iterated Stratonovich integration, we have with $v := \frac{1}{2} I^d$ where $(I^d)^{ij} = \delta^{ij}$, i.e., the identity matrix,

$$(T_v B^I)^{ij}_t = B^{ij}_t.$$

This is nothing but a restatement of the familiar formula $\int_0^t B^i dB^j + \frac{1}{2}\delta^{ij} t = \int_0^t B^i \circ dB^j$. It is a non-trivial exercise to understand the Itô-Stratonovich correction at the level of higher iterated integrals, cf. [Ben89, Proposition 1], and a "branched" version thereof discussed in Section 2.4.1 below. Further examples where such translations serve as a "renormalisation" are discussed in Section 2.4.3, notably the case $\bar{B}^{ij} = (T_a B)^{ij}_t$ with an anti-symmetric 2-tensor $a = (a^{ij})$ which arises in the study of Brownian particles in a magnetic field.

It will be important for us to understand (explicitly) how to formulate (constant speed, higher) order translations, an analytic operation on rough paths, algebraically and "point-wise" terms of the time-space rough path.

2.1.4 Organization of the chapter

This chapter is organized as follows. In Section 2.2, we first discuss renormalization/translation in the by now well-established setting of weakly geometric rough paths. The algebraic background is found for instance in [Reu93, Chapter 0 to 3]. We then, in Section 2.3, move to branched rough paths [Gub10, Definition 7.2], in the notation and formalism from Hairer-Kelly [HK15], and in particular introduce the relevant pre-Lie structures. In Section 2.4 we illustrate the use of the (branched) translation operator (additional examples were already mentioned in Section 2.4.3), while in Section 2.5 we describe the analytic and algebraic effects of such translations on rough paths and associated RDEs. Lastly, Section 2.6 is devoted to the systematic comparison of the translation operator and the "negative renormalization" introduced in [BHZ19].

Acknowledgements. P.K.F. is partially supported by the European Research Council through CoG-683164 and DFG research unit FOR2402. I.C., affiliated to TU Berlin when this project was commenced, was supported by DFG research unit FOR2402, and is currently supported by a Junior Research Fellowship of St John's College, Oxford. Y.B. thanks Martin Hairer for financial support from his Leverhulme Trust award. R.P. is supported by European Research Council through CoG-683164 and was affiliated to Max Planck Institute for Mathematics in the Sciences, Leipzig, in autumn 2018. R.P. also thanks Alexander Schmeding for helpful discussions and information on topological questions.

2.2 Translation of weakly geometric rough paths

We review the algebraic setup for weakly geometric rough paths, as enhancements of $X = (X_0, X_1, ..., X_d)$*, a signal with values in* $V = \mathbb{R}^{1+d}$*. Recall the natural state-space of such rough paths is* $T((V))$*, a space of tensor series (resp. a suitable truncation thereof related to the regularity of* X*). Typically* $\dot{X} \equiv (\xi_0, \xi_1, ..., \xi_d)$ *models noise. Eventually, we will be interested in* $X_0(t) = t$*, so that* X *is a time-space (rough) path, though this plays little role in this section.*

2.2.1 Preliminaries for tensor series

We first establish the notation and conventions used throughout the chapter. Most algebraic aspects used in this section may be found in [Reu93, Chapter 0 to 3] and [FV10, Chapter 7].

Throughout Section 2.2 we let $\{\mathtt{0}, \mathtt{1}, \ldots, \mathtt{d}\}$ be a basis for $\mathbb{R}^{1+d}$. Consider the vector space of formal tensor series over $\mathbb{R}^{1+d}$

$$T((\mathbb{R}^{1+d})) = \prod_{k=0}^{\infty} (\mathbb{R}^{1+d})^{\otimes k}$$

(with the usual convention $(\mathbb{R}^{1+d})^{\otimes 0} = \mathbb{R}$), as well as the vector space of polynomials over $\mathbb{R}^{1+d}$

$$T(\mathbb{R}^{1+d}) = \bigoplus_{k=0}^{\infty} (\mathbb{R}^{1+d})^{\otimes k}.$$

Note that $T(\mathbb{R}^{1+d})$ and $T((\mathbb{R}^{1+d}))$ can equivalently be considered as the vector space of words and non-commutative series respectively in $1 + d$ indeterminates.

Recall that $T((\mathbb{R}^{1+d}))$ can be equipped with a Hopf-type[2] algebra structure

$$(T((\mathbb{R}^{1+d})), \bullet, \Delta_{⧢}, \alpha)$$

with tensor (concatenation) product $\bullet$, coproduct $\Delta_{⧢}$ which is dual to the shuffle product ⧢ on $T(\mathbb{R}^{1+d})$, and antipode α. Recall that $\Delta_{⧢}$ is explicitly given as the unique continuous[3] algebra homomorphism such that

$$\begin{aligned} \Delta_{⧢} &: T((\mathbb{R}^{1+d})) \to T(\mathbb{R}^{1+d}) \overline{\otimes} T(\mathbb{R}^{1+d}) \\ \Delta_{⧢} &: v \mapsto v \otimes 1 + 1 \otimes v, \quad \text{for all } v \in \mathbb{R}^{1+d}. \end{aligned}$$

We shall often refer to elements $\mathtt{i}_1 \bullet \ldots \bullet \mathtt{i}_k$ as words consisting of the letters $\mathtt{i}_1, \ldots, \mathtt{i}_k \in \{\mathtt{0}, \ldots, \mathtt{d}\}$, and shall write $\mathtt{i}_1 \cdots \mathtt{i}_k = \mathtt{i}_1 \bullet \ldots \bullet \mathtt{i}_k$. We likewise denote by

$$(T(\mathbb{R}^{1+d}), ⧢, \Delta_{\bullet}, \tilde{\alpha})$$

the shuffle Hopf algebra. Recall that by identifying $\mathbb{R}^{1+d}$ with its dual through the basis $\{\mathtt{0}, \ldots, \mathtt{d}\}$, there is a natural duality between $T(\mathbb{R}^{1+d})$ and $T((\mathbb{R}^{1+d}))$ in which ⧢ is dual to $\Delta_{⧢}$, and $\bullet$ is dual to $\Delta_{\bullet}$.

We let $G(\mathbb{R}^{1+d})$ and $\mathfrak{g}((\mathbb{R}^{d+1}))$ denote the set of group-like and primitive elements of $T((\mathbb{R}^{1+d}))$ respectively. Recall that $\mathfrak{g}((\mathbb{R}^{d+1}))$ is precisely the space of Lie series over $\mathbb{R}^{1+d}$, and that

$$G(\mathbb{R}^{1+d}) = \exp_{\bullet}(\mathfrak{g}((\mathbb{R}^{d+1}))).$$

[2]The structure here is not exactly of a Hopf algebra since $\Delta_{⧢}$ does not map $T((\mathbb{R}^{1+d}))$ into $T((\mathbb{R}^{1+d}))^{\otimes 2}$, but rather into the complete tensor product $T(\mathbb{R}^{1+d})^{\overline{\otimes} 2} \simeq \prod_{k,m=0}^{\infty} (\mathbb{R}^{1+d})^{\otimes k} \otimes (\mathbb{R}^{1+d})^{\otimes m}$, see [Reu93, Section 1.4].

[3]We equip henceforth $T((\mathbb{R}^{1+d}))$ and $T(\mathbb{R}^{1+d}) \overline{\otimes} T(\mathbb{R}^{1+d})$ with the product topology.

For any integer $N \geq 0$, we denote by $T^N(\mathbb{R}^{1+d})$ the truncated algebra obtained as the quotient of $T((\mathbb{R}^{1+d}))$ by the ideal consisting of all series with no words of length less than N (we keep in mind that the tensor product is always in place on $T^N(\mathbb{R}^{1+d})$). Similarly we let $G^N(\mathbb{R}^{1+d}) \subset T^N(\mathbb{R}^d)$ and $\mathfrak{g}_{\leq N}(\mathbb{R}^{d+1}) \subset T^N(\mathbb{R}^{1+d})$ denote the step-N free nilpotent Lie group and Lie algebra over $\mathbb{R}^{1+d}$ respectively, constructed in analogous ways.

Finally, we identify $\mathbb{R}^d$ with the subspace of $\mathbb{R}^{1+d}$ with basis $\{\mathtt{1}, \dots, \mathtt{d}\}$. From this identification, we canonically treat all objects discussed above built from $\mathbb{R}^d$ as subsets of their counterparts built from $\mathbb{R}^{1+d}$. For example, we treat the algebra $T((\mathbb{R}^d))$ and Lie algebra $\mathfrak{g}_{\leq N}(\mathbb{R}^d)$ as a subalgebra of $T((\mathbb{R}^{1+d}))$ and a Lie subalgebra of $\mathfrak{g}_{\leq N}(\mathbb{R}^{d+1})$ respectively.

2.2.2 Translation of tensor series

The idea for the proof of the following Lemma was given by Mateusz Michałek (back then MPI MIS Leipzig, now Universität Konstanz) in an E-Mail to the author of this thesis from September 18th, 2018. Michałek's E-Mail contained the full argumentation of the final paragraph of the proof, starting from the fact that an algebra endomorphism of $T((\mathbb{R}^{1+d}))$ must map an element with lowest degree k to an element with lowest degree not smaller than k.

Lemma 2.2.1. *Let $M_1, M_2 : T((\mathbb{R}^{1+d})) \to T((\mathbb{R}^{1+d}))$ be algebra homomorphisms with $M_1\restriction_{T(\mathbb{R}^{1+d})} = M_2\restriction_{T(\mathbb{R}^{1+d})}$. Then $M_1 = M_2$.*

Proof. We may assume that $M_1\mathbf{e} = M_2\mathbf{e} = \mathbf{e}$, because otherwise the statement is trivial.

Let $\deg(v) \in \mathbb{N}_0$ denote the first nonzero degree of an element $v \in T((\mathbb{R}^{1+d}))$. Let $M : T((\mathbb{R}^{1+d})) \to T((\mathbb{R}^{1+d}))$ be an algebra homomorphism with $M\mathbf{e} = \mathbf{e}$.

First of all, $\deg(M\mathtt{i}) \geq 1$ for all $i \in \{1, \dots, 1+d\}$. Otherwise, if a_0 denotes the zeroth component of $M\mathtt{i}$, $\mathtt{i} - a_0\mathbf{e}$ would have an inverse $b \in T((\mathbb{R}^d))$, i.e. $(\mathtt{i} - a_0\mathbf{e}) \bullet b = \mathbf{e}$ and thus

$$(M\mathtt{i} - a_0\mathbf{e}) \bullet Mb = \mathbf{e}.$$

But $\deg(M\mathtt{i} - a_0\mathbf{e}) \geq 1$, and thus $M\mathtt{i} - a_0\mathbf{e}$ does not have an inverse, which is the desired contradiction.

As a consequence, $\deg(Mu) \geq n$ for any non-zero homogeneous element of degree n.

Observe that if $v \in T((\mathbb{R}^{1+d}))$ with $\deg(v) = n$, then, for some $m \in \mathbb{N}$, there are homogeneous elements $u_1, \dots, u_m \in T(\mathbb{R}^{1+d})$ of degree n and $v_1, \dots, v_m \in T((\mathbb{R}^{1+d}))$ with $\deg(v_1) = \cdots = \deg(v_m) = 0$ such that

$$v = \sum_{i=1}^{m} u_i \bullet v_i.$$

Since $\deg(Mu_i) \geq n$, we have $\deg(M(u_i) \bullet M(v_i)) \geq n$ and thus $\deg(Mv) \geq n$.

Finally, assume there is $x \in T((\mathbb{R}^{1+d}))$ such that $M_1x \neq M_2x$. Put $k := \deg(M_1x - M_2x) < \infty$ and let x_k denote the canonical projection of x onto $T^{\leq k}(\mathbb{R}^{1+d})$. By assumption, $M_1x_k = M_2x_k$, and thus $M_1x - M_2x = M_1(x - x_k) - M_2(x - x_k)$. But $\deg(x - x_k) > k$ and thus $\deg(M_1x - M_2x) > k$, which is again a contradition. ☐

Note that the argument of Lemma 2.2.1 also works for algebra homomorphisms $M_1, M_2 : T((\mathbb{R}^{1+d})) \to T((\mathbb{R}^{1+d})) \mathbin{\overline{\otimes}} T((\mathbb{R}^{1+d}))$.

Recall that, by the universal property of $T(\mathbb{R}^{1+d})$ any linear map $M : \mathbb{R}^{1+d} \to T((\mathbb{R}^{1+d}))$ extends uniquely to a algebra homomorphism $M : T(\mathbb{R}^{1+d}) \to T((\mathbb{R}^{1+d}))$, and, if $M(\mathtt{i})$ has no component of order zero (i.e., $\langle M(\mathtt{i}), \mathbf{e}\rangle = 0$ for all $\mathtt{i} \in \{\mathtt{0}, \dots, \mathtt{d}\}$, we may put $M(\sum_w a_w w) := \sum_w a_w Mw$ for any $\sum_w a_w w \in T((\mathbb{R}^{1+d}))$, which constitutes an extension of M to an algebra homomorphism on $T((\mathbb{R}^{1+d}))$ and is thus the unique one through Lemma 2.2.1.

Definition 2.2.2. *For a collection of Lie series* $v = (v_0, \dots, v_d) \subset \mathfrak{g}((\mathbb{R}^{d+1}))$, *define* $T_v : T((\mathbb{R}^{1+d})) \to T((\mathbb{R}^{1+d}))$ *as the unique extension to a algebra homomorphism of the linear map*

$$T_v : \mathbb{R}^{1+d} \to \mathfrak{g}((\mathbb{R}^{d+1})) \subset T((\mathbb{R}^{1+d}))$$
$$T_v : \mathtt{i} \mapsto \mathtt{i} + v_i, \quad \text{for all } i \in \{0, \dots, d\}.$$

In the sequel we shall often be concerned with the case that $v_i = 0$ for $i = 1, \dots, d$ and v_0 takes a special form. We shall make precise whenever such a condition is in place by writing, for example, $v = v_0 \in \mathfrak{g}_{\leq N}(\mathbb{R}^d)$.

Remark 2.2.3. We observe the following immediate properties of T_v:

1. Since T_v is a continuous algebra homomorphism which preserves the Lie algebra $\mathfrak{g}((\mathbb{R}^{d+1}))$, it holds that T_v maps $G(\mathbb{R}^{1+d})$ into $G(\mathbb{R}^{1+d})$;

2. $T_v \circ T_u = T_{v+T_v(u)}$, where we write $T_v(u) := (T_v(u_0), \dots, T_v(u_d))$. Indeed,

$$(T_v \circ T_u)(\mathtt{i}) = T_v(\mathtt{i} + u_i) = \mathtt{i} + v_i + T_v(u_i)$$

   for all $i = 0, \dots, d$, and since $T_v \circ T_u$ is a continuous algebra homomorphism again, it must be identical to $T_{v+T_v(u)}$. In particular, $T_{v+u} = T_v \circ T_u$ for all $v = v_0, u = u_0 \in \mathfrak{g}((\mathbb{R}^d))$; i.e., $\{T_v | v = v_0 \in \mathfrak{g}((\mathbb{R}^d))\}$ forms an abelian group isomorphic to $(\mathfrak{g}((\mathbb{R}^d)), +)$.

3. We furthermore have that $\{T_v | v_i \in \hat{\mathfrak{g}}_{\geq 2}(\mathbb{R}^{d+1}) \, \forall i\}$ forms a non-abelian group. Indeed, we may construct the inverse of T_v as $T_{\bar{v}}$ with

$$\bar{v} = -v - \sum_{n=1}^{\infty} (\mathrm{id} - T_v)^{\circ n}(v)$$

   where the sum converges for every $v \in [\hat{\mathfrak{g}}_{\geq 2}(\mathbb{R}^{d+1})]^{d+1}$ as $(\mathrm{id} - T_v)$ strictly raises the minimal homogeneity of any element of $T_{\geq 1}((\mathbb{R}^{d+1}))$, making $\mathsf{proj}_m \circ \sum_{n=1}^{\infty} (\mathrm{id} - T_v)^{\circ n}$ a finite sum for any m.

   Indeed, for arbitrary $v \in [\hat{\mathfrak{g}}_{\geq 2}(\mathbb{R}^{d+1})]^{d+1}$ we then have

$$v + T_v(\bar{v}) = v + \bar{v} + (T_v - \mathrm{id})\bar{v} = v - v - (T_v - \mathrm{id})v - \sum_{n=1}^{\infty} (\mathrm{id} - T_v)^{\circ n}(v) + \sum_{n=2}^{\infty} (\mathrm{id} - T_v)^{\circ n}(v) = 0,$$

   thus $T_v \circ T_{\bar{v}} = T_0 = \mathrm{id}$, which in particular means that $T_v : T((\mathbb{R}^{d+1})) \to T((\mathbb{R}^{d+1}))$ is surjective. However T_v is also injective, as $\mathsf{proj}_n T_v x = \mathsf{proj}_n x$ for any x with lowest non-zero homogeneity n. Thus $T_{\bar{v}} = T_v^{-1}$. (In fact, $\bar{v} = -T_v^{-1}(v)$, the sum in the definition of $\bar{v}$ being a standard construction of a formal series inverse, cf. e.g. [Hai14, Equation (2.22)].)

   Note that $T_v - \mathrm{id}$ for $v \neq 0$ is not a homomorphism and in particular not equal to the homomorphism $T_{\tilde{v}}$ with $\tilde{v}_i = -\mathtt{i} + v$.

4. For every integer $N \geq 0$, T_v induces a well-defined algebra homomorphism $T_v^N : T^N(\mathbb{R}^{1+d}) \to T^N(\mathbb{R}^{1+d})$, which furthermore maps $G^N(\mathbb{R}^{1+d})$ into itself.

The following lemma moreover shows that T_v respects the Hopf algebra structure of $T((\mathbb{R}^{1+d}))$. Note that $T((\mathbb{R}^{1+d}))^{\otimes 2}$ embeds densely into $T(\mathbb{R}^{1+d})^{\overline{\otimes} 2}$, and thus $T_v \otimes T_v$ extends uniquely to a continuous algebra homomorphism $T(\mathbb{R}^{1+d})^{\overline{\otimes} 2} \to T(\mathbb{R}^{1+d})^{\overline{\otimes} 2}$.

Lemma 2.2.4. *The map* $T_v : T((\mathbb{R}^{1+d})) \to T((\mathbb{R}^{1+d}))$ *satisfies* $(T_v \otimes T_v)\Delta_{\sqcup\!\sqcup} = \Delta_{\sqcup\!\sqcup} T_v$ *and commutes with the antipode* α.

Proof. To show that $(T_v \otimes T_v)\Delta_{\sqcup\!\sqcup} = \Delta_{\sqcup\!\sqcup} T_v$, note that both $(T_v \otimes T_v)\Delta_{\sqcup\!\sqcup}$ and $\Delta_{\sqcup\!\sqcup} T_v$ are continuous algebra homomorphisms, and so they are equal provided they agree on $\mathtt{0}, \dots, \mathtt{d}$. Indeed, we have

$$\Delta_{\sqcup\!\sqcup} T_v(\mathtt{i}) = \Delta_{\sqcup\!\sqcup}(\mathtt{i} + v_i) = 1 \otimes (\mathtt{i} + v_i) + (\mathtt{i} + v_i) \otimes 1$$

(here we used that each v_i is a Lie element, i.e., primitive in the sense $\Delta_{\sqcup\!\sqcup} v_i = 1 \otimes v_i + v_i \otimes 1$) and

$$(T_v \otimes T_v)\Delta_{\sqcup\!\sqcup}(\mathtt{i}) = (T_v \otimes T_v)(1 \otimes \mathtt{i} + \mathtt{i} \otimes 1) = 1 \otimes (\mathtt{i} + v_i) + (\mathtt{i} + v_i) \otimes 1,$$

as required. It remains to show that T_v commutes with the antipode α. Actually, this is implied by general principles (e.g. [Pre16, Theorem 2.14], and the references therein), but as it is short to spell out, we give a direct argument: consider the opposite algebra $(T((\mathbb{R}^{1+d})))^{\mathrm{op}}$ (same set and vector space structure as $T((\mathbb{R}^{1+d}))$ but with reverse multiplication). Then $\alpha : T((\mathbb{R}^{1+d})) \to (T((\mathbb{R}^{1+d})))^{\mathrm{op}}$ is an algebra homomorphism, and again it suffices to check that αT_v and $T_v \alpha$ agree on $\mathtt{0}, \dots, \mathtt{d}$. Indeed, since $v_i \in \mathfrak{g}((\mathbb{R}^{d+1}))$, we have $\alpha(v_i) = -v_i$, and thus

$$\alpha T_v(\mathtt{i}) = \alpha(\mathtt{i} + v_i) = -\mathtt{i} - v_i$$

and

$$T_v \alpha(\mathtt{i}) = T_v(-\mathtt{i}) = -\mathtt{i} - v_i.$$

□

2.2.3 Dual action on the shuffle Hopf algebra $T(\mathbb{R}^{1+d})$

We now wish to describe the dual map $T_v^* : T(\mathbb{R}^{1+d}) \to T(\mathbb{R}^{1+d})$ for which

$$\langle T_v x, y \rangle = \langle x, T_v^* y \rangle, \quad \text{for all } x \in T((\mathbb{R}^{1+d})), \quad y \in T(\mathbb{R}^{1+d}).$$

We note immediately that Lemma 2.2.4 implies T_v^* is a Hopf algebra homomorphism from $(T(\mathbb{R}^{1+d}), \sqcup\!\sqcup, \Delta_\bullet, \tilde{\alpha})$ to itself.

For simplicity, and as this is the case most relevant to us, we only consider in detail the case $v = v_0 \in \mathfrak{g}((\mathbb{R}^{d+1}))$, i.e., $v_i = 0$ for $i = 1, \dots, d$ (but see Remark 2.2.6 for a description of the general case).

Let $\mathcal{S}$ denote the unital free commutative algebra generated by the non-empty words $\mathtt{i}_1 \cdots \mathtt{i}_k = \mathtt{i}_1 \bullet \dots \bullet \mathtt{i}_k$ in $T(\mathbb{R}^{1+d})$. We let $\mathbf{1}$ and $\cdot$ denote the unit element and product of $\mathcal{S}$ respectively. For example,

$$\mathtt{01} \cdot \mathtt{2} = \mathtt{2} \cdot \mathtt{01} \in \mathcal{S},$$
$$\mathtt{0} \cdot \mathtt{12} \neq \mathtt{0} \cdot \mathtt{21} \in \mathcal{S}.$$

For a word $w \in T(\mathbb{R}^{1+d})$, we let $D(w)$ denote the set of all elements

$$w_1 \cdot \ldots \cdot w_k \otimes \tilde{w} \in \mathcal{S} \otimes T(\mathbb{R}^{1+d})$$

where $w_1, \dots, w_k$ is formed from disjoint subwords of w and $\tilde{w}$ is the word obtained by replacing every w_i in w with e_0 (note that $\mathbf{1} \otimes w$, corresponding to $k = 0$, is also in $D(w)$).

Consider the linear map $S : T(\mathbb{R}^{1+d}) \to \mathcal{S} \otimes T(\mathbb{R}^{1+d})$ defined for all words $w \in T(\mathbb{R}^{1+d})$ by

$$S(w) = \sum_{w_1 \cdot \ldots \cdot w_k \otimes \tilde{w} \in D(w)} w_1 \cdot \ldots \cdot w_k \otimes \tilde{w}.$$

For example

$$\begin{aligned} S(\texttt{012}) = &\mathbf{1} \otimes \texttt{012} \\ &+ \texttt{0} \otimes \texttt{012} + \texttt{1} \otimes \texttt{002} + \texttt{2} \otimes \texttt{010} \\ &+ (\texttt{0} \cdot \texttt{1}) \otimes \texttt{002} + (\texttt{0} \cdot \texttt{2}) \otimes \texttt{010} + (\texttt{1} \cdot \texttt{2}) \otimes \texttt{000} \\ &+ (\texttt{0} \cdot \texttt{1} \cdot \texttt{2}) \otimes \texttt{000} + \texttt{01} \otimes \texttt{02} + \texttt{12} \otimes \texttt{00} \\ &+ (\texttt{01} \cdot \texttt{2}) \otimes \texttt{00} + (\texttt{0} \cdot \texttt{12}) \otimes \texttt{00} \\ &+ \texttt{012} \otimes \texttt{0}. \end{aligned}$$

Proposition 2.2.5. *Let $v = v_0 \in \mathfrak{g}((\mathbb{R}^{d+1}))$. The dual map $T_v^* : T(\mathbb{R}^{1+d}) \to T(\mathbb{R}^{1+d})$ is given by*

$$T_v^* w = (v \otimes \mathrm{id}) S(w),$$

where $v(w_1 \cdot \ldots \cdot w_k) := \langle w_1, v\rangle \ldots \langle w_k, v\rangle$ and $v(\mathbf{1}) := 1$.

In principle, Proposition 2.2.5 can be proved algebraically by showing that the adjoint of $\Phi := (v \otimes \mathrm{id})S$ is an algebra homomorphism from $T((\mathbb{R}^{1+d}))$ to itself, and check that $\Phi^*(\texttt{i}) = T_v(\texttt{i})$ for every generator $\texttt{i}$. Indeed this is the method used in Section 2.3.3 to prove the analogous result for the translation map on branched rough paths. However, in the current setting of weakly geometric rough paths, we can provide a direct combinatorial proof.

Proof. Note that the claim is equivalent to showing that for every two words $u, w \in T(\mathbb{R}^{1+d})$ (treating $u \in T((\mathbb{R}^{1+d}))$)

$$\langle T_v u, w\rangle = \sum_{w_1 \cdot \ldots \cdot w_k \otimes \tilde{w} \in D(w)} \langle w_1, v\rangle \ldots \langle w_k, v\rangle \langle \tilde{w}, u\rangle. \tag{2.4}$$

Consider a word $u = \texttt{i}_1 \bullet \cdots \bullet \texttt{i}_k \in T(\mathbb{R}^{1+d})$. Then

$$T_v(u) = \texttt{i}_1 \bullet \ldots \bullet (\texttt{0} + v) \bullet \ldots \bullet \texttt{i}_k,$$

where every occurrence of the letter $\texttt{0}$ in u is replaced by $\texttt{0} + v$. We readily deduce that for every $w \in T(\mathbb{R}^{1+d})$

$$\langle T_v(u), w\rangle = \sum_{\substack{w_1 \cdot \ldots \cdot w_k \otimes \tilde{w} \in D(w) \\ u = \tilde{w}}} \langle w_1, v\rangle \ldots \langle w_k, v\rangle. \tag{2.5}$$

For example, with $v = [\texttt{1}, \texttt{2}] = \texttt{12} - \texttt{21}$ and $u = \texttt{012}$, we have

$$T_v(u) = \texttt{012} + \texttt{1212} - \texttt{2112},$$

and we see that indeed for

$$w \in A := \{\texttt{012}, \texttt{1212}, \texttt{2112}\},$$

the right hand side of (2.5) gives $\langle T_v(u), w\rangle$, whilst $\langle w_1, v\rangle \ldots \langle w_k, v\rangle = 0$ for all words w which are not in A and $w_1 \cdot \ldots \cdot w_k \otimes \tilde{w} \in D(w)$ such that $u = \tilde{w}$. But now (2.5) immediately implies (2.4). □

Remark 2.2.6. A similar result to Proposition 2.2.5 holds for the general case $v = (v_0, \ldots, v_d)$. The definition of S changes in the obvious way that in the second tensor, instead of replacing every subword by the letter $\texttt{0}$, one instead replaces every combination of subwords by all combinations of $\texttt{i}$, $i \in \{0, \ldots, d\}$, while in the first tensor, one marks each extracted subword w_j with the corresponding label $i \in \{0, \ldots, d\}$ that replaced it, which gives $(w_j)_i$ (so the left tensor no longer belongs to $\mathcal{S}$ but instead to the free commutative algebra generated by $(w)_i$, for all words $w \in T(\mathbb{R}^{1+d})$ and labels $i \in \{0, \ldots, d\}$). Finally the term $\langle w_1, v\rangle \ldots \langle w_k, v\rangle$ would then be replaced by $\langle (w_1)_{i_1}, v_{i_1}\rangle \ldots \langle (w_k)_{i_k}, v_{i_k}\rangle$.

2.3 Translation of branched rough paths

In the previous section we studied the translation operator T, in the setting relevant for weakly geometric rough paths. Here we extend these results to the branched rough path setting, calling the translation operator M to avoid confusion. Our construction of M faces new difficulties, which we resolve with pre-Lie structures. The dual view then leads us to an extraction procedure of subtrees (a concept familiar from regularity structures, to be explored in Section 2.6).

2.3.1 Preliminaries for forest series

As in the preceding section, we first introduce the notation used throughout the section. Our setup closely follows Hairer-Kelly [HK15, Section 2]. (For additional algebraic background the reader can consult e.g. [GVF01, Chapter 14].)

Recall that a rooted tree is a finite connected graph without cycles with a distinguished node called the root. A rooted tree is unordered if there is no order on the edges leaving a node. We let $\mathcal{B} = \mathcal{B}(\bullet_0, ..., \bullet_d)$ denote the real vector space spanned by the set of unordered rooted trees with vertices labelled from the set $\{0, \dots, d\}$. We denote by $\mathcal{B}^*$ its algebraic dual, which we identify with the space of formal series of labelled trees; we write $\mathcal{B}^* = \mathcal{B}^*(\bullet_0, ..., \bullet_d)$ accordingly. We canonically identify with $\mathbb{R}^{1+d}$ the subspace of $\mathcal{B}$ (and of $\mathcal{B}^*$) spanned by the trees with a single node $\{\bullet_0, \dots, \bullet_d\}$.

We further denote by $\mathcal{H} = \mathcal{H}(\bullet_0, ..., \bullet_d)$ the vector space spanned by (unordered) forests composed of trees from $\mathcal{B}$ (including the empty forest denoted by 1), and let $\mathcal{H}^* = \mathcal{H}^*(\bullet_0, ..., \bullet_d)$ denote its algebraic dual which we identify with the space of formal series of forests. We canonically treat $\mathcal{B}^*$ as a subspace of $\mathcal{H}^*$. Following commonly used notation (e.g. [HK15, Section 2.2]), for trees $\tau_1, \dots, \tau_n \in \mathcal{B}$ we let $[\tau_1 \dots \tau_n]_{\bullet_i} \in \mathcal{B}$ denote the forest $\tau_1 \dots \tau_n \in \mathcal{H}$ grafted onto the node $\bullet_i$.

We equip $\mathcal{H}^*$ with the structure of the Grossman-Larson Hopf-type[4] algebra

$$(\mathcal{H}^*, \star, \Delta_\odot, \alpha)$$

and $\mathcal{H}$ with the structure of the dual graded Hopf algebra (the Connes-Kreimer Hopf algebra)

$$(\mathcal{H}, \odot, \Delta_\star, \tilde{\alpha}).$$

In other words, $\mathcal{H}$ is the free commutative algebra over $\mathcal{B}$ equipped with a coproduct $\Delta_\star$, and graded by the number of vertices in a forest. We shall often drop the product $\odot$ and simply write $\tau \odot \sigma = \tau\sigma$.

The coproduct $\Delta_\star$ may be described in terms of admissible cuts, for which we use the convention to keep the "trunk" on the right: for every tree $\tau \in \mathcal{B}$

$$\Delta_\star \tau = \sum_c \tau_1^c \dots \tau_k^c \otimes \tau_0^c,$$

where we sum over all admissible cuts c of τ, and denote by τ_0^c the trunk and by $\tau_1^c \dots \tau_k^c$ the branches of the cut respectively.

In the sequel, we shall also find it convenient to treat the space $\mathcal{H}$ equipped with $\star$ as a subalgebra of $\mathcal{H}^*$, in which case we explicitly refer to it as the algebra $(\mathcal{H}, \star)$.

[4] Again, $\Delta_\odot$ does not map $\mathcal{H}^*$ into $\mathcal{H}^{*\otimes 2}$, but instead into the complete tensor product $\mathcal{H}^{\bar{\otimes} 2} \simeq \prod_{k,m=0}^\infty \mathcal{H}^{(k)} \otimes \mathcal{H}^{(m)}$, where $\mathcal{H}^{(k)}$ denotes the vector space of forests with k vertices, and therefore the structure is not exactly that of a Hopf algebra. Note also that $\Delta_\odot$ is continuous for the product topologies, which we equip $\mathcal{H}^*$ and $\mathcal{H}^{\bar{\otimes} 2}$ with henceforth.

Recall that the space of series $\mathcal{B}^*$ is exactly the set of primitive elements of $\mathcal{H}^*$. We let $\mathcal{G} = \mathcal{G}(\bullet_0, ..., \bullet_d)$ denote the group-like elements of $\mathcal{H}^*$, often called the Butcher group, for which it holds that

$$\mathcal{G} = \exp_\star(\mathcal{B}^*).$$

All the objects introduced above play an analogous role to those of the previous section. To summarise this correspondence, it is helpful to keep the following picture in mind

"Series space" ...	$\mathcal{H}^*(\bullet_0, ..., \bullet_d) \equiv \mathcal{H}^*$	$\longleftrightarrow$	$T((\mathbb{R}^{1+d}))$
"Polynomial space" ...	$\mathcal{H}(\bullet_0, ..., \bullet_d) \equiv \mathcal{H}$	$\longleftrightarrow$	$T(\mathbb{R}^{1+d})$
Lie elements ...	$\mathcal{B}^*(\bullet_0, ..., \bullet_d) \equiv \mathcal{B}^* \subset \mathcal{H}^*$	$\longleftrightarrow$	$\mathfrak{g}((\mathbb{R}^{d+1}))$
Group-like elements ...	$\mathcal{G}(\bullet_0, ..., \bullet_d) \equiv \mathcal{G} \subset \mathcal{H}^*$	$\longleftrightarrow$	$G(\mathbb{R}^{1+d})$.

As in the previous section, for any integer $N \geq 0$ we let $\mathcal{H}^N$ denote "truncated" algebra obtained by the quotient of $\mathcal{H}^*$ by the ideal consisting of all series with no forests having less than N vertices (we keep in mind that the product $\star$ is always in place for the truncated objects). Similarly, we let $\mathcal{G}^N \subset \mathcal{H}^N$ and $\mathcal{B}^N \subset \mathcal{H}^N$ denote the step-N Butcher Lie group over $\mathbb{R}^{1+d}$ its and Lie algebra respectively, constructed in analogous ways.

Finally, as before, we write "$(\mathbb{R}^d)$" to denote the analogous objects built over $\mathbb{R}^d$, treated as subsets of their "full" counterparts built over $\mathbb{R}^{1+d}$ (by identifying $\mathbb{R}^d$ with the subspace of $\mathbb{R}^{1+d}$ with basis $\{e_1, \ldots, e_d\}$). For example, we treat $\mathcal{H}^*(\mathbb{R}^d)$ and $\mathcal{B}^N(\mathbb{R}^d)$ as a subalgebra of $\mathcal{H}^*$ and a Lie subalgebra of $\mathcal{B}^N$ respectively.

2.3.2 Translation of forest series

2.3.2.1 Non-uniqueness of algebra extensions

In the previous section, we defined a map T_v which "translated" elements in $T((\mathbb{R}^{1+d}))$ in directions $(v_0, \ldots, v_d) \subset \mathfrak{g}((\mathbb{R}^{d+1}))$, and which mapped the set of group-like elements $G(\mathbb{R}^{1+d})$ into itself. In the same spirit, we aim to define a map M_v which translates elements in $\mathcal{H}^*$ in directions $(v_0, \ldots, v_d) \subset \mathcal{B}^*$, and which likewise maps $\mathcal{G}$ into itself.

Note that our construction of T_v relied on the fact that any linear map $M : \mathbb{R}^{1+d} \to T((\mathbb{R}^{1+d}))$ such that $\langle Mv, 1\rangle = 0$ for $v \in \mathbb{R}^{1+d}$ extended uniquely to a continuous algebra homomorphism $M : T((\mathbb{R}^{1+d})) \to T((\mathbb{R}^{1+d}))$ (for the product $\bullet$). We note here that no such universal property holds for $\mathcal{H}^*$; indeed, there exists a canonical injective algebra homomorphism

$$\begin{aligned} \mathfrak{a} &: T((\mathbb{R}^{1+d})) \to \mathcal{H}^* \\ \mathfrak{a} &: \mathtt{i} \mapsto \bullet_i \end{aligned} \tag{2.6}$$

which embeds $T((\mathbb{R}^{1+d}))$ into a *strict* subalgebra of $\mathcal{H}^*$.

Specifically, we can see that $\mathfrak{a}$ is injective by considering the space $\mathcal{B}^*_\ell \subset \mathcal{B}^*$ of linear trees, i.e., trees of the form $[\ldots[\bullet_{i_1}]_{\bullet_{i_2}}]\ldots]_{\bullet_{i_k}}$. Then there is a natural projection $\pi_\ell : \mathcal{H}^* \to \mathcal{B}^*_\ell$, and one can readily see that $\pi_\ell \circ \mathfrak{a}$ is a bijective linear map (this is the same map as described in [HK15, Remark 2.7]). To see further that the image of $T((\mathbb{R}^{1+d}))$ under $\mathfrak{a}$ is not all of $\mathcal{H}^*$, it suffices to observe that the linear tree $[\bullet_i]_{\bullet_j}$ is not in the algebra generated by $\{\bullet_i\}_{i=1}^{1+d}$.

Remark 2.3.1. The embedding $\mathfrak{a}$ arises naturally in the context of branched rough paths as this is essentially the embedding $\imath$ used in [HK15, Section 4.1 and Section 5.1] to realise weakly geometric rough paths as branched rough paths (though note $\imath$ in [HK15] denotes $\pi_\ell \circ \mathfrak{a}$ in our notation).

Remark 2.3.2. While the above argument shows that $(\mathcal{B}, [\cdot,\cdot])$ is clearly not isomorphic to $\mathfrak{g}(\mathbb{R}^{d+1})$ as a Lie algebra, it is a curious and non-trivial fact that $(\mathcal{B}, [\cdot,\cdot])$ *is* isomorphic to a free Lie algebra generated by another subspace of $\mathcal{B}$. Correspondingly, $(\mathcal{H}, \star)$, being isomorphic to the universal enveloping algebra of $\mathcal{B}$, is isomorphic to a tensor algebra (see [Foi02, Section 8]; see also [Cha10, Section 6], where the statement about free pre Lie algebras is Corollary 6.3). This was used in [BC19] to show that the space of branched p-rough paths is canonically isomorphic to a space of weakly geometric Π-rough paths over an enlarged vector space (see [BC19, Theorem 4.3] for the statement about the isomorphism between branched and weakly geometric rough paths).

It follows from the above discussion that given a map $M : \mathbb{R}^{1+d} \to \mathcal{H}^*$, even one whose range is in $\mathcal{B}^*$, there is in general no canonical choice of how to extend M to elements outside $\mathfrak{a}(T((\mathbb{R}^{1+d})))$ if we only demand that the extension $M : \mathcal{H}^* \to \mathcal{H}^*$ is an algebra morphism (moreover, without calling on Remark 2.3.2, it is *a priori* not even clear that such an extension always exists).

Example 2.3.3. *Consider the case of a single label* 0 *(i.e.* $d = 0$*), and the map* $M : \{\bullet_0, [\bullet_0]_{\bullet_0}\} \to \mathcal{B}^*$ *given by*

$$M : \bullet_0 \mapsto \bullet_0$$
$$M : [\bullet_0]_{\bullet_0} \mapsto \bullet_0.$$

Since

$$\bullet_0 \star \bullet_0 = [\bullet_0]_{\bullet_0} + 2 \bullet_0 \bullet_0,$$

we may extend M *to an algebra homomorphism on the truncated space* $\mathcal{H}^2 \to \mathcal{H}^2$ *by setting*

$$M(\bullet_0 \bullet_0) = \frac{1}{2}\left([\bullet_0]_{\bullet_0} + 2 \bullet_0 \bullet_0 - \bullet_0\right).$$

This example shows that, on the level of the truncated algebras, there is not a unique algebra homomorphism above the identity map $\mathrm{id} : \bullet_0 \mapsto \bullet_0$*.*

Of course, it is not clear from the above that the identity map $\mathrm{id} : \bullet_0 \mapsto \bullet_0$ *can extend in a non-trivial way to an algebra homomorphism on all of* $\mathcal{H}^* \mapsto \mathcal{H}^*$*, but such extensions will always exist due to Remark 2.3.2.*

In what follows, we address this non-uniqueness issue by demanding a finer structure on the extension of M, namely that $M : \mathcal{B}^* \to \mathcal{B}^*$ is a *pre-Lie algebra* homomorphism. The notion of a pre-Lie algebra will be recalled in the following subsection, and the significance of preserving the pre-Lie product on $\mathcal{B}^*$ is first seen when establishing a dual characterization of M (Proposition 2.3.14), and then again in Section 2.5.2 when studying the impact on (rough) differential equations. For now, we simply state that this is a natural condition to demand given the role of pre-Lie algebras in control theory and Butcher series ([CEM11, Section 10],[Man11, Section 3.4]).

2.3.2.2 The free pre-Lie algebra over $\mathbb{R}^{1+d}$

Definition 2.3.4. *A (left)* pre-Lie algebra *is a vector space* V *with a bilinear map* $\triangleright : V \times V \to V$*, called the* pre-Lie product*, such that*

$$(x \triangleright y) \triangleright z - x \triangleright (y \triangleright z) = (y \triangleright x) \triangleright z - y \triangleright (x \triangleright z), \quad \text{for all } x, y, z \in V.$$

That is, the associator $(x, y, z) := (x \triangleright y) \triangleright z - x \triangleright (y \triangleright z)$ *is invariant under exchanging* x *and* y*.*

One can readily check that every pre-Lie algebra $(V, \triangleright)$ induces a Lie algebra $(V, [\cdot, \cdot])$ consisting of the same vector space V with bracket $[x, y] := x \triangleright y - y \triangleright x$.

Example 2.3.5. *A basic example of a pre-Lie algebra is the space of smooth vector fields on $\mathbb{R}^e$ with the product $(f_i \partial_i) \blacktriangleright (f_j \partial_j) := (f_i \partial_i f_j)\partial_j$. The induced bracket is the usual Lie bracket of vector fields.*

The space of trees $\mathcal{B}$ can be equipped with a (non-associative) pre-Lie product $\curvearrowright : \mathcal{B} \times \mathcal{B} \to \mathcal{B}$ defined by

$$\tau_1 \curvearrowright \tau_2 = \sum_{\tau} n(\tau_1, \tau_2, \tau)\tau, \tag{2.7}$$

where the sum is over all trees $\tau \in \mathcal{B}$ and $n(\tau_1, \tau_2, \tau)$ is the number of *single* admissible cuts of τ for which the branch is τ_1 and the trunk is τ_2. Equivalently, $\curvearrowright$ is given in terms of $\star$ by

$$\tau_1 \curvearrowright \tau_2 = \pi_{\mathcal{B}}(\tau_1 \star \tau_2), \tag{2.8}$$

where $\pi_{\mathcal{B}} : \mathcal{H} \to \mathcal{B}$ is the projection onto $\mathcal{B}$.

It holds that $(\mathcal{B}, \curvearrowright)$ indeed defines a Lie algebra for which

$$[\tau_1, \tau_2] := \tau_1 \curvearrowright \tau_2 - \tau_2 \curvearrowright \tau_1 = \tau_1 \star \tau_2 - \tau_2 \star \tau_1,$$

i.e., the Lie algebra structures on $\mathcal{B}$ induced by $\star$ and $\curvearrowright$ coincide. Moreover since $\curvearrowright$ respects the grading of $\mathcal{B}$, we can naturally extend $\curvearrowright$ to a bilinear map on the space of series, so that $(\mathcal{B}^*, \curvearrowright)$ is also a pre-Lie algebra.

We now recall the following universal property of $(\mathcal{B}, \curvearrowright)$ first established by Chapoton and Livernet in [CL01, Corollary 1.10] (see also [DL02, Theorem 6.3]).

Theorem 2.3.6. *The space $(\mathcal{B}, \curvearrowright)$ is the free pre-Lie algebra over $\mathbb{R}^{1+d}$.*

An equivalent formulation of Theorem 2.3.6 is that for any pre-Lie algebra $(V, \triangleright)$ and linear map $M : \mathbb{R}^{1+d} \to V$, there exists a unique extension of M to a pre-Lie algebra homomorphism $M : (\mathcal{B}, \curvearrowright) \to (V, \triangleright)$.

2.3.2.3 Construction of the translation map

An immediate consequence of Theorem 2.3.6 is the following.

Theorem 2.3.7. *Every linear map $M : \mathbb{R}^{1+d} \to \mathcal{B}^*$ extends to a unique continuous algebra homomorphism $M : \mathcal{H}^* \to \mathcal{H}^*$ whose restriction to $\mathcal{B}^*$ is a pre-Lie algebra homomorphism from $\mathcal{B}^*$ to itself.*

Proof. By Theorem 2.3.6, M extends uniquely to a pre-Lie algebra homomorphism $M : \mathcal{B} \to \mathcal{B}^*$. Recall also that, by the Milnor-Moore theorem, $(\mathcal{H}, \star)$ is isomorphic to the universal enveloping algebra of $(\mathcal{B}, [\cdot, \cdot])$. It follows that M extends further to a unique algebra homomorphism $M : (\mathcal{H}, \star) \to (\mathcal{H}^*, \star)$. Finally, since M necessarily does not decrease the degree of every element $x \in \mathcal{H}$, we obtain a unique continuous extension $M : \mathcal{H}^* \to \mathcal{H}^*$ for which the restriction $M : \mathcal{B}^* \to \mathcal{B}^*$ is a pre-Lie algebra homomorphism as desired. □

We can finally define a natural translation map $M_v : \mathcal{H}^* \to \mathcal{H}^*$ analogous to T_v.

Definition 2.3.8. *For $v = (v_0, \ldots, v_d) \subset \mathcal{B}^*$, define $M_v : \mathcal{H}^* \to \mathcal{H}^*$ as the unique continuous algebra homomorphism obtained in Theorem 2.3.7 from the linear map*

$$M_v : \mathbb{R}^{1+d} \to \mathcal{B}^*$$
$$M_v : \bullet_i \mapsto \bullet_i + v_i, \quad \textit{for all } i \in \{0, \ldots, d\}.$$

Note: Do not confuse M_v with the map M_p from Definition 3.3.1, they are different in nature and not even defined on the same space!

Example 2.3.9. *Let us illustrate how the construction works in the case of two nodes with a single label* 0*. Since* M_v *is constructed as pre-Lie algebra homomorphism, we compute*

$$M_v\left([\bullet_0]_{\bullet_0}\right) = M_v\left(\bullet_0 \curvearrowright \bullet_0\right) = M_v\left(\bullet_0\right) \curvearrowright M_v\left(\bullet_0\right) = \left(\bullet_0 + v_0\right) \curvearrowright \left(\bullet_0 + v_0\right).$$

Since M_v *is in addition an algebra homomorphism w.r.t.* $\star$ *we have*

$$(\bullet_0 + v_0) \star (\bullet_0 + v_0) = (M_v \bullet_0) \star (M_v \bullet_0) = M_v\left(\bullet_0 \star \bullet_0\right) = M_v\left(2\,\bullet_0\,\bullet_0 + [\bullet_0]_{\bullet_0}\right)$$

from which we can uniquely determine $M_v\left(\bullet_0 \bullet_0\right)$.

As in the previous section, we shall often be concerned with the case that $v_i = 0$ for $i = 1, \ldots, d$ and v_0 takes a special form. We again make precise whenever such a condition is in place by writing, for example, $v = v_0 \in \mathcal{B}^N(\mathbb{R}^d)$.

Remark 2.3.10. We observe the following immediate properties of M_v, analogous to those of T_v:

1. Since M_v is an algebra homomorphism which preserves the Lie algebra $\mathcal{B}^*$, it holds that M_v maps $\mathcal{G}$ into $\mathcal{G}$;
2. $M_v \circ M_u = M_{v+M_v(u)}$, where we write $M_v(u) = (M_v(u_0), \ldots, M_v(u_d))$. In particular, $M_{v+u} = M_v \circ M_u$ for all $v = v_0, u = u_0 \in \mathcal{B}^*(\mathbb{R}^d)$; i.e., $\{M_v | v = v_0 \in \mathcal{B}(\mathbb{R}^d)^*\}$ forms an abelian group isomorphic to $(\mathcal{B}(\mathbb{R}^d)^*, +)$.
3. We furthermore have that $\{M_v | v_i \in \mathcal{B}^*_{\geq 2}(\mathbb{R}^{d+1})\, \forall i\}$ forms a non-abelian group. Indeed, we may construct the inverse of M_v as $M_{\bar{v}}$ with
$$\bar{v} = -v - \sum_{n=1}^{\infty} (\mathrm{id} - M_v)^{\circ n}(v).$$
4. For every integer $N \geq 0$, M_v induces a well-defined algebra homomorphism $M_v^N : \mathcal{H}^N \to \mathcal{H}^N$, which maps $\mathcal{G}^N$ into $\mathcal{G}^N$;
5. Recall the embedding $\mathfrak{a} : T((\mathbb{R}^{1+d})) \to \mathcal{H}^*$ from (2.6). Then for all $v = (v_0, \ldots, v_d) \subset \mathfrak{g}((\mathbb{R}^{d+1}))$, it holds that $M_{\mathfrak{a}(v)} \circ \mathfrak{a} = \mathfrak{a} \circ T_v$ (as both are continuous algebra homomorphisms from $T((\mathbb{R}^{1+d}))$ to $\mathcal{H}^*$ which agree on $e_0, \ldots, e_d$).

As in the remark before Lemma 2.2.4, note that $\mathcal{H}^{*\otimes 2}$ embeds densely into $\mathcal{H}^{\overline{\otimes} 2}$, and thus $M_v \otimes M_v$ extends uniquely to a continuous algebra homomorphism $\mathcal{H}^{\overline{\otimes} 2} \to \mathcal{H}^{\overline{\otimes} 2}$.

Lemma 2.3.11. *The map* $M_v : \mathcal{H}^* \to \mathcal{H}^*$ *satisfies* $(M_v \otimes M_v)\Delta_{\odot} = \Delta_{\odot} M_v$ *and commutes with the antipode* α.

Remark 2.3.12. We note that in the following proof, we only use the fact that M_v is a continuous algebra homomorphism from $\mathcal{H}^*$ to itself which preserves the space of primitive elements $\mathcal{B}^*$, and so do not directly use the fact that M_v preserves the pre-Lie product of $\mathcal{B}^*$.

Proof. To show that the maps $(M_v \otimes M_v)\Delta_{\odot}$ and $\Delta_{\odot} M_v$ agree, by continuity it suffices to show they agree on $\mathcal{H}$. In turn, their restrictions to $\mathcal{H}$ are algebra homomorphisms on $(\mathcal{H}, \star)$, and, since $(\mathcal{H}, \star)$ is the universal enveloping algebra of its space of primitive elements $\mathcal{B}$ by the Milnor-Moore theorem, it suffices to show that

$$(M_v \otimes M_v)\Delta_{\odot}\tau = \Delta_{\odot} M_v \tau, \quad \text{for all } \tau \in \mathcal{B}.$$

But this is immediate since M_v maps $\mathcal{B}^*$ into itself and $M_v(1) = 1$. It remains to show that M_v commutes with the antipode, which follows from the same argument as in the proof of Lemma 2.2.4. □

2.3.3 Dual action on the Connes–Kreimer Hopf algebra $\mathcal{H}$

As in Section 2.2.3, we now wish to describe the dual map $M_v^* : \mathcal{H} \to \mathcal{H}$ for which

$$\langle M_v x, y\rangle = \langle x, M_v^* y\rangle, \quad \text{for all } x \in \mathcal{H}^*, \quad y \in \mathcal{H}.$$

For simplicity, we again consider in detail only the special case $v_i = 0$ for $i = 1, \ldots, d$ (but see Remark 2.3.16 for a description of the general case).

Let $\mathcal{A}$ denote the unital free commutative algebra generated by the trees $\tau \in \mathcal{B}$. We let $\mathbf{1}$ and $\cdot$ denote the unit element and product of $\mathcal{A}$ respectively. The algebra $\mathcal{A}$ plays here the same role as the algebra $\mathcal{S}$ in Section 2.2.3.

Remark 2.3.13. Although the algebras $(\mathcal{A}, \cdot)$ and $(\mathcal{H}, \odot)$ are isomorphic, they should be thought of as separate spaces and thus we make a clear distinction between the two.

For a tree $\tau \in \mathcal{B}$, we let $D(\tau)$ denote the set of all elements

$$\tau_1 \cdot \ldots \cdot \tau_k \otimes \tilde{\tau} \in \mathcal{A} \otimes \mathcal{B}$$

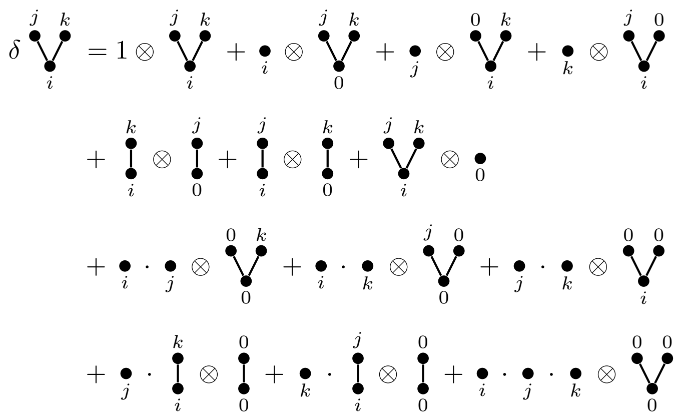
where $\tau_1, \ldots, \tau_k$ is formed from all disjoint collections of non-empty subtrees of τ (including subtrees consisting of a single node), and $\tilde{\tau}$ is the tree obtained by contracting every subtree τ_i to a single node which is then labelled by 0 (note that $\mathbf{1} \otimes \tau$, corresponding to $k = 0$, is also in $D(\tau)$).

Consider the linear map $\delta : \mathcal{H} \to \mathcal{A} \otimes \mathcal{H}$ defined for all trees $\tau \in \mathcal{B}$ by

$$\delta\tau = \sum_{\tau_1 \cdot \ldots \cdot \tau_k \otimes \tilde{\tau} \in D(\tau)} \tau_1 \cdot \ldots \cdot \tau_k \otimes \tilde{\tau},$$

and then extended multiplicatively to all of $\mathcal{H}$, where we canonically treat $\mathcal{A} \otimes \mathcal{H}$ as an algebra with multiplication $\mathcal{M}_{\mathcal{A}\otimes\mathcal{H}}(\tau_1 \otimes \hat{\tau}_1 \otimes \tau_2 \otimes \hat{\tau}_2) := (\tau_1 \cdot \tau_2) \otimes (\hat{\tau}_1 \odot \hat{\tau}_2)$ for $\tau_1, \tau_2 \in \mathcal{A}$, $\hat{\tau}_1, \hat{\tau}_2 \in \mathcal{H}$.

For example,



Proposition 2.3.14. *Let $v = v_0 \in \mathcal{B}^*$. The dual map $M_v^* : \mathcal{H} \to \mathcal{H}$ is given by*

$$M_v^* \tau = (v \otimes \mathrm{id}) \circ \delta(\tau),$$

where $v(\tau_1 \cdot \ldots \cdot \tau_k) := \langle \tau_1, v\rangle \ldots \langle \tau_k, v\rangle$ and $v(\mathbf{1}) := 1$.

For the proof of Proposition 2.3.14, we require the following combinatorial lemma. We note that similar "cointeraction" results appear for closely related algebraic structures in [CEM11, Theorem 8] and [BHZ19, Proposition 3.27 and Theorem 5.37]. We will particularly discuss in further detail the link with the work of [BHZ19] in Section 2.6.

Lemma 2.3.15. *Let $\Delta_{\curvearrowright} : \mathcal{B} \to \mathcal{B} \otimes \mathcal{B}$ denote the adjoint of $\curvearrowright$. It holds that*

$$(\mathrm{id} \otimes \Delta_{\curvearrowright})\delta = \mathcal{M}_{1,3}(\delta \otimes \delta)\, \Delta_{\curvearrowright}, \tag{2.9}$$

where $\mathcal{M}_{1,3} : \mathcal{A} \otimes \mathcal{B} \otimes \mathcal{A} \otimes \mathcal{B} \to \mathcal{A} \otimes \mathcal{B} \otimes \mathcal{B}$ is the linear map defined by $\mathcal{M}_{1,3}(\tau_1 \otimes \tau_2 \otimes \tau_3 \otimes \tau_4) = \tau_1\tau_3 \otimes \tau_2 \otimes \tau_4$.

Proof. Note that

$$\Delta_{\curvearrowright} \tau = \sum_c b_c \otimes \tau_c$$

where the sum runs of all *single* admissible cuts c of τ, and b_c is the branch, τ_c the trunk of c. Consider a single cut c of τ across an edge e. Let τ^c denote the sum of the terms of $(\mathrm{id} \otimes \Delta_{\curvearrowright})\delta\tau$ obtained by contracting all collections of subtrees of τ which do not contain e, followed by a cut (on the second tensor) along the edge e (which necessarily remains). One immediately sees that τ^c is equivalently given by first cutting along e, and then contracting along all collections of subtrees of b_c and τ_c, and then grouping the extracted subtrees together, i.e., $\tau^c = \mathcal{M}_{1,3}(\delta \otimes \delta)(b_c \otimes \tau_c)$. It finally remains to observe that summing over all single cuts c gives (2.9). □

Proof of Proposition 2.3.14. Denote by

$$\Phi = (v \otimes \mathrm{id}) \circ \delta : \mathcal{B} \to \mathcal{B}.$$

By duality, it follows from Lemma 2.3.11 that M_v^* is a Hopf algebra homomorphism. In particular, it suffices to show that $\Phi\tau = M_v^*\tau$ for every tree $\tau \in \mathcal{B}$.

To this end, observe that Lemma 2.3.15 implies $\Delta_{\curvearrowright} \Phi = (\Phi \otimes \Phi)\, \Delta_{\curvearrowright}$, from which it follows that $\Phi^* : \mathcal{B}^* \to \mathcal{B}^*$ is a pre-Lie algebra homomorphism. Furthermore, for every tree $\tau \in \mathcal{B}$

$$\text{for all } i \in \{1, \ldots, d\},\ \langle \Phi^* \bullet_i, \tau\rangle = \langle \bullet_i, \Phi\tau\rangle = \langle \bullet_i, \tau\rangle = \langle M_v \bullet_i, \tau\rangle;$$
$$\langle \Phi^* \bullet_0, \tau\rangle = \langle \bullet_0, \Phi\tau\rangle = \langle \bullet_0, \tau\rangle + \langle v, \tau\rangle = \langle M_v \bullet_0, \tau\rangle.$$

It follows that Φ^* is a pre-Lie algebra homomorphism on $(\mathcal{B}^*, \curvearrowright)$ which agrees with M_v on the set $\{\bullet_0, \ldots, \bullet_d\} \subset \mathcal{B}^*$. Hence, by the universal property of $(\mathcal{B}, \curvearrowright)$ (Theorem 2.3.6), Φ^* agrees with M_v on all of $\mathcal{B}^*$, which concludes the proof. □

Remark 2.3.16. A similar result to Proposition 2.3.14 holds for the general case $v = (v_0, \ldots, v_d)$. The definition of δ changes in the obvious way that in the second tensor, instead of replacing every subtree by the node $\bullet_0$, one instead replaces every combination of subtrees by all combinations of $\bullet_i$, $i \in \{0, \ldots, d\}$, while in the first tensor, one marks each extracted subtree τ_j with the corresponding label $i \in \{0, \ldots, d\}$ that replaced it, which gives $(\tau_j)_i$ (so the left tensor no longer belongs to $\mathcal{A}$ but instead to the free commutative algebra generated by $(\tau)_i$, for all trees $\tau \in \mathcal{B}$ and labels $i \in \{0, \ldots, d\}$). Finally the term $\langle \tau_1, v\rangle \ldots \langle \tau_k, v\rangle$ would then be replaced by $\langle (\tau_1)_{i_1}, v_{i_1}\rangle \ldots \langle (\tau_k)_{i_k}, v_{i_k}\rangle$.

2.4 Examples

In the following examples, we assume that we are given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual hypotheses and to which all mentioned stochastic processes are adapted.

2.4.1 Itô-Stratonovich conversion

As an application of Proposition 2.3.14, we illustrate how to re-express iterated Stratonovich integrals (and products thereof) over some interval $[s,t]$ as Itô integrals. Consider the $\mathbb{R}^{1+d}$-valued process $B_t = (B^0_t, B^1_t, \dots, B^d_t)$, where $(B^1_t, \dots, B^d_t)$ is an $\mathbb{R}^d$-valued Brownian motion with covariance $[B^i, B^j]_t = C^{i,j}t$, and $B^0_t \equiv t$ denotes the time component. Let $\mathbf{B}^{\text{Strat}}$ denote the enhancement of B_t to an α-Hölder branched rough path, $\alpha \in (0, 1/2)$, using Stratonovich iterated integrals. For example,

$$\begin{aligned} \langle \mathbf{B}^{\text{Strat}}_{s,t}, \tau \rangle &= \int \cdots \int_{s<t_1<\cdots<t_m<t} \circ\, dB^{i_1}_{t_1} \circ \cdots \circ dB^{i_m}_{t_m} \\ \text{for the linear tree } \tau &= [\dots[\bullet_{i_1}]_{\bullet_{i_2}} \dots]_{\bullet_{i_m}}, \quad i_1, \dots, i_m \in \{0, \dots, d\}, \end{aligned} \tag{2.10}$$

and

$$\begin{aligned} \langle \mathbf{B}^{\text{Strat}}_{s,t}, \tau \rangle &= \int_s^t B^j_u B^k_u \circ dB^i_u \\ \text{for } \tau &= [\bullet_j \bullet_k]_{\bullet_i}, \quad i,j,k \in \{0, \dots, d\}. \end{aligned}$$

Similarly, we define $\mathbf{B}^{\text{Itô}}$ in exactly the same way using Itô integrals.

For a tree $\tau \in \mathcal{B}$, recall the definition of $D(\tau) \subset \mathcal{A} \otimes \mathcal{B}$ from Section 2.3.3 (which was used to define δ). Consider the function $C : D(\tau) \to \mathbb{R}$ defined by

$$C(\tau_1 \cdot \ldots \cdot \tau_k \otimes \tilde{\tau}) = \begin{cases} 1 & \text{if } \tau_1 \cdot \ldots \cdot \tau_k \otimes \tilde{\tau} = \mathbf{1} \otimes \tau \\ 2^{-k} \prod_{n=1}^k C^{i_n, j_n}_n & \text{if } \tau_n = [\bullet_{i_n}]_{\bullet_{j_n}} \text{ for all } n = 1, \dots, k \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 2.4.1. *For every tree $\tau \in \mathcal{B}$ it holds that*

$$\langle \mathbf{B}^{\text{Strat}}_{s,t}, \tau \rangle = \sum_{\tau_1 \cdot \ldots \cdot \tau_k \otimes \tilde{\tau} \in D(\tau)} C(\tau_1 \cdot \ldots \cdot \tau_k \otimes \tilde{\tau}) \langle \mathbf{B}^{\text{Itô}}_{s,t}, \tilde{\tau} \rangle. \tag{2.11}$$

Proof. Consider the sum of linear trees $v = v_0 = \frac{1}{2} \sum_{i,j=1}^d C^{i,j} [\bullet_i]_{\bullet_j} \in \mathcal{B}^2(\mathbb{R}^d)$. One can readily verify that $\mathbf{B}^{\text{Strat}} = M_v(\mathbf{B}^{\text{Itô}})$, understood in the pointwise sense $\mathbf{B}^{\text{Strat}}_{s,t} = M_v(\mathbf{B}^{\text{Itô}}_{s,t})$. Indeed, both $\mathbf{B}^{\text{Strat}}$ and $M_v(\mathbf{B}^{\text{Itô}})$ are a.s. "full" α-Hölder rough paths, where this fact - in the case of $M_v(\mathbf{B}^{\text{Itô}})$ - either requires an (easy) check by hand, or an appeal to Theorem 2.5.1, (ii), below. Since, by construction, both agree on the first two levels, and $\alpha \in (1/2, 1/3)$, we see that $\mathbf{B}^{\text{Strat}}$ and $M_v(\mathbf{B}^{\text{Itô}})$ must be equal, a.s., thanks to the uniqueness part of the extension theorem.

It then follows by Proposition 2.3.14 that

$$\langle \mathbf{B}^{\text{Strat}}_{s,t}, \tau \rangle = \langle \mathbf{B}^{\text{Itô}}_{s,t}, M^*_v \tau \rangle = \sum_{\tau_1 \cdot \ldots \cdot \tau_k \otimes \tilde{\tau} \in D(\tau)} \langle \mathbf{B}^{\text{Itô}}_{s,t}, \langle v, \tau_1 \rangle \dots \langle v, \tau_k \rangle \tilde{\tau} \rangle.$$

Since $\langle v, \mathbf{1} \rangle = 1$, while $\langle v, \tau_n \rangle = \frac{1}{2} C^{i,j}$ if $\tau_n = [\bullet_i]_{\bullet_j}$ and zero otherwise, we obtain precisely (2.11). □

Example 2.4.2. *Suppose B is a standard Brownian motion, i.e., $C^{i,j} = \delta_{ij}$. Consider the tree $\tau = [\bullet_j \bullet_k]_{\bullet_i}$, so that*

$$\langle \mathbf{B}^{\mathrm{Strat}}_{s,t}, \tau \rangle = \int_s^t B^j_u B^k_u \circ dB^i_u.$$

*Recalling the explicit form of $\delta\tau$ in (**??**), we see that if i is distinct from both j, k, then only $\mathbf{1} \otimes \tau$ remains in $D(\tau)$ for which C is non-zero, and so (in trivial agreement with stochastic calculus)*

$$\langle \mathbf{B}^{\mathrm{Strat}}_{s,t}, \tau \rangle = \langle \mathbf{B}^{\mathrm{It\hat{o}}}_{s,t}, \tau \rangle.$$

On the other hand,if $i = j \neq k$, an additional term $[\bullet_i]_{\bullet_i} \otimes [\bullet_k]_{\bullet_0}$ appears in $D(\tau)$ at which C is $\frac{1}{2}$, and so

$$\begin{aligned} \langle \mathbf{B}^{\mathrm{Strat}}_{s,t}, \tau \rangle &= \langle \mathbf{B}^{\mathrm{It\hat{o}}}_{s,t}, \tau \rangle + \frac{1}{2} \iint_{s<t_1<t_2<t} dB^k_{t_1} dB^0_{t_2} \\ &= \langle \mathbf{B}^{\mathrm{It\hat{o}}}_{s,t}, \tau \rangle + \frac{1}{2} \int_s^t B^k_u du. \end{aligned}$$

The case $i = k \neq j$ is identical. At last, in the case $i = j = k$, looking at $\delta\tau$ shows that

$$\begin{aligned} \langle \mathbf{B}^{\mathrm{Strat}}, \tau \rangle &= \langle \mathbf{B}^{\mathrm{It\hat{o}}}, \tau \rangle + \frac{1}{2} \iint_{s<t_1<t_2<t} dB^i_{t_1} dB^0_{t_2} + \frac{1}{2} \iint_{s<t_1<t_2<t} dB^i_{t_1} dB^0_{t_2} \\ &= \langle \mathbf{B}^{\mathrm{It\hat{o}}}, \tau \rangle + \int_s^t B^i_u du. \end{aligned}$$

Remark 2.4.3. When $\tau = [\ldots[\bullet_{i_1}]_{\bullet_{i_2}} \ldots]_{\bullet_{i_m}}$ is a linear tree, this is in agreement with [Ben89, Proposition 1]. In fact, by considering general semi-martingales $X^1_t, \ldots, X^d_t$ and adding extra labels $\bullet_{i,j}$, $1 \leq i \leq j \leq d$ (thus increasing the underlying dimension from d to $d + d(d+1)/2$) to encode the quadratic variants $[X_i, X_j]$, the above procedure (in the more general setting with elements $v_{ij} = [\bullet_i]_{\bullet_j} \in \mathcal{B}^2(\mathbb{R}^d)$, see Remark 2.3.16), immediately provides an Itô-Stratonovich conversion formula for general semi-martingales.

2.4.2 Lévy rough paths

Note that the example in the previous section can be viewed as follows: $\mathbf{B}^{\mathrm{It\hat{o}}}$ and $\mathbf{B}^{\mathrm{Strat}}$ are both $\mathcal{G}^2$-valued Lévy processes which are branched p-rough paths, $2 < p < 3$, and one can recover the signature of one from the other by a suitable (deterministic) translation map $M_v : \mathcal{G} \to \mathcal{G}$. We now consider a generalisation of this setting to arbitrary $\mathcal{G}^N$-valued Lévy processes, which have already been studied in the context of rough paths in [FS17, Section 3] and [Che18].

Let $\tau_1, \ldots, \tau_m$ be a basis for $\mathcal{B}^N$ consisting of trees, which we identify with left-invariant vector fields on $\mathcal{G}^N$, where we suppose for convenience that $\tau_1 = \bullet_0$. Recall that $\mathcal{G}^N$ is a homogenous group in the sense of [FS82, Section 1.A] (cf. [HK15, Remark 2.15]).

Recall that to every (left) Lévy process $\mathbf{X}$ in $\mathcal{G}^N$ without jumps and with identity starting point (i.e., $\mathbf{X}_0 = 1_{\mathcal{G}^N}$ a.s.) there is an associated Lévy tuple (A, B), where $B = \sum_{i=1}^m B^i \tau_i$ is an element of $\mathcal{B}^N$ and $(A^{i,j})_{i,j=1}^m$ is a correlation matrix. Then the generator of $\mathbf{X}$ is given for all $f \in C^2_0(\mathcal{G}^N)$ by (see, e.g., [Lia04, Theorem 1.1])

$$\lim_{t \to 0} t^{-1} \mathbb{E}\left[f(x \star \mathbf{X}_t) - f(x)\right] = \sum_{i=1}^m B^i (\tau_i f)(x) + \frac{1}{2} \sum_{i,j=1}^m A^{i,j} (\tau_i \tau_j f)(x).$$

Lemma 2.4.4. *Let $M : (\mathcal{H}^N, \star) \to (\mathcal{H}^N, \star)$ be an algebra homomorphism which preserves $\mathcal{G}^N$ and $\mathbf{X}$ a Lévy process in $\mathcal{G}^N$ with Lévy tuple (A, B).*

Then $M(\mathbf{X})$ is the (unique in law) $\mathcal{G}^N$-valued (left) Lévy process with generator given for all $f \in C_0^2(\mathcal{G}^N)$ by

$$\lim_{t\to 0} t^{-1}\mathbb{E}\left[f(x \star M\mathbf{X}_t) - f(x)\right] = \sum_{i=1}^m B^i(M\tau_i f)(x) + \frac{1}{2}\sum_{i,j=1}^m A^{i,j}(M\tau_i M\tau_j f)(x). \tag{2.12}$$

Proof. The fact that $M\mathbf{X}$ is a Lévy process is immediate from the fact that $\mathbf{X}$ is a Lévy process and that $M : \mathcal{G}^N \to \mathcal{G}^N$ is a (continuous) group homomorphism. It thus only remains to show (2.12), where we may suppose without loss of generality that $x = 1_{\mathcal{G}^N}$. To this end, define $h = f \circ M$ and observe that

$$\lim_{t\to 0} t^{-1}\mathbb{E}\left[f(M\mathbf{X}_t) - f(1_{\mathcal{G}^N})\right] = \sum_{i=1}^m B^i(\tau_i h)(1_{\mathcal{G}^N}) + \frac{1}{2}\sum_{i,j=1}^m A^{i,j}(\tau_i\tau_j h)(1_{\mathcal{G}^N})$$

(note that in general h might fail to decay at infinity and thus not be an element of $C_0^2(\mathcal{G}^N)$, however the above limit is readily justified by taking suitable approximations). Using the fact that $(\tau h)(x) = \frac{d}{dt}h(x \star e^{t\tau}) \mid_{t=0}$, one can easily verify that for all $\sigma, \tau \in \mathcal{B}^N$ and $x \in \mathcal{G}^N$

$$(\tau h)(x) = (M\tau f)(Mx),$$
$$(\sigma\tau h)(x) = ((M\sigma)(M\tau)f)(Mx),$$

from which (2.12) follows. □

We now specialise to the case that $(A^{i,j})_{i,j=1}^m$ is a correlation matrix for which $A^{i,i} = 0$ whenever τ_i has more than $\lfloor N/2 \rfloor$ nodes, which is a necessary and sufficient condition for the sample paths of $\mathbf{X}$ to a.s. have finite p-variation for all $N < p < N+1$ [Che18, Theorem 5.1]. Assume also that $A^{i,i} = 0$ whenever τ_i contains a node with label 0, and that $B = \tau_1 = \bullet_0$, so that for all $f \in C_0^2(\mathcal{G}^N)$

$$\lim_{t\to 0} t^{-1}\mathbb{E}\left[f(x \star \mathbf{X}_t) - f(x)\right] = (\tau_1 f)(x) + \frac{1}{2}\sum_{i,j=1}^m A^{i,j}(\tau_i\tau_j f)(x).$$

The drift term $(\tau_1 f)(x)$ should be interpreted as the time component of the branched rough path $\mathbf{X}$ (which also explains the zero-diffusion condition in the direction of trees with a label 0).

Any other $\mathcal{G}^N$-valued Lévy process $\tilde{\mathbf{X}}$ without jumps and the same correlation matrix $(A^{i,j})_{i,j=1}^m$ is also a branched p-rough path, and its generator differs from that of $\mathbf{X}$ only by a drift term. As a consequence of Lemma 2.4.4, we see that every such $\tilde{\mathbf{X}}$ can be constructed by applying a (deterministic) translation map M_v to $\mathbf{X}$. In particular, the full signature of $\tilde{\mathbf{X}}$ can be recovered from that of $\mathbf{X}$, generalising the example from Section 2.4.1.

Corollary 2.4.5. *Let $v = v_0 \in \mathcal{B}^N$ and $M_v : \mathcal{H}^N \to \mathcal{H}^N$ the truncation of the translation map from Section 2.3.2.3.*

Then $M_v(\mathbf{X})$ is the (unique in law) $\mathcal{G}^N$-valued (left) Lévy process with generator given for all $f \in C_0^2(\mathcal{G}^N)$ by

$$\lim_{t\to 0} t^{-1}\mathbb{E}\left[f(x \star M_v(\mathbf{X}_t)) - f(x)\right] = (\bullet_0 + v)f(x) + \frac{1}{2}\sum_{i,j=1}^m A^{i,j}(\tau_i\tau_j f)(x).$$

Remark 2.4.6. The statement of the corollary likewise holds for every algebra homomorphism $M : \mathcal{H}^N \to \mathcal{H}^N$ satisfying $M\bullet_0 = \bullet_0 + v$ and $M\tau = \tau$ for all forests $\tau \in \mathcal{H}^N$ without a label 0, which is a manifestation of the final point of the upcoming Theorem 2.5.1 (ii).

2.4.3 Higher-order translation and renormalization in finite-dimensions

In [BCF18], from which we give an excerpt in this subsection, two examples are studied of families of random bounded variation paths $(X^\varepsilon)_{\varepsilon>0}$ whose canonical lifts to geometric rough paths $(\mathbf{X}^\varepsilon)_{\varepsilon>0}$ diverge as $\varepsilon \to 0$. In particular, ODEs driven by X^ε in general also fail to converge. However, for suitably chosen $v^\varepsilon = v_0^\varepsilon \in \mathfrak{g}_{\leq N}(\mathbb{R}^d)$, for which in general $\lim_{\varepsilon\to 0} |v^\varepsilon| = \infty$, one obtains convergence of the translated rough paths $T_{v^\varepsilon}\mathbf{X}^\varepsilon$. In particular, it follows from the upcoming Theorem 2.5.10 that solutions to modified ODEs driven by X^ε, with terms generally diverging as $\varepsilon \to 0$, converge to well-defined limits. In this specific context, the translation maps T_{v^ε} are precisely the renormalization maps occurring in regularity structures when applied to the setting of SDEs; we shall make this connection precise in Section 2.6.

2.4.3.1 Physical Brownian motion in a (large) magnetic field.

It was shown in [FGL15, Theorem 1] that the motion of a charged Brownian particle, in the zero mass limit, in a magnetic field which is kept constant while taking the limit, naturally leads to a perturbed second level, of the form $\bar{\mathbb{B}}_{s,t} = \mathbb{B}^{\text{Strat}}_{s,t} + v\,(t-s)$ for some $0 \neq v \in \mathfrak{so}\,(d)$, v being proportional to the strength of the magnetic field. We now want to look at the evolution of the system under the blow-up of the magnetic field.

Consider a physical Brownian motion in a magnetic field with dynamics given by

$$m\ddot{x} = -A\dot{x} + B\dot{x} + \xi, \quad x(t) \in \mathbb{R}^d,$$

where A is a symmetric matrix with strictly positive spectrum (representing friction), B is an anti-symmetric matrix (representing the Lorentz force due to a magnetic field), and ξ is an $\mathbb{R}^d$-valued white noise in time. We shall consider the situation that A is constant whereas B is a function of the mass m.

We rewrite these dynamics as

$$\begin{aligned} dX_t &= \frac{1}{m} P_t dt, \quad X_0 = 0, \\ dP_t &= -\frac{1}{m} M P_t dt + dW_t, \quad P_0 = 0, \end{aligned}$$

where $M = A - B$, and we have chosen the starting point as zero simply for convenience. We furthermore introduce the parameter $\varepsilon^2 = m$ and write $X^\varepsilon_t, P^\varepsilon_t$, and $M^\varepsilon = A - B^\varepsilon$ to denote the dependence on ε.

We are interested in the convergence of the processes P^ε and $M^\varepsilon X^\varepsilon$ in rough path topologies. As before in Section 2.2, let $G^2(\mathbb{R}^d)$ and $\mathfrak{g}_{\leq 2}(\mathbb{R}^d)$ denote the step-2 free nilpotent Lie group and Lie algebra respectively. Let us also write $\mathfrak{g}_{\leq 2}(\mathbb{R}^d) = \mathbb{R}^d \oplus \mathfrak{g}_2(\mathbb{R}^d)$ for the decomposition of $\mathfrak{g}_{\leq 2}(\mathbb{R}^d)$ into the first and second levels, where we identify $\mathfrak{g}_2(\mathbb{R}^d)$ with the space of anti-symmetric $d \times d$ matrices.

For every $\varepsilon > 0$, define the matrix

$$C^\varepsilon = \int_0^\infty e^{-M^\varepsilon s} e^{-(M^\varepsilon)^* s} ds,$$

and the element

$$v^\varepsilon = -\frac{1}{2}(M^\varepsilon C^\varepsilon - C^\varepsilon (M^\varepsilon)^*) \in \mathfrak{g}_2(\mathbb{R}^d).$$

For $\alpha \in (1/3, 1/2]$, due to the extension theorem, any α-Hölder weakly geometric rough path $\mathbf{Z} : [0,T]^2 \to G(\mathbb{R}^d)$ is fully characterized by the truncation $\pi_2\mathbf{Z} : [0,T]^2 \to G^2(\mathbb{R}^d)$. Thus, for the purpose of this example, we represent any such rough path $\mathbf{Z}$ by the increments $Z_{s,t}$ of the underlying path and the second level $\mathbb{Z}_{s,t}$, i.e.

$$Z^i_{s,t} = \langle \mathbf{Z}_{s,t}, \mathtt{i} \rangle, \qquad \mathbb{Z}^{j,k}_{s,t} = \langle \mathbf{Z}_{s,t}, \mathtt{jk} \rangle.$$

In this special case and for any $v = v_0 \in \mathfrak{g}_2(\mathbb{R}^d)$, the translation map introduced in Definition 2.2.2 is given by

$$T_v(Z_{s,t}, \mathbb{Z}_{s,t}) = (Z_{s,t}, \mathbb{Z}_{s,t} + (t-s)v). \tag{2.13}$$

Consider the $G^2(\mathbb{R}^d)$-valued processes

$$\begin{aligned}
(P^\varepsilon_{s,t}, \mathbb{P}^\varepsilon_{s,t}) &= \left(P^\varepsilon_{s,t}, \int_s^t P^\varepsilon_{s,r} \bullet \circ dP^\varepsilon_r \right), \\
(Z^\varepsilon_{s,t}, \mathbb{Z}^\varepsilon_{s,t}) &= \left(M^\varepsilon X^\varepsilon_{s,t}, \int_s^t M^\varepsilon X_{s,r} \bullet d(M^\varepsilon X^\varepsilon)_r \right),
\end{aligned}$$

and the canonical lift of the Brownian motion W

$$(W_{s,t}, \mathbb{W}_{s,t}) = \left(W_{s,t}, \int_s^t W_{s,r} \bullet \circ dW_r \right),$$

where the integrals in the definition of $\mathbb{P}^\varepsilon_{s,t}$ and $\mathbb{W}_{s,t}$ are in the Stratonovich sense.

Contrary to [FGL15], we allow blow-up of the magnetic field with rate $B^\varepsilon \lesssim \varepsilon^{-\kappa}$, $\kappa \in [0,1]$, as a method to model magnetic fields which are large (in a quantified way) in comparison to the (small) mass. The paths Z^ε then form approximations of Brownian motion, whose canonical rough path lifts $(Z^\varepsilon, \mathbb{Z}^\varepsilon)$ do not converge in rough path space (due to divergence of the Lévy's area). The following result establishes convergence of the "renormalised" paths $T_{v^\varepsilon}(P^\varepsilon_{s,t}, \mathbb{P}^\varepsilon_{s,t})$ and $T_{v^\varepsilon}(Z^\varepsilon_{s,t}, \mathbb{Z}^\varepsilon_{s,t})$.

Theorem 2.4.7 ([BCF18, Theorem 1])**.** *Suppose that*

$$\lim_{\varepsilon \to 0} |M^\varepsilon| \varepsilon^\kappa = 0 \ \text{ for some } \kappa \in [0,1]. \tag{2.14}$$

Then for any $\alpha \in [0, 1/2 - \kappa/4)$ and $q < \infty$, it holds that $T_{v^\varepsilon}(P^\varepsilon, \mathbb{P}^\varepsilon) \to (0,0)$ and $T_{v^\varepsilon}(Z^\varepsilon, \mathbb{Z}^\varepsilon) \to (W, \mathbb{W})$ in L^q and α-Hölder topology as $\varepsilon \to 0$. More precisely, as $\varepsilon \to 0$, in L^q

$$\sup_{s,t \in [0,T]} \frac{|P^\varepsilon_{s,t}|}{|t-s|^\alpha} + \sup_{s,t \in [0,T]} \frac{|\mathbb{P}^\varepsilon_{s,t} + (t-s)v^\varepsilon|}{|t-s|^{2\alpha}} \to 0.$$

and

$$\sup_{s,t \in [0,T]} \frac{|Z^\varepsilon_{s,t} - W_{s,t}|}{|t-s|^\alpha} + \sup_{s,t \in [0,T]} \frac{|\mathbb{Z}^\varepsilon_{s,t} + (t-s)v^\varepsilon - \mathbb{W}_{s,t}|}{|t-s|^{2\alpha}} \to 0.$$

In particular, if $\kappa \in [0, \frac{2}{3})$, one can take $\alpha \in (\frac{1}{3}, \frac{1}{2} - \frac{\kappa}{4})$ and convergence takes place in α-Hölder rough path topology.

Lastly, we would like to point out that higher-order renormalization can be expected in the presence of highly oscillatory fields, which also points to some natural connections with homogenization theory.

2.4.3.2 Fractional delay / Hoff process

Viewed as two-dimensional rough paths, Brownian motion and its ε-delay, $t \mapsto (B_t, B_{t-\varepsilon})$, does not converge to (B, B), with - as one may expect - zero area. Instead, the quadratic variation of Brownian motion leads to a rough path limit of the form $(B, B; A)$ with area of order one [FV10, Theorem 13.31]. It is then possible to check that, replacing B by a fractional Brownian motion with Hurst parameter $H < 1/2$, the same construction will yield exploding Lévy area as $\varepsilon \downarrow 0$.

The same phenomenon is seen in lead-lag situations, popular in time series analysis. As in the case of physical Brownian motion in a (large) magnetic field, these divergences can be cured by applying suitable (second-level) translation / renormalization operators, as we shall now see; for details on the (non-divergent) Brownian / semi-martingale case, see e.g. [FV10, Section 13.3.5] and [FHL16].

Consider a path $X : [0,1] \mapsto \mathbb{R}^d$. Let $n \geq 1$ be an integer and write for brevity $X_i^n = X_{i/n}$. Consider the piecewise linear path $\tilde{X}^n : [0,1] \mapsto \mathbb{R}^{2d}$ defined by

$$\tilde{X}^n_{2i/2n} = (X_i^n, X_i^n),$$
$$\tilde{X}^n_{(2i+1)/2n} = (X_i^n, X_{i+1}^n),$$

and linear on the intervals $\left[\frac{2i}{2n}, \frac{2i+1}{2n}\right]$ and $\left[\frac{2i+1}{2n}, \frac{2i+2}{2n}\right]$ for all $i = 0, \ldots, n-1$. Note that this is a variant of the Hoff process considered in [FHL16] (as given in [FHL16, Definition 2.1]).

Denote by $\tilde{\mathbf{X}}^n_{s,t} = \pi_2 \exp_\bullet(\tilde{X}^n_{s,t} + \mathbb{A}^n_{s,t})$ the level-2 lift of $\tilde{X}^n$, where $\mathbb{A}^n_{s,t}$ is the $(2d) \times (2d)$ anti-symmetric Lévy area matrix given by

$$\mathbb{A}^n_{s,t} = \frac{1}{2}\left(\int_s^t \tilde{X}^n_{s,r} \bullet d\tilde{X}^n_r - \int_s^t \tilde{X}^n_{s,r} \bullet d\tilde{X}^n_r\right).$$

Let $H \in (0,1)$ and consider a fractional Brownian motion B^H with covariance $R(s,t) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H})$. Let $X : [0,1] \mapsto \mathbb{R}^d$ be d independent copies of B^H.

Recall the definition of T_v from (2.13). We are interested in the convergence in rough path topologies of $T_{\tilde{v}^n}(\tilde{\mathbf{X}}^n)$ where $\tilde{v}^n \in \mathfrak{g}_2(\mathbb{R}^{2d})$ is appropriately chosen. Define the (diagonal) $d \times d$ matrix

$$v^n = \frac{1}{2}\mathbb{E}\left[\sum_{k=0}^{n-1}(X^n_{k+1} - X^n_k) \otimes (X^n_{k+1} - X^n_k)\right] = \frac{n^{1-2H}}{2} I,$$

and the anti-symmetric $(2d) \times (2d)$ matrix

$$\tilde{v}^n = \begin{pmatrix} 0 & -v^n \\ v^n & 0 \end{pmatrix} \in \mathfrak{g}_2(\mathbb{R}^{2d}).$$

Finally, consider the path $\tilde{X} = (X, X) : [0,1] \mapsto \mathbb{R}^{2d}$, its canonically defined Lévy area $\mathbb{A}$ (which exists for $1/4 < H \leq 1$), and its level-2 lift $\tilde{\mathbf{X}} = \pi_2 \exp_\bullet(\tilde{X} + \mathbb{A})$. The following result establishes convergence of the "renormalised path" $T_{\tilde{v}^n}(\tilde{\mathbf{X}}^n)$.

Theorem 2.4.8 ([BCF18, Theorem 5])**.** *Suppose $1/4 < H \leq 1/2$. Then for all $\alpha \in [0, H)$ and $q < \infty$, it holds that $T_{\tilde{v}^n}(\tilde{\mathbf{X}}^n) \to \tilde{\mathbf{X}}$ in L^q and α-Hölder topology. More precisely, as $n \to \infty$, in L^q*

$$\sup_{s,t \in [0,T]} \frac{|\tilde{X}^n_{s,t} - \tilde{X}_{s,t}|}{|t-s|^\alpha} + \sup_{s,t \in [0,T]} \frac{|\mathbb{A}^n_{s,t} + (t-s)\tilde{v}^n - \mathbb{A}_{s,t}|}{|t-s|^{2\alpha}} \to 0.$$

2.4.3.3 Rough stochastic volatility and robust Itô integration.

Applications from quantitative finance recently led to the pathwise study of the (1-dimensional) Itô-integral,

$$\int_0^T f(\hat{B}_t)dB_t \text{ with } \hat{B}_t = \int_0^t |t-s|^{H-1/2}\, dB_s$$

where $f:\mathbb{R}\to\mathbb{R}$ is of the form $x\mapsto \exp(\eta x)$. When $H\in(0,1/2)$, the case relevant in applications, this stochastic integration is singular in the sense that the mollifier approximations actually diverge (infinite Itô-Stratonovich correction, due to infinite quadratic variation of $\hat{B}$ when $H<1/2$). The integrand $f(\hat{B}_t)$, which plays the role of a stochastic volatility process ($\eta>0$ is a volatility-of-volatility parameter) is *not* a controlled rough path, nor has the pair $(\hat{B},B)$ a satisfactory rough path lift (the Itô integral $\int \hat{B}dB$ is well-defined, but $\int Bd\hat{B}$ is not). The correct "Itô rough path" in this context is then an $\mathbb{R}^{n+1}$-valued "partial" branched rough path of the form

$$\left(B,\hat{B},\int \hat{B}dB,\dots\int \hat{B}^n dB\right)$$

where $n\sim 1/H$. Again, mollifier approximations will diverge but it is possible to see that one can carry out a renormalization which restores convergence to the Itô limit. (We note the similarity with SPDE situations like KPZ.) See [BFG+20] for details.

2.5 Rough differential equations

2.5.1 Translated rough paths are rough paths

We now show that the maps T_v and M_v act on the spaces of weakly geometric and branched rough paths. Throughout, we regard these rough paths as fully lifted, as can always (and uniquely) be done thanks to the extension theorem. The action of our translation operator is then pointwise, i.e.

$$(M_v\mathbf{X})_{s,t} := M_v(\mathbf{X}_{s,t}),$$

and similarly for the weakly geometric rough path translation operator T. In the following, we let $|w|$ denote the length of a word $w\in T(\mathbb{R}^{1+d})$ (resp. number of nodes in a forest $w\in\mathcal{H}$), and equip the space of α-Hölder weakly geometric (resp. branched) rough paths with the inhomogeneous Hölder norm

$$||\mathbf{X}||_{\alpha\text{-Höl};[s,t]} = \max_{|w|\le\lfloor 1/\alpha\rfloor}\ \sup_{u\neq v\in[s,t]} \frac{|\langle \mathbf{X}_{u,v},w\rangle|}{|v-u|^{|w|\alpha}},$$

where the max runs over all words $w\in T(\mathbb{R}^{1+d})$ (resp. forests $w\in\mathcal{H}$) with $|w|\le\lfloor 1/\alpha\rfloor$.

Theorem 2.5.1. *Let $\alpha\in(0,1]$ and $\mathbf{X}$ a α-Hölder weakly geometric (resp. branched) rough path over $\mathbb{R}^{1+d}$.*

(i) *Let $v=(v_0,v_1\dots,v_d)$ be a collection of elements in $\mathfrak{g}_{\le N}(\mathbb{R}^{d+1})$ (resp. in $\mathcal{B}^N$).*

Then $T_v\mathbf{X}$ (resp. $M_v\mathbf{X}$) is a α/N-Hölder weakly geometric (resp. branched) rough path satisfying

$$||T_v\mathbf{X}||_{\alpha/N\text{-Höl};[s,t]} \text{ (resp. } ||M_v\mathbf{X}||_{\alpha/N\text{-Höl};[s,t]}) \le C_v\,||\mathbf{X}||_{\alpha\text{-Höl};[s,t]} \tag{2.15}$$

for a constant C_v depending polynomially on v.

(ii) *Let* $v = (v_0, 0, \dots 0)$ *for* $v_0 \in \mathfrak{g}_{\leq N}(\mathbb{R}^{d+1})$ *(resp.* $v_0 \in \mathcal{B}^N$*). Suppose that* $\mathbf{X}$ *satisfies*

$$||\mathbf{X}||_{(1,\alpha)\text{-Höl};[s,t]} := \max_{|w| \leq \lfloor 1/\alpha \rfloor} \sup_{u \neq v \in [s,t]} \frac{|\langle \mathbf{X}_{u,v}, w \rangle|}{|v-u|^{(1-\alpha)|w|_0 + \alpha|w|}} < \infty, \tag{2.16}$$

where the max *runs over all words* $w \in T(\mathbb{R}^{1+d})$ *(resp. forests* $w \in \mathcal{H}$*) with* $|w| \leq \lfloor 1/\alpha \rfloor$ *and* $|w|_0$ *denotes the number of times the letter* e_0 *(resp. label* 0*) appears in* w*.*

Then $T_v\mathbf{X}$ *(resp.* $M_v\mathbf{X}$*) is a* $\alpha \wedge (1/N)$*-Hölder weakly geometric (resp. branched) rough path over* $\mathbb{R}^{1+d}$ *satisfying*

$$||T_v\mathbf{X}||_{\alpha\wedge(1/N)\text{-Höl};[s,t]} \ \left(\text{resp. } ||M_v\mathbf{X}||_{\alpha\wedge(1/N)\text{-Höl};[s,t]}\right) \ \leq C_v \, ||\mathbf{X}||_{(1,\alpha)\text{-Höl};[s,t]}$$

for a constant C_v *depending polynomially on* v*.*

Finally, in the setting of branched rough paths, let $M : \mathcal{H}^* \to \mathcal{H}^*$ *be any continuous algebra homomorphism which preserves* $\mathcal{G}$ *and such that* $M\tau = \tau$ *for every forest* $\tau \in \mathcal{H}$ *without a label* 0*, and* $M\bullet_0 = M_v\bullet_0 = \bullet_0 + v_0$*. Then* $M\mathbf{X} = M_v\mathbf{X}$*.*

Before the proof of the theorem, several remarks are in order.

Remark 2.5.2. In Theorem 2.5.1 we treat α-Hölder weakly geometric rough paths as already enhanced with their iterated integrals. Thus $\mathbf{X}_{s,t}$ is an element of $T((\mathbb{R}^{1+d}))$ and $(T_v\mathbf{X})_{s,t}$ is just the image of $\mathbf{X}_{s,t}$ under T_v. Therefore the statement of the proposition is that not only does $(T_v\mathbf{X})_{s,t}$ have the correct regularity on the first $n = \lfloor 1/\alpha \rfloor$ levels to qualify as a rough path but that all further iterated integrals are already given, in a purely algebraic way, by $(T_v\mathbf{X})$. That said, if one takes the level-n view, writing $\pi_n(T_v\mathbf{X})$ for the translation only defined as a level-n rough path, the extension theorem asserts that there is a unique full rough path lift, say $\mathbf{Z}$. But then, by the uniqueness part of the extension theorem, $\mathbf{Z} = T_v\mathbf{X}$ so that our construction is compatible with the rough path extension.

The same remark applies to branched rough paths, where we recall that, as a particular consequence of the sewing lemma, every α-Hölder branched rough path admits a unique lift (extension) to all of $\mathcal{H}^*$ ([Gub10, Theorem 7.3], or [HK15, page 223]). We would also like to point out that Boedihardjo in [Boe18, Theorem 4] recently extended a result on the factorial decay of lifts of geometric rough paths (first shown as part of [Lyo98, Theorem 2.2.1]) to the branched setting, answering a conjecture in [Gub10, Remark 7.4].

Remark 2.5.3. In the case of geometric rough paths, the previous remark points to an alternative (analytic) construction of the translation operator, first defined on a smooth path X identified with its full lift $X \equiv (1, X^1, X^2, \ldots)$, and subsequently extended to geometric rough paths by continuity. We stick to the case of one Lie polynomial $v_0 = v = (v^1, v^2, \ldots v^N)$ which we want to add at constant speed to X. At level 1, obviously $(T_vX)^1_{s,t} = X^1_{s,t} + (t-s)v^1$ and (T_vX) is a Lipschitz path (a 1-rough path). We then perturb the canonically obtained (extended) 2-rough path which in turn we can perturb on the second level by adding $(t-s)v^2$, thereby obtaining a (non-canonical) 2-rough path. Iterating this construction allows us to "feed in, level-by-level" the perturbation v until we arrive at a rough path $T_v\mathbf{X}$ with regularity α-Höl $\wedge\, (1/N)$. We leave it to the reader to check that this construction yields indeed $T_v\mathbf{X}$. The severe downside of this construction is that it's restricted to geometric rough paths, not to mention its repeated use of the (analytic) extension theorem, in a situation that is within reach of purely algebraic methods.

Remark 2.5.4. The condition on $\mathbf{X}$ in equation (2.16) is very natural and arises by "colifting" a Lipschitz path X^0 with a d-dimensional α-Hölder weakly geometric rough path. Moreover, this is a special case of a weakly geometric (p, q)-rough path (see [FV10, Section 9.4] with the definition of a path of "finite mixed (p,q)-variation" in [FV10, Definition 9.19]), and the statement can readily be extended to this general setting. One can also make a statement about the continuity of the maps $(v,\mathbf{X}) \mapsto T_v\mathbf{X}$ and $(v,\mathbf{X}) \mapsto M_v\mathbf{X}$ in suitable rough path topologies. However these points will not be explored here further.

Remark 2.5.5. The proof of Theorem 2.5.1 part (i) will reveal that the only properties required of T_v (resp. M_v) are that it is an algebra homomorphism, preserves group-like (or equivalently primitive) elements, is upper-triangular (increases grading), and that it increases the grade of every word of length k (resp. forest with k nodes) to at most Nk. While already the first of these conditions uniquely determines T_v once $T_v(e_i) = e_i + v_i$ is chosen, we emphasise that without demanding that M_v is a pre-Lie algebra homomorphism, there is freedom to how M_v can be extended to satisfy these properties even after $M_v(\bullet_i) = \bullet_i + v_i$ is chosen.

In general, different choices of M_v will give rise to different branched rough paths $M_v(\mathbf{X})$. There is a notable exception to this, which is when $\mathbf{X}$ is the canonical lift of a Lipschitz (or more generally α-Hölder, $\alpha \in (1/2,1]$) path in $\mathbb{R}^{1+d}$. Then for every algebra homomorphism $M : \mathcal{H}^* \to \mathcal{H}^*$ such that $M\bullet_i = M_v\bullet_i = \bullet_i + v_i$, it holds that $M\mathbf{X} = M_v\mathbf{X}$. Indeed, in this case $\mathbf{X}$ is necessarily in the image of $G(\mathbb{R}^{1+d}) \subset T((\mathbb{R}^{1+d}))$ under the embedding (2.6), and since M and M_v agree on the generators $\bullet_i$, it follows that $M\mathbf{X} = M_v\mathbf{X}$ (this discussion relates of course to the final point of Theorem 2.5.1 part (ii), where upon demanding additional structure on $\mathbf{X}$, we see that all maps M satisfying the specified properties agree on $\mathbf{X}$).

Remark 2.5.6. Observe that the level-N lift of a weakly geometric rough path is precisely the solution to the linear RDE

$$dY_t = L(Y_t)d\mathbf{X}_t, \;\; Y_0 = 1 \in T^N(\mathbb{R}^{1+d}),$$

where $L = (L_0,\dots,L_d)$ are the linear vector fields on $T^N(\mathbb{R}^{1+d})$ given by right-multiplication by $(\mathtt{0},\dots,\mathtt{d})$ respectively. In much the same way, the level-N truncation of the translated path $Y_t := \pi_N(T_v\mathbf{X}_t)$ is the solution to the modified linear RDE

$$dY_t = L^v(Y_t)d\mathbf{X}_t, \;\; Y_0 = 1 \in T^N(\mathbb{R}^{1+d}),$$

where now $L^v = (L_{\mathtt{0}+v_0},\dots,L_{\mathtt{d}+v_d})$ are given by right-multiplication by $(\mathtt{0}+v_0,\dots,\mathtt{d}+v_d)$ (which is a special case of the upcoming Theorem 2.5.10).

We note however that the same conclusion does not hold for branched rough paths. Indeed, even the level-N lift of a branched rough path $\mathbf{X}$, $N \geq \lfloor 1/\alpha \rfloor$, is in general not the solution of a linear RDE driven by $\mathbf{X}$, which can easily be seen from the fact that linear RDEs are completely determined by the values $\langle \mathbf{X}_{s,t}, \tau \rangle$ where τ ranges over all linear trees $\tau = [\dots[\bullet_{i_1}]_{\bullet_{i_2}}\dots]_{\bullet_{i_m}}$ (see, e.g., [HK15, Example 3.11]). A simple example is any branched rough path $\mathbf{X}$ for which $\langle \mathbf{X}, \tau \rangle = 0$ for all linear trees τ (e.g., the $\frac{1}{3}$-Hölder branched rough path for which $\langle \mathbf{X}_{s,t}, \tau \rangle = t-s$ for some $\tau = [\bullet_i\bullet_j]_{\bullet_k}$ and zero for every other tree τ of size $|\tau| \leq 3$), so that every linear RDE driven by $\mathbf{X}$ is constant.

Proof of Theorem 2.5.1. (i) We are required to show that

1. $T_v\mathbf{X}$ takes values in $G(\mathbb{R}^{1+d})$,
2. Chen's relation $(T_v\mathbf{X})_{s,t} \bullet (T_v\mathbf{X})_{t,u} = (T_v\mathbf{X})_{s,u}$ holds, and
3. the analytic condition (2.15).

The first two properties follow immediately from the analogous properties of $\mathbf{X}$ and the fact that $T_v\big|_{G(\mathbb{R}^{1+d})} : G(\mathbb{R}^{1+d}) \to G(\mathbb{R}^{1+d})$ is group homomorphism. To verify the final property, fix a word

$w \in T(\mathbb{R}^{1+d})$. It readily follows from Proposition 2.2.5 and Remark 2.2.6 that $T_v^* w = \sum_i \lambda_i w_i$ where $\lambda_i \in \mathbb{R}$ and w_i is a word which satisfies $N|w_i| \geq |w|$. However

$$|\langle \mathbf{X}_{s,t}, w_i \rangle| \leq ||\mathbf{X}||_{\alpha\text{-Höl};[s,t]} |t-s|^{\alpha|w_i|},$$

and thus

$$|\langle (T_v\mathbf{X})_{s,t}, w\rangle| = |\langle \mathbf{X}_{s,t}, T_v^* w\rangle| \leq C\, ||\mathbf{X}||_{\alpha\text{-Höl};[s,t]} |t-s|^{\alpha|w|/N}$$

with C depending only on w and (polynomially) on v. It follows that $T_v\mathbf{X}$ is indeed a α/N-Hölder rough path, and the desired estimate (2.15) follows by running over all w with $|w| \leq \lfloor N/\alpha \rfloor$. The proof for the case of branched rough paths is identical, using now Proposition 2.3.14.

The proof of the first statement of (ii) is virtually the same, except we now observe that Proposition 2.2.5 and the condition $v = v_0 \in \mathfrak{g}_{\leq N}(\mathbb{R}^{d+1})$ imply that $T_v^* w = \sum_i \lambda_i w_i$ where $\lambda_i \in \mathbb{R}$ and w_i is a word which satisfies

$$N|w_i|_0 + (|w_i| - |w_i|_0) \geq |w|.$$

The first statement of (ii) now follows from (2.16), and the proof for the case of branched rough paths is again identical.

To show the last point of (ii), consider the subspace $\mathcal{H}^k(\mathbb{R}^d) \oplus \langle \bullet_0 \rangle \subset \mathcal{H}^k$ spanned by $\bullet_0$ and all forests $\tau \in \mathcal{H}^k$ without a label 0. Observe that it suffices to show that for every $k \geq 0$, the level-k truncation $\pi_k \mathbf{X}$ takes values in the subalgebra of $\mathcal{H}^k$ generated by $\mathcal{H}^k(\mathbb{R}^d) \oplus \langle \bullet_0 \rangle$.

To this end, consider the space $\tilde{C}^\infty$ defined as the collection of all piecewise smooth paths $\mathbf{x} : [0,T] \to \mathcal{G}^k$ for which $\dot{\mathbf{x}} \in \mathcal{H}^k(\mathbb{R}^d) \oplus \langle \bullet_0 \rangle$ (so that in fact $\dot{\mathbf{x}} \in \mathcal{B}^k(\mathbb{R}^d) \oplus \langle \bullet_0 \rangle$). For every partition $D = (t_0, \ldots, t_m) \subset [0,T]$, we can construct $\mathbf{x}^D \in \tilde{C}^\infty$ as the piecewise geodesic path (for the Riemannian structure of $\mathcal{G}^k$) whose increment over $[t_i, t_{i+1}]$ is $\exp(\pi_{\mathcal{B}^k(\mathbb{R}^d)\oplus\langle \bullet_0 \rangle} \log \mathbf{X}_{t_i,t_{i+1}})$. One can verify that condition (2.16) guarantees that $\mathbf{x}^D \to \pi_k \mathbf{X}$ uniformly as $|D| \to 0$. The conclusion now follows since, by construction, $\mathbf{x}^D$ takes values in the subalgebra generated by $\mathcal{B}^k(\mathbb{R}^d) \oplus \langle \bullet_0 \rangle$. $\square$

2.5.2 Effects of translations on RDEs

Throughout this section, we assume that $f = (f_0, \ldots, f_d)$ is a collection of vector fields on $\mathbb{R}^e$ which are as regular as required for all stated operations and RDEs to make sense.

Observe that f induces a canonical map from $\mathfrak{g}_{\leq N}(\mathbb{R}^{d+1})$ to the space of vector fields $\mathrm{Vect}(\mathbb{R}^e)$ which extends the map $\mathtt{i} \mapsto f_i$. Write f_u for the image of $u \in \mathfrak{g}_{\leq N}(\mathbb{R}^d)$ under this map, e.g., for $u = [\mathtt{1}, \mathtt{2}]$, we have the vector field $f_{[\mathtt{1},\mathtt{2}]} := [f_1, f_2]$. Given a collection $v = (v_0, \ldots, v_d) \subset \mathfrak{g}_{\leq N}(\mathbb{R}^{d+1})$, we write

$$f^v = (f_0^v, \ldots, f_d^v) = (f_{\mathtt{0}+v_0}, \ldots, f_{\mathtt{d}+v_d}).$$

Similarly, f induces a canonical map from $\mathcal{B}^N$ to $\mathrm{Vect}(\mathbb{R}^e)$ which extends $\bullet_i \mapsto f_i$ using the pre-Lie product $\blacktriangleright$ on $\mathrm{Vect}(\mathbb{R}^e)$ (recall from Example 2.3.5 that in coordinates $\left(f^i \partial_i\right) \blacktriangleright \left(g^j \partial_j\right) := \left(f^i \partial_i g^j\right) \partial_j$). Once more write $\hat{f}_u$ for the image of $u \in \mathcal{B}^N$ under this map, e.g., for $u = [\bullet_1]_{\bullet_2} = \bullet_1 \curvearrowright \bullet_2$, we have the vector field

$$\hat{f}_{\bullet_1 \curvearrowright \bullet_2} = \hat{f}_{[\bullet_1]_{\bullet_2}} := f_1 \blacktriangleright f_2$$

Again given a collection $v = (v_0, \ldots, v_d) \subset \mathcal{B}^N$, we write

$$f^v = (f_0^v, \ldots, f_d^v) = (\hat{f}_{\bullet_0 + v_0}, \ldots, \hat{f}_{\bullet_d + v_d}).$$

Furthermore, for $h := f^v = (f_0^v, \dots, f_d^v)$, put $\hat{f}_u^v := \hat{h}_u$.

Now, to cite an argument from the original formulation of the proof of Theorem 2.5.10 in the publication [BCFP19, Proof of Theorem 38], assuming for a moment that the vector fields f_i are all smooth, both $x \mapsto \hat{f}_x^v$ and $x \mapsto \hat{f}_{M_v x}$ constitute pre-Lie homomorphisms from $\mathcal{B}$ to the smooth vector fields on $\mathbb{R}^e$ with $\hat{f}^v_{\bullet_i} = \hat{f}_{\bullet_i + v_i} = \hat{f}_{M_v \bullet_i}$ for all i, and thus due to the universal property of the free pre-Lie algebra, we must in fact have

$$\hat{f}_x^v = \hat{f}_{M_v x} \tag{2.17}$$

for all $x \in \mathcal{B}$. Then, going back to the situation where we only assume all the f_i to be $\lfloor N/\alpha \rfloor - 1$ times continuously differentiable, we realise that Equation (2.17) stays valid for $x \in \mathcal{B}^{\leq \lfloor 1/\alpha \rfloor}$, as it only depends on the pre-Lie identity for the vector field product $\blacktriangleright$ which continues to hold as long as we only form $\blacktriangleright$ products which are well-defined given the regularity assumption for the f_i.

Remark 2.5.7. The map $u \mapsto \hat{f}_u$ is closely related to the notion of elementary differentials in B-series [CEM11, Section 10] and has already been used to study solutions of branched RDEs in the works of Cass–Weidner [CW17, Section 5] and Hairer–Kelly [HK15, Section 3.2 ff.] (note also that our notation $\hat{f}_u$ agrees with that of [**?**, []ection 3]HairerKelly15, except that they denote by f_u what we denote by $\hat{f}_u$ in this thesis).

Remark 2.5.8. Treating $\mathfrak{g}_{\leq N}(\mathbb{R}^{d+1})$ (resp. $\mathcal{B}^N$) as a nilpotent Lie (resp. pre-Lie) algebra, the map considered above is not in general a Lie (resp. pre-Lie) algebra homomorphism into $\mathrm{Vect}(\mathbb{R}^e)$.

Let us now look at a rough differential equation (RDE) of the form

$$dY = f(Y)d\mathbf{X} + g(Y)dZ. \tag{2.18}$$

We need the definition of another linear map $\mathfrak{b} : T(\mathbb{R}^{d+1}) \to \mathcal{B}(\mathbb{R}^{d+1})$,

$$\mathfrak{b}(\mathbf{e}) := 0, \quad \mathfrak{b}(\mathtt{i}) := \bullet_i, \quad \mathfrak{b}(\mathtt{i}w) := \bullet_i \curvearrowright \mathfrak{b}(w).$$

Definition 2.5.9. *For the purpose of this chapter, denote a continuous path $Y : [0,T] \to \mathbb{R}^e$ an* RDE solution *to the Equation* (2.18) *driven by an α-Hölder weakly geometric rough path $\mathbf{X}$ and a Lipschitz continuous path $Z : [0,T] \to \mathbb{R}^m$ (very often $m = 1$ and Z is just time: $Z_t = t$) if there exists $r : [0,T]^2 \to \mathbb{R}^e$ with $\|r_{st}\|_2 \in o(|t-s|)$ such that, in accordance with the Euler RDE estimate in e.g. [FV10, Corollary 10.15] and [HK15, Proposition 5.2],*

$$Y_t - Y_s = \sum_{|w| \leq \lfloor 1/\alpha \rfloor} \langle \mathbf{X}_{st}, w \rangle \hat{f}_{\mathfrak{b}(w)}(Y_s) + \sum_{j=1}^m (Z_t^j - Z_s^j) g_j(Y_s) + r_{st},$$

and an RDE solution to the equation (2.18) *driven by an α-Hölder branched rough path $\mathbf{X}$ and a Lipschitz path $Z : [0,T] \to \mathbb{R}^m$ if there exists $r : [0,T]^2 \to \mathbb{R}^e$ with $\|r_{st}\|_2 \in o(|t-s|)$ such that, in accordance[5] with the Euler RDE estimate derived in [HK15, Proposition 3.8],*

$$Y_t - Y_s = \sum_{|\tau| \leq \lfloor 1/\alpha \rfloor} \langle \mathbf{X}_{st}, \tau \rangle \hat{f}_\tau(Y_s) + \sum_{j=1}^m (Z_t^j - Z_s^j) g_j(Y_s) + r_{st}.$$

[5] see also the formal series [BCE20, Equation (2)], where a slightly different but isomorphic definition (actually the original one from [CL01]) of the free pre-Lie algebra product on trees leads to a symmetry factor "$\sigma(\tau)$" appearing

For a $(1,\alpha)$ weakly geometric (resp. branched) rough path $\mathbf{X}$ and a Lipschitz path $Z:[0,T]\to\mathbb{R}^m$, we call $Y:[0,T]\to\mathbb{R}^e$ an RDE solution to Equation 2.18 if it is an RDE solution to

$$dY = f(Y)d\check{\mathbf{X}} + f_0(Y)dX^0 + g(Y)dZ$$

in the above sense driven by the α weakly geometric (resp. branched) rough path $\check{\mathbf{X}}$ and the Lipschitz path $(X^0,Z):[0,T]\to\mathbb{R}^{m+1}$. Here $\check{\mathbf{X}} := \mathsf{proj}_{T((\mathbb{R}^d))}\mathbf{X}$ (resp. $\mathsf{proj}_{\mathcal{H}^*(\mathbb{R}^d)}\mathbf{X}$), considering $T((\mathbb{R}^d))$ as a subalgebra of $T((\mathbb{R}^{d+1}))$ (resp. $\mathcal{H}^*(\mathbb{R}^d)$ as a subalgebra of $\mathcal{H}^*(\mathbb{R}^{d+1})$), i.e. $\mathbf{X}\mapsto\check{\mathbf{X}}$ is realised by putting every word (resp. forest) containing a letter $\mathtt{0}$ (resp. label 0) in the expansion of $\mathbf{X}$ to zero.

Theorem 2.5.10 (cf. [Sus91, Theorem 1], [FO09, Theorem 2])**.** (i) *Let notation be as in Theorem 2.5.1 part* (i)*. Then Y is an RDE solution to*

$$dY = f(Y)\,d(T_v\mathbf{X}) + g(Y)dZ \ \ (\text{resp. } dY = f(Y)\,d(M_v\mathbf{X}) + g(Y)dZ)$$

if and only if Y is an RDE solution to

$$dY = f^v(Y)\,d\mathbf{X} + g(Y)dZ \ \ (\text{resp. } dY = \hat{f}^v(Y)\,d(M_v\mathbf{X}) + g(Y)dZ)$$

(ii) *Let notation be as in Theorem 2.5.1 part* (ii)*. Then Y is an RDE solution to*

$$dY = f(Y)\,d(T_v\mathbf{X}) + g(Y)dZ \ \ (\text{resp. } dY = f(Y)\,d(M_v\mathbf{X}) + g(Y)dZ)$$

if and only if Y is an RDE solution to

$$\begin{aligned} &dY = f^v(Y)d\mathbf{X} + g(Y)dZ \equiv f(Y)\,d\mathbf{X} + f_{v_0}(Y)\,dX^0 + g(Y)dZ \\ &(\text{resp. } dY = f^v(Y)d\mathbf{X} + g(Y)dZ \equiv f(Y)d\mathbf{X} + \hat{f}_{v_0}(Y)dX^0 + g(Y)dZ). \end{aligned}$$

Lemma 2.5.11. *For any $u\in\mathfrak{g}(\mathbb{R}^{d+1})$ and any $x\in T^{\geq 1}(\mathbb{R}^{d+1})$ (the space of non-empty words), we have*

$$\mathfrak{b}(u\bullet x) = \mathfrak{b}(u)\curvearrowright\mathfrak{b}(x).$$

Proof. We proceed by induction over the length of u. For u a letter, the statement holds by definition of $\mathfrak{b}$. Assume the statement holds for all u of length n. Then, to check it for all homogeneous Lie elements of word length $n+1$, since left bracketings span the free Lie algebra, it suffices to look at Lie elements of the form $[u,\mathtt{i}]$. So,

$$\begin{aligned} \mathfrak{b}([u,\mathtt{i}]\bullet x) &= \mathfrak{b}(u\bullet\mathtt{i}\bullet x) - \mathfrak{b}(\mathtt{i}\bullet u\bullet x) = \mathfrak{b}(u)\curvearrowright(\mathfrak{b}(\mathtt{i})\curvearrowright\mathfrak{b}(x)) - \mathfrak{b}(\mathtt{i})\curvearrowright(\mathfrak{b}(u)\curvearrowright\mathfrak{b}(x)) \\ &= [\mathfrak{b}(u),\mathfrak{b}(\mathtt{i})]_\star\curvearrowright\mathfrak{b}(x) = \mathfrak{b}([u,\mathtt{i}])\curvearrowright\mathfrak{b}(x). \end{aligned}$$

□

In particular, $\mathfrak{b}$ restricted to $\mathfrak{g}(\mathbb{R}^{d+1})$ is the unique Lie homomorphism from the free Lie algebra to the free pre Lie algebra mapping $\mathtt{i}$ to $\bullet_i$, i.e. $\mathfrak{a}(x)=\mathfrak{b}(x)$ for all $x\in\mathfrak{g}(\mathbb{R}^{d+1})$, or, more generally,

$$\mathfrak{b}(y) = \mathsf{proj}_{\mathcal{B}}\mathfrak{a}(y), \tag{2.19}$$

for all $y\in T((\mathbb{R}^d))$, an immediate consequence of $\mathsf{proj}_{\mathcal{B}}(\tau_1\star\tau_2)=\tau_1\curvearrowright\tau_2$ and the fact that $\mathsf{proj}_{\mathcal{B}}(\zeta_1\star\zeta_2)=0$ whenever ζ_1 or ζ_2 is a non-tree forest. Note that $\mathfrak{b}$ is not an algebra homomorphism on $T(\mathbb{R}^{d+1})$ in contrast to $\mathfrak{a}$. Nevertheless $\mathfrak{b}$ is injective on the whole of $T(\mathbb{R}^{d+1})$, but $\mathfrak{b}(T(\mathbb{R}^{d+1}))$ is a strict subspace of $\mathcal{B}$.

For our Lie and pre-Lie vector field constructions, we have $f_u=\hat{f}_{\mathfrak{b}(u)}=\hat{f}_{\mathfrak{a}(u)}$, from which we conclude

$$f_i^v = f_{\mathtt{i}}^v = \hat{f}_{\bullet_i}^{\mathfrak{b}(v)} = \hat{f}_{\bullet_i}^{\mathfrak{a}(v)} \tag{2.20}$$

and

$$f_u^v = \hat{f}_{\mathfrak{b}(u)}^{\mathfrak{b}(v)} = \hat{f}_{\mathfrak{a}(u)}^{\mathfrak{a}(v)}$$

for all $u, v \in \mathfrak{g}(\mathbb{R}^{d+1})$ and the $\mathbb{R}^e$ vector fields $(f_i)_i$ sufficiently many times continuously differentiable.

Corollary 2.5.12. *For any $v = (v_0, \ldots, v_d) \subset \mathfrak{g}(\mathbb{R}^{d+1})$ and $x \in T(\mathbb{R}^{d+1})$, we have*

$$M_{\mathfrak{b}(v)} \mathfrak{b}(x) = \mathfrak{b}(T_v x).$$

Proof. The statement obviously holds for x the empty word or a letter. Assume it holds for a word w. Then, using Lemma 2.5.11,

$$\begin{aligned} M_{\mathfrak{b}(v)} \mathfrak{b}(\mathtt{i}w) &= M_{\mathfrak{b}(v)}(\bullet_i \curvearrowright \mathfrak{b}(w)) = (\bullet_i + \mathfrak{b}(v_i)) \curvearrowright M_{\mathfrak{b}(v)} \mathfrak{b}(w) = \mathfrak{b}(\mathtt{i} + v_i) \curvearrowright \mathfrak{b}(T_v w) \\ &= \mathfrak{b}((\mathtt{i} + v_i) \bullet T_v w) = \mathfrak{b}(T_v(\mathtt{i}w)). \end{aligned}$$

The general result follows by induction over word length and linearity. □

Remark 2.5.13. Since the space of weakly geometric rough paths embeds into the space of branched rough paths using the map (2.6), the statements in Theorem 2.5.10 for weakly geometric rough paths are a special case of those for branched rough paths. Indeed, for a weakly geometric rough path $\mathbf{X}$ and its branched rough path $\mathfrak{a}\mathbf{X}$ we have for all n and all collections of C^{n-1} $\mathbb{R}^e$-vector fields $(f_i)_{i=0\ldots d}$, through Equation (2.19) and the fact that $\mathfrak{a}$ preserves the grading, cf. [HK15, Section 5.1]

$$\begin{aligned} \sum_{|\tau| \le n} \langle \mathfrak{a}\mathbf{X}_{s,t}, \tau \rangle \hat{f}_\tau &= \sum_{|\tau| \le n} \langle \mathbf{X}_{s,t}, \mathfrak{a}^\top(\tau) \rangle \hat{f}_\tau = \sum_w \sum_{|\tau| \le n} \langle \mathbf{X}_{s,t}, w \rangle \langle \mathfrak{a}^\top(\tau), w \rangle \hat{f}_\tau \\ &= \sum_{|w| \le n} \langle \mathbf{X}_{s,t}, w \rangle \sum_\tau \langle \tau, \mathfrak{a}(w) \rangle \hat{f}_\tau = \sum_{|w| \le n} \langle \mathbf{X}_{s,t}, w \rangle \sum_\tau \langle \tau, \mathfrak{b}(w) \rangle \hat{f}_\tau \\ &= \sum_{|w| \le n} \langle \mathbf{X}_{s,t}, w \rangle \hat{f}_{\mathfrak{b}(w)}. \end{aligned}$$

We make a distinction between the two cases only for clarity.

Proof of Theorem 2.5.10. For clarity, we first prove the statement for weakly geometric rough paths and then generalise to branched rough paths (although by Remark 2.5.13, it suffices to prove the statement only in the branched case).

For all $x \in T(\mathbb{R}^{d+1})$, we have, by Equation (2.17) and Corollary 2.5.12,

$$\sum_w \langle x, w \rangle \hat{f}_{\mathfrak{b}(w)}^{\mathfrak{b}(v)} = \hat{f}_{\mathfrak{b}(x)}^{\mathfrak{b}(v)} = \hat{f}_{M_{\mathfrak{b}(v)} \mathfrak{b}(x)} = \hat{f}_{\mathfrak{b}(T_v x)} = \sum_w \langle T_v x, w \rangle \hat{f}_{\mathfrak{b}(w)}.$$

Thus, for any continuous $Y : [0, T] \to \mathbb{R}^e$, with $h = (h_0, \ldots, h_d) := (f_0^v, \ldots, f_d^v) = f^v$, using

Equation (2.20),

$$\begin{aligned}\sum_{|w|\le\lfloor 1/\alpha\rfloor}\langle \mathbf{X}_{s,t},w\rangle \hat{h}_{\mathfrak{b}(w)}(Y_s) &= \sum_{|w|\le\lfloor 1/\alpha\rfloor}\langle \mathbf{X}_{s,t},w\rangle \hat{f}^{\mathfrak{b}(v)}_{\mathfrak{b}(w)}(Y_s)\\ &= \sum_w \langle \mathsf{proj}_{\le\lfloor 1/\alpha\rfloor}\mathbf{X}_{s,t},w\rangle \hat{f}^{\mathfrak{b}(v)}_{\mathfrak{b}(w)}(Y_s)\\ &= \sum_w \langle T_v\mathsf{proj}_{\le\lfloor 1/\alpha\rfloor}\mathbf{X}_{s,t},w\rangle \hat{f}_{\mathfrak{b}(w)}(Y_s)\\ &= \sum_w \langle \mathsf{proj}_{\le\lfloor N/\alpha\rfloor}T_v\mathbf{X}_{s,t},w\rangle \hat{f}_{\mathfrak{b}(w)}(Y_s)+R_{st}\\ &= \sum_{|w|\le\lfloor N/\alpha\rfloor}\langle T_v\mathbf{X}_{s,t},w\rangle \hat{f}_{\mathfrak{b}(w)}(Y_s)+R_{st},\end{aligned}$$

where the 2-norm of the $\mathbb{R}^e$ vector

$$\begin{aligned}R_{st} &= \sum_{|w|\le\lfloor N/\alpha\rfloor}\langle T_v\mathsf{proj}_{\le\lfloor 1/\alpha\rfloor}\mathbf{X}_{s,t}-T_v\mathbf{X}_{s,t},w\rangle \hat{f}_{\mathfrak{b}(w)}(Y_s)\\ &= \sum_{|w|\le\lfloor N/\alpha\rfloor}\langle \mathsf{proj}_{>\lfloor 1/\alpha\rfloor}\mathbf{X}_{s,t},T_v^*w\rangle \hat{f}_{\mathfrak{b}(w)}(Y_s)\end{aligned}$$

is uniformly bounded by $C|t-s|^\gamma$ for some constant $C>0$, with γ the smallest integer multiple of α strictly greater than one, since $\mathbf{X}$ is an α-Hölder rough path, $f_{\mathfrak{b}(w)}(Y_s)$ is a continuous function of s bounded on the compact interval $[0,T]$ for any w, and $\sum_{|w|\le\lfloor N/\alpha\rfloor}T_v^*w$ is a linear combination of a finite number of words. Thus, $\|R_{st}\|_2\in o(|t-s|)$, which shows the claim of the equivalence of the two RDEs in terms of solutions $Y:[0,T]\to\mathbb{R}^e$.

The branched rough path case then follows completely analogously, just that instead of $\sum_w\langle x,w\rangle \hat{f}^{\mathfrak{b}(v)}_{\mathfrak{b}(w)}=\sum_w\langle T_vx,w\rangle\hat{f}_{\mathfrak{b}(w)}$ for all $x\in T(\mathbb{R}^{d+1})$, we now use $\sum_\tau\langle x,\tau\rangle\hat{f}^v_\tau=\sum_\tau\langle M_vx,\tau\rangle\hat{f}_\tau$ for all $x\in\mathcal{H}(\mathbb{R}^{d+1})$ which we already saw in Equation (2.17).

Considering now the setup of a $(1,\alpha)$ branched rough path $\mathbf{X}$,

$$\begin{aligned}&\sum_{|\tau|\le 1/\alpha}\langle \check{\mathbf{X}}_{s,t},\tau\rangle\hat{f}_\tau(Y_s)+(X^0_t-X^0_s)\hat{f}_{\bullet_0+v_0}(Y_s)\\ &=\sum_{|\tau|\le 1/\alpha}\langle \mathbf{X}_{s,t},\tau\rangle\hat{f}^v_\tau(Y_s)+R_{st}\\ &=\sum_{|\tau|\le N\vee 1/\alpha}\langle M_v\mathbf{X}_{s,t},\tau\rangle\hat{f}_\tau(Y_s)+R'_{st}\\ &=\sum_{|\tau|\le N\vee 1/\alpha}\langle \mathsf{proj}_{\mathcal{H}^*(\mathbb{R}^d)}M_v\mathbf{X}_{s,t},\tau\rangle\hat{f}_\tau(Y_s)+(X^0_t-X^0_s)f_0(Y_s)+R''_{st},\end{aligned}$$

where the first and last equality hold as any tree other than $\bullet_0$ involving a label 0 corresponds to a rough path entry of order $\le|t-s|^{1+1/\alpha}$ and the second equality holds as now $(M_v\mathsf{proj}_{\le 1/\alpha}-\mathsf{proj}_{N\vee 1/\alpha}M_v)\mathbf{X}_{st}$ is of order $\le|t-s|^\gamma$ for some $\gamma>1$.

We proceed along the same lines for a $(1,\alpha)$ weakly geometric rough path.

□

Remark 2.5.14. Recall that $M_v:\mathcal{B}^*\to\mathcal{B}^*$ was constructed, from Section 2.3.2.2 on, as a pre-Lie algebra homomorphism. This matters in part (i) of Theorem 2.5.10 above, where this property is needed to obtain a universal conversion formula for translated RDEs. For example, consider that M_v was replaced by an algebra homomorphism M (which satisfies the conditions of Remark 2.5.5) such that $M(\bullet_i) = \bullet_i$ for all $i = 0, ..., d$, but acted non-trivially on some higher order trees (so that M is not a pre-Lie homomorphism). Then given vector fields f, in general there does not exist another collection of vector fields f_v such that for every branched rough path $\mathbf{X}$, the RDE driven by $M(\mathbf{X})$ along vector fields f agrees with the RDE driven by $\mathbf{X}$ along f_v. Indeed, if such f_v existed, then for every weakly geometric (branched) rough path $\mathbf{X}$ (so that $M(\mathbf{X}) = \mathbf{X}$), the RDEs driven by $M(\mathbf{X})$ and $\mathbf{X}$ agree without the need to change the vector fields f, so that necessarily $f_v = f$. However if $\mathbf{X}$ is a non-geometric branched rough path, the RDE driven by $M(\mathbf{X})$ along vector fields f will not in general agree with the RDE driven by $\mathbf{X}$ along f.

2.6 Link with renormalization in regularity structures

We now recall several notions from the theory of regularity structures and draw a link between the map δ from Section 2.3.3 and the coproduct Δ^- associated to negative renormalization in [BHZ19, Section 4 ff.] and [Hai16, Section 2 and 3]. In particular, we demonstrate how negative renormalization maps on the regularity structure associated to branched rough paths carry a natural interpretation as rough path translations (see Theorem 2.6.10 below).

2.6.1 Regularity structures

Regularity structures usually deal with (e.g. SPDE solutions) $u = u(z)$ where $z \in \mathbb{R}^n$ (e.g. space-time), u takes values in $\mathbb{R}$ (or $\mathbb{R}^e$). Equations further involve a β-regularizing kernel, and there are d sources of noise, say $\xi_1, ..., \xi_d$, of arbitrary (negative) order $\alpha_{\min}$, as long as the equation is subcritical.

2.6.1.1 Generalities

We review the general (algebraic) setup in the case $n = 1$, $\beta = 1$ and $\alpha_{\min} \in (-1, 0)$.

In the spirit of Hairer's formalism, consider the equation

$$u(t) = u(0) + \left(K * \sum_{i=1}^{d} f_i(u(\cdot))\xi_i(\cdot) \right)(t), \quad t \in \mathbb{R}, \tag{2.21}$$

where $u(t)$ is a real-valued function for which we solve, $\xi_i(t)$ are driving noises, f_i are smooth functions on $\mathbb{R}$ (one could readily extend to the case that u takes values in $\mathbb{R}^e$ and f_i are vector fields on $\mathbb{R}^e$), and K is a kernel which improves regularity by order $\beta = 1$.

Remark 2.6.1. The example to have in mind here is $K(s) = \exp(-\lambda s)1_{s>0}$, which allows to incorporate an additional linear drift term ("$-\lambda u dt$"), or of course the case $\lambda = 0$, i.e. the Heaviside step function, which leads to the usual setting of controlled differential equations. We shall indeed specialize to the Heaviside case in subsequent sections, as this simplifies some algebraic constructions and so provides a clean link to rough path structures. For the time being, however, we find it instructive to work with a general 1-regularizing K, as this illustrates the need for polynomials decorations as well as symbols $\mathcal{J}_k$, representing k-th derivatives of the kernel.

Our driving noises $\xi_i(t)$ should be treated as distributions on $\mathbb{R}$ of regularity $C^{\alpha-1}$ for some $\alpha \in (0, 1)$ (which will later correspond to the case of α-Hölder branched rough paths). In the case that $\alpha \leq 1/2$, due to the product $f_i(u)\xi_i$, (2.21) is singular and thus cannot in general be solved analytically. However, the equation is evidently sub-critical in the sense of [Hai14, Assumption 8.3] if we for the moment assume that the vector fields f_i are polynomial maps , and so one can build an associated regularity structure.

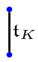
Introducing the symbols

We first collect all the symbols of the regularity structure required to both solve (2.21) and to form a set of symbols which is stable under the renormalization maps in the sense of [BHZ19, Section 4 ff.]. Define the linear space

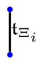
$$\mathcal{T} = \operatorname{span}_{\mathbb{R}} \mathcal{W},$$

where $\mathcal{W}$ is the set of all rooted trees where every node carries a "polynomial" decoration $k \in \mathbb{N} \cup \{0\}$ and where every edge which ends on a leaf may be (but is not necessarily) assigned a type $\mathfrak{t}_{\Xi_i}$, $i \in \{1, \dots, d\}$. An edge with type $\mathfrak{t}_{\Xi_i}$ corresponds to the driving noise ξ_i. Every other edge has a type $\mathfrak{t}_K$ which means that it is associated to the kernel K. (For now, we only assume K is 1-regularizing, later we will take it to be the Heaviside step function.) Also, each node has at most one incoming edge with type belonging to $\{1, \dots, d\}$.[6] With regard to [BHZ19], we also note the absence of edge decorations.[7]

To avoid confusion between the different meaning of trees in $\mathcal{W}$ and those introduced in Section 2.3, we will color every tree in $\mathcal{W}$ blue. Every such tree has a corresponding symbol representation (cf. $H_\circ$ in [BHZ19, Remark 4.16]), e.g.,



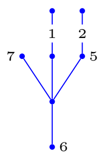
$\leftrightarrow$

$\leftrightarrow \mathcal{I}$,


$\leftrightarrow$

$\leftrightarrow \Xi_i$,

$\leftrightarrow X^k$,



$\leftrightarrow \mathcal{I}(\mathcal{I}(\Xi_1)\mathcal{I}(\Xi_2 X^5)\mathcal{I}(X^7))X^6$,

where we implicitly drop the 0 decoration ($\leftrightarrow X^0$) from the nodes. It is instructive to check that $\mathcal{W}$ provides an example of a structure built from a subcritical complete rule (in the sense of [BHZ19, Section 5.2 and 5.3]) arising from the equation (2.21). Indeed, we can write down the rule used for the construction of $\mathcal{W}$ as

$$R(\Xi_i) = \{()\}, \quad R(\mathcal{I}) = \{([\mathcal{I}]_\ell), ([\mathcal{I}]_\ell, \Xi_i), \ell \in \mathbb{N} \cup \{0\}, i \in \{1, ..., d\}\}. \tag{2.22}$$

The notation $[\mathcal{I}]_\ell$ is a shorthand notation for $\mathcal{I}, ..., \mathcal{I}$ where $\mathcal{I}$ is repeated ℓ times.

We define a degree $|\cdot|$ associated to an edge type and a decorated tree. For edge types and polynomials, we have

$$|\Xi_i| = \alpha - 1, \quad |\mathcal{I}| = 1, \quad |X^k| = k.$$

Then by recursion,

$$|\mathcal{I}(\tau)| = |\tau| + |\mathcal{I}|, \quad \left|\prod_i \tau_i\right| = \sum_i |\tau_i|.$$

[6] This rules out symbols corresponding to products of noise, such as $\Xi_i \Xi_j$ with $i, j \in \{1, \dots, d\}$.

[7] This is in contrast to, say, KPZ or Φ^4_3, where edge decorations appear in view of $Du \to \mathcal{I}'$ or negative renormalization, respectively.

For a non-recursive definition see [BHZ19, Definition 5.3], where the degree is described through a summation over all the edge types and the decorations in the tree[8].

Remark 2.6.2. Remark that $\mathcal{W} \equiv \mathcal{W}_{\mathrm{BHZr}}$ (the "r" in BHZr refers to *reduced*, in the terminology of [BHZ19, Section 6.4] these are trees without any extended decorations) will contain certain symbols which do *not* arise if one follows the original procedure of [Hai14, Section 8.1] (which, in some sense, is the most economical way to build the structure)[9]:

$$\mathcal{W}_{\mathrm{Hai14}} \subset \mathcal{W}_{\mathrm{BHZr}} \subset \mathcal{W}_{\mathrm{BHZ}}.$$

For example, $\mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)$, $\mathcal{I}(\mathcal{I}(\Xi_k))$, and $\mathcal{I}() \equiv \mathcal{I}(X^0)$ do not appear in $\mathcal{W}_{\mathrm{Hai14}}$, but all of these appear in $\mathcal{W}_{\mathrm{BHZr}}$. These in turn are embedded in $\mathcal{W}_{\mathrm{BHZ}}$, a set of trees with extended decorations on the nodes and also colourings of the nodes which give more algebraic properties. In the setting of [BHZ19], we would work with an additional symbol $\mathbf{1}_\alpha$ for $\alpha \in \mathbb{N}_0(\{\Xi_1,\dots,\Xi_d,\mathcal{I}\}) \cong \mathbb{N}_0^{d+1}$, representing an extended decoration, which provides information on some "singular" (negative degree) tree which has been removed, and all of these extended decorations would be placed based on the rule (2.22), see [BHZ19, Definitions 5.23 and 5.24].

Introducing $\mathcal{T}_-$

We define the space $\mathcal{T}_-$ as (cf. H_1 in [BHZ19, Remark 4.16])

$$\mathcal{T}_- = \mathrm{span}_{\mathbb{R}}\,\{\tau_1 \odot \cdots \odot \tau_n,\ \tau_i \in \mathcal{W},\ |\tau_i| < 0\}. \tag{2.23}$$

where $\odot$ is the forest product and the unit is given by the empty forest. (In other words, $\mathcal{T}_-$ is the free unital commutative algebra generated by elements in $\mathcal{W}$ of negative degree.) We now recall that $\mathcal{T}_-$ can be equipped with a Hopf algebra structure $\mathcal{T}_-$ for which there exists a coaction $\Delta^- : \mathcal{T} \to \mathcal{T}_- \otimes \mathcal{T}$ such that $(\mathcal{T}, \Delta^-)$ is a (left) comodule over $\mathcal{T}_-$. Then the action of a character $\ell \in \mathcal{T}_-^*$ on $x \in \mathcal{T}$, termed "negative renormalization", is given by $M_\ell x = (\ell \otimes \mathrm{id})\Delta^- x$.

Following [Hai16, page 10, in particular Equation (2.9)] we can describe the coaction Δ^- as follows. Fix a tree $\tau \in \mathcal{W}$, consider a subforest $A \subset \tau$, i.e., an arbitrary subgraph of τ which contains no isolated vertices. We then write $R_A\tau$ for the tree obtained by contracting the connected components of A in τ. With this notation at hand, we then define a linear map, the coaction,

$$\Delta^- : \mathcal{T} \to \mathcal{T}_- \otimes \mathcal{T}$$

by setting, for $\tau \in \mathcal{W}$,

$$\Delta^-\tau = \sum_{A \subset \mathcal{T}_-} A \otimes R_A\tau. \tag{2.24}$$

Unfortunately, this is not quite the correct coaction as it does not handle correctly the powers of X. However, upon restriction to $\tilde{\mathcal{T}} \subset \mathcal{T}$, as done in detail in the next section, this is precisely the form of the coaction (now on $\tilde{\mathcal{T}}$). When moving to a coproduct this fortunately plays no role (since $\mathcal{T}_-$ does not contain any non-zero powers of X or a factor of the form $\mathcal{I}()$). Following [Hai16, page 11], by abuse of notation, Δ^- also acts as a coproduct, that is

$$\Delta^- : \mathcal{T}_- \to \mathcal{T}_- \otimes \mathcal{T}_-. \tag{2.25}$$

To be explicit, given $f = \tau_1 \cdots \tau_n \in \mathcal{T}$, we have $\Delta^-(f) = \Delta^-(\tau_1)...\Delta^-(\tau_n)$ with each $\Delta^-(\tau_i)$ as defined above, but with an additional projection to the negative trees on the right-hand side of the tensor-product.

[8] Note that we only need one notion of degree here instead of the two $|\cdot|_-$ and $|\cdot|_+$ in [BHZ19, Definition 5.3] because we do not work with extended decorations

[9] $\mathcal{W}_{\mathrm{Hai14}}$ is called $\mathcal{H}_{\mathcal{F}}$ in [Hai14, Section 8.1] itself.

Remark 2.6.3. The spaces $\mathcal{T}_- \equiv \mathcal{T}^-_{\mathrm{BHZr}}$, $\mathcal{T}^-_{\mathrm{BHZ}}$ and $\mathcal{T}^-_{\mathrm{Hai14}}$ are the same *in this framework* (cf. assumptions from the beginning of this subsection). Indeed, all negative trees of $\mathcal{W}$ have a degree of the form $N\alpha - 1$. Then if we remove one negative subtree, of degree $M\alpha - 1$ say, from a negative tree, we obtain a degree $(N-M)\alpha$ which is positive and hence the "cured" tree does not belong to $\mathcal{T}_-$.

Introducing $\mathcal{T}_+$

In order to describe the space $\mathcal{T}_+$ as in [BHZ19], we need to associate to each edge a decoration $k \in \mathbb{N} \cup \{0\}$ viewed as a derivation of the kernels or the driving noises. Such a decoration does not appear in $\mathcal{T}$. Thus we will replace the letter $\mathcal{I}$ by $\mathcal{J}$ in this context. We do not give any graphical notation for $\mathcal{J}_k$, the edge with type $\mathfrak{t}_K$ and (edge) decoration k representing $K^{(k)}$, because these symbols ultimately will not appear in our context.

We define $\mathcal{T}_+$ as the linear span of (cf. $\hat{H}_2$ in [BHZ19, Remark 4.16])

$$\{X^k \prod_{i=1}^n \mathcal{J}_{k_i}(\tau_i) \mid k, n \in \mathbb{N} \cup \{0\}, k_i \in \mathbb{N} \cup \{0\}, \tau_i \in \mathcal{W},\ |\tau_i| + 1 - k_i > 0\}.$$

In other words, $\mathcal{T}_+$ is the free unital commutative algebra generated by

$$\mathcal{W}_+ := \{X\} \cup \{\mathcal{J}_k \tau \mid \tau \in \mathcal{W}, |\tau| + 1 - k > 0\}.$$

We use a different letter $\mathcal{J}$ to stress that $\mathcal{W}$ is different from $\mathcal{W}_+$. The use of this letter is viewed in [BHZ19] as a blue colouration of the root, see Section 4.3 there. We also define the degree of a term

$$\tau = X^k \prod_{i=1}^n \mathcal{J}_{k_i}(\tau_i) \in \mathcal{T}_+,\ \ |\tau| = k + \sum_{i=1}^n 1 - k_i + |\tau_i|.$$

The space $\mathcal{T}_+$ is used in the description of the structure group associated to $\mathcal{T}$. More precisely, recall that $\mathcal{T}_+$ can be equipped with a Hopf algebra structure for which there exists a coaction $\Delta^+ : \mathcal{T} \to \mathcal{T} \otimes \mathcal{T}_+$ such that $(\mathcal{T}, \Delta^+)$ is a (right) comodule over $\mathcal{T}_+$. Following Hairer's survey [Hai16, Equation (2.2) and (2.3)], the coaction

$$\Delta^+ \colon \mathcal{T} \to \mathcal{T} \otimes \mathcal{T}_+ \tag{2.26}$$

is given by

$$\Delta^+ X_i = X_i \otimes \mathbf{1} + \mathbf{1} \otimes X_i\ , \qquad \Delta^+ \Xi_i = \Xi_i \otimes \mathbf{1}\ , \tag{2.27}$$

and then recursively by

$$\Delta^+ \mathcal{I}(\tau) = (\mathcal{I} \otimes \mathrm{id})\Delta^+ \tau + \sum_{\ell \in \mathbb{N} \cup \{0\}} \frac{X^\ell}{\ell!} \otimes \mathcal{J}_\ell(\tau) \tag{2.28}$$

and

$$\Delta^+(\tau\bar{\tau}) = \Delta^+ \tau\, \Delta^+ \bar{\tau}. \tag{2.29}$$

The coproduct $\Delta^+ : \mathcal{T}_+ \to \mathcal{T}_+ \otimes \mathcal{T}_+$ is then defined in the same way by replacing (2.28) with

$$\Delta^+ \mathcal{J}_k(\tau) = (\mathcal{J}_k \otimes \mathrm{id})\Delta^+ \tau + \sum_{\ell \in \mathbb{N} \cup \{0\}} \frac{X^\ell}{\ell!} \otimes \mathcal{J}_{k+\ell}(\tau),$$

in which $\Delta^+ \tau$ is understood as the coaction $\Delta^+ : \mathcal{T} \to \mathcal{T} \otimes \mathcal{T}_+$.

Then the action of a character $g \in \mathcal{T}_+^*$ on $x \in \mathcal{T}$, termed "positive renormalization", is given by

$$\Gamma_g x = (\mathrm{id} \otimes g)\Delta^+ x.$$

Remark 2.6.4. The space $\mathcal{T}_+ \equiv \mathcal{T}^+_{\mathrm{BHZr}}$ depends strongly on the space $\mathcal{W}$. We have

$$\mathcal{T}^+_{\mathrm{Hai14}} \subset \mathcal{T}^+_{\mathrm{BHZr}} \subset \mathcal{T}^+_{\mathrm{BHZ}}.$$

These two inclusions are Hopf subalgebra inclusions. Indeed, as proved in [BHZ19, Section 6.4], the second one, with $\mathcal{T}_+$ equipped with coproduct Δ^+ is a Hopf subalgebra inclusion (with Δ^+_{BHZ} found in [BHZ19, Proposition 4.17 and Corollary 5.32]). The same is also true for $\mathcal{T}^+_{\mathrm{Hai14}}$. The key point for the Hopf algebra structure is thatthe symbols defined in [BHZ19, Section 5] are obtained by a "complete rule" which guarantees the invariance under Δ^+, see [BHZ19, Lemma 5.28]. Note that $\mathcal{T}^+_{\mathrm{Hai14}}$ is called "$\mathcal{H}^+_{\mathcal{F}}$" in [Hai14] itself, which is shown to be stable under Δ^+ in Lemma 8.22 there. In the case of $\mathcal{T}^+_{\mathrm{BHZ}}$, we use the degree $|\cdot|_+$ which is exactly $|\cdot|$ when we restrict ourselves to $\mathcal{T}^+_{\mathrm{BHZr}}$.

Remark 2.6.5. Unfortunately, there is a problem here in that, with the definition in equation (2.28), a desirable cointeraction between Δ^+ and Δ^- fails as we shall explain momentarily. The "official" remedy, following [BHZ19, Sections 4 and 5], is to use the extended decorations through another degree $|\cdot|_+$ which takes into account these decorations and behaves the same as $|\cdot|$ for the rest. For example, one has $|\mathcal{I}(\mathbf{1}_\beta\tau)|_+ = |\tau|_+ + 1 + \beta$. The "correct" coaction Δ^+ (see [BHZ19, Proposition 4.17]) then also involves these extended decorations. The extended decorations are crucial in [BHZ19, Theorem 5.37] for obtaining a cointeraction between the two Hopf algebras $(\mathcal{T}_+, \Delta^+)$ and $(\mathcal{T}_-, \Delta^-)$:

$$\mathcal{M}^{(13)(2)(4)}\left(\Delta^-\otimes\Delta^-\right)\Delta^+ = \left(\mathrm{id}\otimes\Delta^+\right)\Delta^-$$

where $\mathcal{M}^{(13)(2)(4)}$ is given as $\mathcal{M}^{(13)(2)(4)}\left(\tau_1\otimes\tau_2\otimes\tau_3\otimes\tau_4\right) = (\tau_1\odot\tau_3)\otimes\tau_2\otimes\tau_4$. This identity is both true on $\mathcal{T}$ through the comodule structures and on $\mathcal{T}_+$ when the coproduct Δ^- is viewed as an action on $\mathcal{T}_+$. We have already came across something similar in Lemma 2.3.15, but in that case the maps involved were not really coproducts. In our simple framework, this property is not satisfied if we just consider the reduced structure. One can circumvent this issue without introducing extended decorations by changing the coproduct Δ^+ to the form (2.30) given below. This approach is possible in our context (specifically, minimal degree $\alpha - 1 > -1$ and 1-regularizing kernel) because we know *a priori* that each edge type $\mathcal{I}$ in the elements of $\mathcal{W}$ with negative degree has the same "Taylor expansion" of length 1 in (2.28) ($\ell = 0$). In general, we would use the extended decorations to maintain this property, however, in the specific setting of the Heaviside kernel, *to which we will specialize from this moment on to the rest of the chapter,* we can just fix the length in the coproduct and not use the extended decorations. That is, we can get away by replacing (2.28) with the same formula, but only keeping $\ell = 0$ in the sum. Specifically, with $\mathcal{J} \equiv \mathcal{J}_0$ this amounts to make the (recursive) definition of Δ^+ with (2.28) replaced by

$$\Delta^+\mathcal{I}(\tau) = (\mathcal{I}\otimes\mathrm{id})\Delta^+\tau + 1\otimes\mathcal{J}(\tau). \tag{2.30}$$

We can also get rid of colours when we have no derivatives on the edges at the root: if we want to extract from $\mathcal{I}(\tau_1\Xi_i)\mathcal{I}(\tau_2\Xi_j)$ all the negative subtrees, we observe that it is not possible to extract one at the root, and thus are only left with negative subtrees in $\tau_1\Xi_i$ and $\tau_2\Xi_j$, which ensures that

$$M_\ell\mathcal{I}(\tau_1\Xi_i)\mathcal{I}(\tau_2\Xi_j) = \mathcal{I}\left(M_\ell\left(\tau_1\Xi_i\right)\right)\mathcal{I}\left(M_\ell\left(\tau_2\Xi_j\right)\right).$$

In the setting of [BHZ19], this multiplicativity property is encoded by a blue colour at the root which avoids the extraction of a tree containing the root, see Remark 4.15 there.

2.6.1.2 The case of rough differential equations

As in the last subsection: $n = 1$, $\beta = 1$ and noise degree $\alpha_{\min} \in (-1, 0) > -1$. We further specialize the algebraic set in that no symbols $\mathcal{J}_k$ and polynomials X^k with $k > 0$ are required in describing $\mathcal{T}_+$.

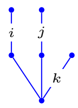
Assuming K to be the Heaviside step function, all derivatives (away from the origin) are zero, hence there is no need (with regard to $\mathcal{W}$) to have any polynomial symbols (X^k with $k > 0$). Removing these from $\mathcal{W}$ leaves us with $\tilde{\mathcal{W}} \subset \mathcal{W}$ which we may list as

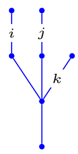
$$\begin{aligned}\tilde{\mathcal{W}} = \{&\Xi_i, ..., \mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)\Xi_k, ..., 1, \mathcal{I}(\Xi_i), \mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j), ... \\ &..., \mathcal{I}(\mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)\Xi_k), \mathcal{I}(\mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)), ..., \mathcal{I}()\mathcal{I}(), \mathcal{I}(\mathcal{I}()), ...\},\end{aligned} \tag{2.31}$$

(all indices are allowed to vary from $1, ..., d$), with associated degrees $|\tau|$ as follows:[10]

$$\alpha - 1, ..., 3\alpha - 1, ..., 0, \alpha, 2\alpha, ... \; ..., 3\alpha, 2\alpha + 1, ..., 2, 2, ...$$

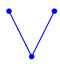
As in the case of $\mathcal{W}$, elements of $\tilde{\mathcal{W}}$ can be viewed as rooted trees, but *without* node decorations. For instance,

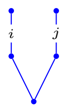


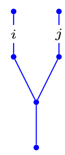
$\leftrightarrow \mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)\Xi_k$,


$\leftrightarrow \mathcal{I}(\mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)\Xi_k)$,

are trees ($\leftrightarrow$ symbols) contained in $\mathcal{W}$, and also in $\mathcal{W}_{\text{Hai14}}$, the symbols arising in the construction of [Hai14, Section 8.1], whereas


$\leftrightarrow \mathcal{I}()\mathcal{I}()$,

$\leftrightarrow \mathcal{I}(\mathcal{I}())$,


$\leftrightarrow \mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)$,


$\leftrightarrow \mathcal{I}(\mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j))$,

are contained in $\mathcal{W}$, following the above construction taken from [BHZ19], in order to obtain stability under the negative renormalization maps (but not included in $\mathcal{W}_{\text{Hai14}}$.)

A linear subspace of $\mathcal{T} = \text{span}_{\mathbb{R}} \mathcal{W}$ is then given by

$$\tilde{\mathcal{T}} := \text{span}_{\mathbb{R}} \tilde{\mathcal{W}}. \tag{2.32}$$

Symbols for negative renormalization

Recall that, thanks to $\beta = 1$, noise degree $\alpha - 1 \in (-1, 0)$, no terms X, X^2 or $\mathcal{I}(), ...$ arise as symbol in $\mathcal{W}_- := \{\tau \in \mathcal{W} \mid |\tau| < 0\}$. (As a consequence, replacing $\mathcal{W}$ by $\mathcal{W}_{\text{Hai14}}, \tilde{\mathcal{W}}$ or $\mathcal{W}_{\text{BHZ}}$ in the definition of the negative symbols makes no difference.) In particular,

$$\mathcal{W}_- = \{\Xi_i, \mathcal{I}(\Xi_i)\Xi_j, ..., \mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)\Xi_k, ...\}.$$

[10] tacitly assuming $\alpha < 1/3$

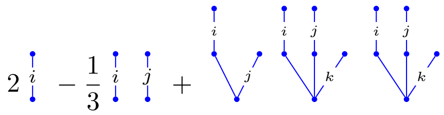
(where $\mathcal{W}_-$ "ends" right before the element 1 in (2.31) above) contains no powers of X, (hence no need to introduce "$\tilde{\mathcal{W}}_-$"). As previously defined (see (2.23)), we have

$$\mathcal{T}_- = \text{free unital commutative algebra generated by } \mathcal{W}_-.$$

For instance, writing $\odot$ for the (free, commutative) product in $\mathcal{T}_-$,

$$2\Xi_i - \frac{1}{3}\Xi_i \odot \Xi_j + \mathcal{I}(\Xi_i)\Xi_j \odot (\mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)\Xi_k)^{\odot 2} \quad \in \mathcal{T}_-.$$

Interpreting $\odot$ as the *forest product*, elements in $\mathcal{T}_-$ can then be represented as linear combinations of forests, such as



One can readily verify that $\Delta^- : \mathcal{T} \to \mathcal{T}_- \otimes \mathcal{T}$ restricted to $\tilde{\mathcal{T}}$ maps $\tilde{\mathcal{T}} \to \mathcal{T}_- \otimes \tilde{\mathcal{T}}$, also denoted by Δ^- so that $(\tilde{\mathcal{T}}, \Delta^-)$ is a subcomodule of $(\mathcal{T}, \Delta^-)$.

Symbols for positive renormalization and $\mathcal{T}_+$.

Recall that $\mathcal{T}_+$ was generated, as a free commutative algebra, by

$$\mathcal{W}_+ := \{X\} \cup \{\mathcal{J}_k\tau \mid \tau \in \mathcal{W}, |\tau| + 1 - k > 0\}.$$

Writing $\mathcal{J} \equiv \mathcal{J}_0$ as usual, we define a subset $\tilde{\mathcal{W}}_+ \subset \mathcal{W}_+$ as follows

$$\begin{aligned} \tilde{\mathcal{W}}_+ &:= \{\mathcal{J}\tau \mid \tau \in \tilde{\mathcal{W}}\} \qquad (2.33)\\ &= \{1, \mathcal{J}(\Xi_i), \mathcal{J}(\mathcal{I}(\Xi_i)\Xi_j), \mathcal{J}(\mathcal{I}(\mathcal{I}(\Xi_i)\Xi_j)\Xi_k), \mathcal{J}(\mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)\Xi_k), ..., \mathcal{J}(\mathcal{I}(\Xi_i)\mathcal{I}(\Xi_j)), ...\} \end{aligned}$$

with degrees $0, \alpha, 2\alpha, 3\alpha, 3\alpha, ..., 2\alpha + 1, ...$ here.

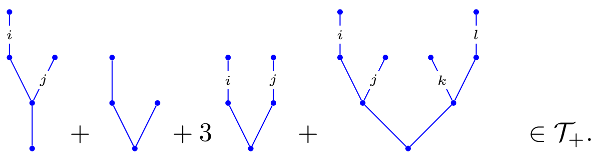
Recall that elements in $\mathcal{W}_+$ can be represented by *elementary* trees, in the sense that - disregarding the trivial (empty) tree 1 - only one edge departs from the root. The same is true for elements in $\tilde{\mathcal{W}}_+$. Set

$$\tilde{\mathcal{T}}_+ := \text{free unital commutative algebra generated by } \tilde{\mathcal{W}}_+.$$

For example, writing $\tau_1\tau_2$ for the (free, commutative) product of $\tau_1, \tau_2 \in \tilde{\mathcal{T}}_+$, an example of an element in this space would be

$$\mathcal{J}(\mathcal{I}(\Xi_i)\Xi_j) + \mathcal{J}(\mathcal{I}())\mathcal{J}(1) + 3\ \mathcal{J}(\Xi_i)\mathcal{J}(\Xi_j) + \mathcal{J}(\mathcal{I}(\Xi_i)\Xi_j)\mathcal{J}(\mathcal{I}(\Xi_k)\Xi_l) \quad \in \tilde{\mathcal{T}}_+.$$

Fortunately, every such element can still be represented as a tree; it suffices to interpret the free product in $\mathcal{T}_+$ as the "root-joining" product (which is possible since all constituting trees are elementary). The (abstract) unit element $1 \in \mathcal{T}_+$ is then indeed given by the (trivial) tree $\bullet \leftrightarrow X^0$, where we recall our convention to drop the node decoration "0". For instance, the above element becomes[11]



[11] Remark that $\mathcal{J}(1)$, which corresponds to the right branch of the second term, could also have been written as $\mathcal{J}()$, reflecting our convention to drop the decoration 0 from nodes (here: $1 \equiv X^0$). By the same logic, we could also write $\mathcal{I}()$, one of the symbols arising in $\mathcal{W}$, as $\mathcal{I}(1)$.

Remark 2.6.6. Though we used the same formalism to draw trees as in the case of $\tilde{\mathcal{W}}$ above, the interpretation here is slightly different in that all root-touching edges refer to $\mathcal{J}$ rather than $\mathcal{I}$. As mentioned before, in [BHZ19], this is indicated by a blue colouring of the root.

As before, we define a coaction of $\tilde{\mathcal{T}}_+$ on $\tilde{\mathcal{T}}$ (which we again denote $\Delta^+ : \tilde{\mathcal{T}} \to \tilde{\mathcal{T}} \otimes \tilde{\mathcal{T}}_+$) by (2.27), (2.29), and (2.30) as well as a coproduct $\Delta^+ : \tilde{\mathcal{T}}_+ \to \tilde{\mathcal{T}}_+ \otimes \tilde{\mathcal{T}}_+$ defined in the same way, but with $\mathcal{I}$ changed to $\mathcal{J}$ in (2.30). (In contrast to the case of Δ^- discussed above, it is not the case that $(\tilde{\mathcal{T}}, \Delta^+)$ is a subcomodule of $(\mathcal{T}, \Delta^+)$.)

We note already that $(\tilde{\mathcal{T}}_+, \Delta^+)$ is isomorphic to the Connes-Kreimer Hopf algebra $\mathcal{H}$ arising from the identifications laid out in the following subsection (and which will be used crucially in the proof of the upcoming Proposition 2.6.8).

2.6.2 Link with translation of rough paths

2.6.2.1 Identification of spaces

We now give a precise description the map Δ^- in our context as well as its connection to the map δ from Section 2.3.3. To do so, we first need to introduce several identifications of vector spaces and algebras, as well as appropriately identify branched rough paths as models on a regularity structure.

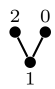
Recall the space $\mathcal{H} = \mathcal{H}(\bullet_0, ..., \bullet_d)$ from Section 2.3 spanned by labelled forests with label set $\{0, 1, \dots, d\}$. Consider now the enlarged vector space

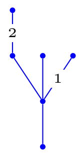
$$\tilde{\mathcal{H}} := \mathcal{H} \oplus \mathcal{H}\Xi_1 \oplus .... \oplus \mathcal{H}\Xi_d. \tag{2.34}$$

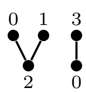
driven by branched rough paths (c.f. [Pre16, Section 5.3.2]). With $\tilde{\mathcal{T}}$ as defined in (2.32), and in particular with noise types $\Xi_1, ..., \Xi_d$, we then have a vector space isomorphism

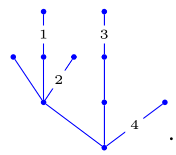
$$\tilde{\mathcal{H}} \leftrightarrow \tilde{\mathcal{T}}$$

obtained by adding an extra edge to indicate a noise Ξ_i, $i \neq 0$, and by "forgetting" the label 0 (which is equivalent to setting the noise Ξ_0 to the constant 1). For example,



$\leftrightarrow \mathcal{I}\left[\mathcal{I}(\Xi_2)\mathcal{I}(1)\Xi_1\right] =$




$\Xi_4 \leftrightarrow \mathcal{I}\left[\mathcal{I}(1)\mathcal{I}(\Xi_1)\Xi_2\right]\mathcal{I}\left[\mathcal{I}(\Xi_3)\right]\Xi_4 =$


.

Recall that $\mathcal{B} = \mathcal{B}(\bullet_0, ..., \bullet_d)$ denotes the subspace of $\mathcal{H}$ spanned by trees, and define

$$\mathcal{B}_- = \mathcal{B}_-(\bullet_1, ..., \bullet_d) \subset \mathcal{B} \subset \mathcal{H}$$

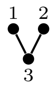
as the subspace of $\mathcal{B}$ spanned by trees with no label 0 and with at most $\lfloor 1/\alpha \rfloor$ nodes. Observe that there is a canonical vector space isomorphism

$$\phi : \mathcal{B}_- \to \langle \mathcal{W}_- \rangle \quad \subset \quad \mathcal{H}\Xi_1 \oplus .... \oplus \mathcal{H}\Xi_d \quad \subset \quad \tilde{\mathcal{H}}, \tag{2.35}$$

where we have used the identification $\tilde{\mathcal{H}} \leftrightarrow \tilde{\mathcal{T}} \supset \mathrm{span}_{\mathbb{R}} \mathcal{W}_-$ for the first inclusion (and both inclusions being strict: for the first, just consider the element  $\Xi_1 \notin \langle \mathcal{W}_- \rangle$). We denote this isomorphism also by

$$\tau \mapsto \dot{\tau} := \phi(\tau).$$

For example,

$$\phi :$$  $$\mapsto \mathcal{I}(\Xi_1)\mathcal{I}(\Xi_2)\Xi_3,$$

where we assume $\alpha \in (0, 1/3)$ so the tree appearing on the left is indeed an element in $\mathcal{B}_-$. Correspondingly, the symbol on the right has negative degree as an element of $\mathcal{W}$, hence is an element of $\mathcal{W}_-$.

Write $\mathcal{B}_-^*$ for the dual of the (finite-dimensional) vector space $\mathcal{B}_-$. Of course, $\mathcal{B}_-^* \cong \mathcal{B}_-$ which allows us to identify $\mathcal{B}_-^*$ with $\langle \mathcal{W}_- \rangle$. Recall that $(\mathcal{T}_-, \odot)$ was defined as the free unital commutative algebra generated by $\mathcal{W}_-$, and let $\mathcal{G}_- \subset \mathcal{T}_-^*$ denote the group of characters on $\mathcal{T}_-$. By definition of $\mathcal{T}_-$, we then have a bijection

$$\mathcal{B}_-^* \leftrightarrow \mathcal{G}_-. \tag{2.36}$$

To be fully explicit about this, recall that

$$\mathcal{T}_- = \langle \dot{\tau}_1 \odot .... \odot \dot{\tau}_n : \dot{\tau}_i \in \mathcal{W}_-, \ n = 1, 2, ... \rangle,$$

so writing $\tau_i = \phi^{-1}(\dot{\tau}_i) \in \mathcal{B}_-$, we have that associated to $v \in \mathcal{B}_-^*$ the character $\ell \in \mathcal{G}_-$ given explicitly by the formula

$$\ell(\dot{\tau}_1 \cdot .... \cdot \dot{\tau}_n) = \ell(\dot{\tau}_1)...\ell(\dot{\tau}_n) = \langle v, \tau_1 \rangle ... \langle v, \tau_n \rangle.$$

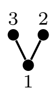
Define now

$$(\mathcal{H}_-, \cdot)$$

as the free commutative algebra generated by the subspace $\mathcal{B}_-$ of $\tilde{\mathcal{H}}$ (remark that the product in $\mathcal{H}_-$ has nothing to do with the product in $\mathcal{H}$ itself), so that there is an algebra isomorphism

$$\mathcal{H}_- \leftrightarrow \mathcal{T}_-.$$

A typical element of $\mathcal{H}_-$ looks like:

 $\Xi_1 + \Xi_2 + \Xi_2 \cdot$  $\Xi_3,$

whereas one has  $\notin \mathcal{H}_-$.

Note that we can also make the identification of algebras

$$\mathcal{H} \leftrightarrow \tilde{\mathcal{T}}_+.$$

For instance, using the bracket notation,

$$[\bullet_0]_{\bullet_0}\,\bullet_0 + [\bullet_i]_{\bullet_j}[\bullet_k]_{\bullet_l} \leftrightarrow \mathcal{J}(\mathcal{I}())\mathcal{J}(1) + \mathcal{J}(\mathcal{I}(\Xi_i)\Xi_j)\mathcal{J}(\mathcal{I}(\Xi_k)\Xi_l) \quad \in \tilde{\mathcal{T}}_+.$$

We denote by $\tilde{\mathcal{G}}_+ \subset \tilde{\mathcal{T}}_+^*$ the characters on $\tilde{\mathcal{T}}_+$ and note that there is also a bijection $\mathcal{G} \leftrightarrow \tilde{\mathcal{G}}_+$, where we recall that $\mathcal{G} \subset \mathcal{H}^*$ is the Butcher group over $\mathbb{R}^{1+d}$, i.e., the set of characters on $\mathcal{H}$.

To summarise, we have the following identifications in place

$$\begin{aligned}
\tilde{\mathcal{H}} &\leftrightarrow \tilde{\mathcal{T}},\\
\mathcal{H}_- &\leftrightarrow \mathcal{T}_-,\\
\mathcal{H} &\leftrightarrow \tilde{\mathcal{T}}_+,\\
\mathcal{B}_-^* &\leftrightarrow \langle \mathcal{W}_- \rangle \leftrightarrow \mathcal{G}_- \subset \mathcal{T}_-^*\\
\mathcal{G} &\leftrightarrow \tilde{\mathcal{G}}_+ \subset \tilde{\mathcal{T}}_+^*.
\end{aligned}$$

2.6.2.2 Renormalization as rough path translations

It now only remains to identify (a family of) branched rough paths with a class of models on a suitable regularity structure. Define the index set $A := \{0\} \cup \alpha\mathbb{N} \cup (\alpha\mathbb{N} - 1)$. Recall that the action of $g \in \tilde{\mathcal{G}}_+$ on $\tilde{\mathcal{T}}$ is given exactly as before by

$$\Gamma_g \tau = (\mathrm{id} \otimes g)\Delta^+ \tau, \ \text{ for all } \tau \in \tilde{\mathcal{T}}.$$

Note that Γ_g indeed maps $\tilde{\mathcal{T}}$ to itself due to the definition of $\tilde{\mathcal{G}}_+$. Note further that $\Gamma_g \Gamma_h$ (as a composition of linear maps) is exactly $\Gamma_{g\circ h}$ (with $\circ$ the product in $\tilde{\mathcal{G}}_+$ given as the dual of Δ^+), and so

$$G := \{\Gamma_g : g \in (\tilde{\mathcal{G}}_+, \circ)\}.$$

is indeed a group of endomorphisms of $\tilde{\mathcal{T}}$.

Recall now the definition of a regularity structure from [Hai14, Definition 2.1] (see also Section 1.3 of this thesis).

Lemma 2.6.7. *The triplet $(A, \tilde{\mathcal{T}}, G)$ is a regularity structure.*

Proof. The only non-trivial property to check is that for all $\tau \in \tilde{\mathcal{T}}$ of degree $\alpha \in A$ and $\Gamma \in G$, $\Gamma\tau - \tau$ is a linear combination of terms of degree strictly less than α, which in turn is a direct consequence of the definition of $\Delta^+ : \tilde{\mathcal{T}} \to \tilde{\mathcal{T}} \otimes \tilde{\mathcal{T}}_+$ from (2.30) (see end of Section 2.6.1.2). □

Recall also the definition of a model on a regularity structure (see [Hai14, Definition 2.17] and Section 1.3 of this thesis). Let $\mathscr{M}_{[0,T]}$ denote the set of all models (Π, Γ) for $(A, \tilde{\mathcal{T}}, G)$ on $\mathbb{R}$ such that (cf. [Pre16, Section 5.3.2])

(i) $\Pi_t 1$ is the constant function 1 for all $t \in \mathbb{R}$,

(ii) $\Gamma_{st} = \mathrm{id}$ for $s, t \in (-\infty, 0]$ and for $s, t \in [T, \infty)$,

(iii) $(\Pi_t \mathcal{I} y)' = \Pi_t y$ for all $t \in \mathbb{R}$ and $y \in \tilde{\mathcal{T}}$. (Here $(..)'$ denotes the Schwartz derivative.).

On the other hand, let $\mathscr{R}^{\alpha}_{[0,T]}$ be the set of all $(1+d)$-dimensional α-Hölder branched rough paths $\mathbf{X} : [0,T]^2 \to \mathcal{G}$ whose zeroth component is time, i.e., $\langle \mathbf{X}_{s,t}, \bullet_0 \rangle = t - s$ and

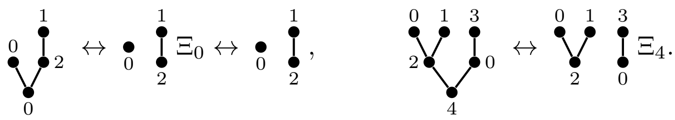
$$\langle \mathbf{X}_{s,t}, [\tau]_{\bullet_0} \rangle = \int_s^t \langle \mathbf{X}_{s,u}, \tau \rangle du, \quad \text{for all } \tau \in \mathcal{H},\ s,t \in [0,T]. \tag{2.37}$$

Observe that this condition necessarily implies that $\mathbf{X}$ satisfies condition (2.16) from Theorem 2.5.1 (cf. Remark 2.5.4). Note that $\mathbf{X}_{s,t}$ can be identified with an element of $\tilde{\mathcal{G}}_+$ due to the identification $\mathcal{G} \leftrightarrow \tilde{\mathcal{G}}_+$,.

Finally, observe that ϕ defined in (2.35) may be extended to a vector space isomorphism

$$\phi : \mathcal{B} \leftrightarrow \mathcal{H}\Xi_0 \oplus \mathcal{H}\Xi_1 \oplus \dots \oplus \mathcal{H}\Xi_d \cong \mathcal{H} \oplus \mathcal{H}\Xi_1 \oplus \dots \oplus \mathcal{H}\Xi_d \equiv \tilde{\mathcal{H}} \tag{2.38}$$

which maps a tree $\tau \in \mathcal{B}$ into a forest $\phi(\tau) \equiv \dot{\tau}$, as illustrated in the following two examples:



Conversely, ϕ^{-1} adds an extra node (which becomes the root) and should be thought of as taking the integral of a symbol in $\tilde{\mathcal{H}}$. The following result makes this precise by giving a bijection between $\mathscr{M}_{[0,T]}$ and $\mathscr{R}^{\alpha}_{[0,T]}$.

Proposition 2.6.8 (cf. [Pre16, Theorem 5.15])**.** *There is a bijective map $I : \mathscr{R}^{\alpha}_{[0,T]} \to \mathscr{M}_{[0,T]}$ which maps a branched rough path $\mathbf{X}$ to the unique model $(\Pi, \Gamma) \in \mathscr{M}_{[0,T]}$ with the property that*

$$(\Pi_s \mathcal{I}\dot{\tau})(t) = \langle \mathbf{X}_{s,t}, \tau \rangle \quad \textit{for all } \tau \in \mathcal{B},\ \ s,t \in [0,T],$$

where we have made the identifications $\phi(\tau) \equiv \dot{\tau} \in \tilde{\mathcal{H}} \leftrightarrow \tilde{\mathcal{T}}$. Furthermore, the model (Π, Γ) satisfies $\Gamma_{ts} = \Gamma_{\mathbf{X}_{s,t}}$ (where we have made the identification $\mathbf{X}_{s,t} \in \mathcal{G} \cong \tilde{\mathcal{G}}_+$) and the multiplicativity property

$$\Pi_t((\mathcal{I}y_1)\dots(\mathcal{I}y_n)) = \Pi_t(\mathcal{I}y_1)\dots\Pi_t(\mathcal{I}y_n), \quad \textit{for all } n \in \mathbb{N},\ \ y_i \in \tilde{\mathcal{T}}. \tag{2.39}$$

Proof. Consider $\mathbf{X} \in \mathscr{R}^{\alpha}_{[0,T]}$. For all $s,t \in [0,T]$ define $\Gamma_{ts} = \Gamma_{\mathbf{X}_{s,t}}$ and $(\Pi_s \mathcal{I}\dot{\tau})(t) = \langle \mathbf{X}_{s,t}, \tau \rangle$ for all $\tau \in \mathcal{B}$. Observe that we may further impose on (Π, Γ) that properties (i) and (ii) hold. Furthermore, for every $\tau \notin \mathcal{I}\tilde{\mathcal{T}}$, we may define $\Pi_t \tau = (\Pi_t \mathcal{I}\tau)'$, which completely characterises Π. It remains to verify (2.39), that property (iii) holds for all $\tau \in \mathcal{I}\tilde{\mathcal{T}}$, and that (Π, Γ) is indeed a model.

For (2.39), note that from (2.37) we have

$$\begin{aligned}
\Pi_t(\mathcal{I}\dot{\tau}_1 \dots \mathcal{I}\dot{\tau}_n) &= (\Pi_t \mathcal{I}(\mathcal{I}\dot{\tau}_1 \dots \mathcal{I}\dot{\tau}_n))' \\
&= (\langle \mathbf{X}_{t,\cdot}, \phi^{-1}(\mathcal{I}\dot{\tau}_1 \dots \mathcal{I}\dot{\tau}_n) \rangle)' \\
&= (\langle \mathbf{X}_{t,\cdot}, [\tau_1 \dots \tau_n]_{\bullet_0} \rangle)' \\
&= \langle \mathbf{X}_{t,\cdot}, \tau_1 \dots \tau_n \rangle \\
&= \langle \mathbf{X}_{t,\cdot}, \tau_1 \rangle \dots \langle \mathbf{X}_{t,\cdot}, \tau_n \rangle = \Pi_t(\mathcal{I}\dot{\tau}_1) \dots \Pi_t(\mathcal{I}\dot{\tau}_n).
\end{aligned}$$

To show property (iii) for $\dot{\tau} = \mathcal{I}\dot{\bar{\tau}} \in \mathcal{I}\tilde{\mathcal{T}}$, where $\dot{\bar{\tau}} \in \tilde{\mathcal{T}}$, observe that $\phi([\bar{\tau}]_{\bullet_0}) = \dot{\tau}$, so that again by (2.37)

$$\begin{aligned}\Pi_t\dot{\tau} &= \Pi_t\mathcal{I}\dot{\bar{\tau}}\\ &= \langle \mathbf{X}_{t,\cdot}, \bar{\tau}\rangle\\ &= (\langle \mathbf{X}_{t,\cdot}, [\bar{\tau}]_{\bullet_0}\rangle)'\\ &= (\Pi_t\mathcal{I}\phi([\bar{\tau}]_{\bullet_0}))'\\ &= (\Pi_t\mathcal{I}\dot{\tau})'.\end{aligned}$$

It remains to show that (Π, Γ) is a model. We first verify that $\Pi_s\Gamma_{s,t} = \Pi_t$. Let $\tau \in \mathcal{B}$, so that $\mathcal{I}(\dot{\tau}) \in \tilde{\mathcal{T}}$. Recall that the Connes-Kreimer coproduct $\Delta_\star : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ as was introduced in Section 2.3.1 can be defined recursively by

$$\Delta_\star[\tau_1 \dots \tau_n]_{\bullet_i} = [\tau_1 \dots \tau_n]_{\bullet_i} \otimes 1 + (\mathrm{id} \otimes [\cdot]_{\bullet_i})\Delta_\star(\tau_1 \dots \tau_n), \text{ for all } \tau_1, \dots, \tau_n \in \mathcal{B},\ i \in \{0, \dots, d\}.$$

With this recursion, one can verify that

$$\Delta^+ : \mathcal{I}(\tilde{\mathcal{T}}) \to \mathcal{I}(\tilde{\mathcal{T}}) \otimes \tilde{\mathcal{T}}_+$$

agrees with the "reversed" Connes-Kreimer coproduct

$$\sigma_{1,2}\Delta_\star : \mathcal{B} \to \mathcal{B} \otimes \mathcal{H},$$

where $\sigma_{1,2} : \mathcal{H} \otimes \mathcal{B} \to \mathcal{B} \otimes \mathcal{H}$, $\sigma_{1,2} : \tau \otimes \bar{\tau} \mapsto \bar{\tau} \otimes \tau$, and where we make the usual identification $\mathcal{H} \leftrightarrow \tilde{\mathcal{T}}_+$ as well as $\phi^{\mathcal{I}} : \mathcal{B} \to \mathcal{I}(\tilde{\mathcal{T}})$ via $\phi^{\mathcal{I}} : \tau \mapsto \mathcal{I}(\dot{\tau})$ (which is of course just $\mathcal{I} \circ \phi$). Therefore, treating $\mathbf{X}_{s,t}$ as a character on $\mathcal{H} \leftrightarrow \tilde{\mathcal{T}}_+$, we have for all $\tau \in \mathcal{B}$

$$\begin{aligned}(\Pi_t\Gamma_{ts}\mathcal{I}\dot{\tau})(u) &= (\Pi_t(\mathrm{id} \otimes \mathbf{X}_{s,t})\Delta^+\mathcal{I}\dot{\tau})(u)\\ &= \langle \mathbf{X}_{t,u}, (\phi^{\mathcal{I}})^{-1}(\mathrm{id} \otimes \mathbf{X}_{s,t})\Delta^+\mathcal{I}\dot{\tau})\rangle\\ &= \langle \mathbf{X}_{t,u}, (\mathbf{X}_{s,t} \otimes \mathrm{id})\Delta_\star\tau)\rangle\\ &= \langle \mathbf{X}_{s,t} \otimes \mathbf{X}_{t,u}, \Delta_\star\tau\rangle\\ &= \langle \mathbf{X}_{s,t} \bullet \mathbf{X}_{t,u}, \tau\rangle\\ &= \langle \mathbf{X}_{s,u}, \tau\rangle\\ &= \Pi_s(\mathcal{I}\dot{\tau})(u).\end{aligned} \tag{2.40}$$

Observe now that for $\tau \in \tilde{\mathcal{T}}$, we have

$$\Gamma_{ts}\mathcal{I}\tau = \mathcal{I}\Gamma_{ts}\tau + \langle \mathbf{X}_{s,t}, \mathcal{I}\tau\rangle 1,$$

where we emphasize the symbol $1 \in \tilde{\mathcal{T}}$. Therefore, by the (already established) properties (i) and (iii), it follows that for any $\tau \in \tilde{\mathcal{T}}$

$$\Pi_t\Gamma_{ts}\tau = (\Pi_t\mathcal{I}\Gamma_{ts}\tau)' = (\Pi_t(\Gamma_{ts}\mathcal{I}\tau - \langle \mathbf{X}_{s,t}, \mathcal{I}\tau\rangle 1))' = (\Pi_t\Gamma_{ts}\mathcal{I}\tau)' = (\Pi_s\mathcal{I}\tau)' = \Pi_s\tau,$$

which shows that $\Pi_t\Gamma_{ts} = \Pi_s$.

It remains to verify the analytic bounds on (Π, Γ). As in Theorem 2.5.1, denote by $|\tau|$ the number of nodes in τ and by $|\tau|_0$ the number of nodes with the label 0. It follows that the degree of $\mathcal{I}\dot{\tau}$ is given by $|\mathcal{I}\dot{\tau}| = |\tau|_0(1-\alpha) + |\tau|\alpha$. Since $\mathbf{X}$ satisfies (2.16), we have the analytic bound

$$|(\Pi_s\mathcal{I}\dot{\tau})(t)| = |\langle \mathbf{X}_{s,t}, \tau\rangle| \lesssim |t-s|^{|\mathcal{I}\dot{\tau}|}.$$

Since $\Pi_s\tau = (\Pi_s\mathcal{I}\tau)'$ by property (iii), we see that Π satisfies the correct analytic bounds. The exact same argument applies to Γ upon using the identification of Δ^+ with $\sigma_{1,2}\Delta_\star$ above. Therefore (Π,Γ) is a model in $\mathscr{M}_{[0,T]}$ as claimed.

Finally, it remains to observe that we may reverse the construction. Indeed, starting with a model (Π,Γ) in $\mathscr{M}_{[0,T]}$, we may define $\mathbf{X}$ by $\langle \mathbf{X}_{s,t},\tau\rangle = (\Pi_s\mathcal{I}\dot\tau)(t)$. The fact that $\mathbf{X}$ satisfies (2.37) follows from property (iii), while the required analytic bounds for $\mathbf{X}$ to be an α-Hölder branched rough path follow from the analytic bounds associated to Π. To conclude, it suffices to verify that $\mathbf{X}$ thus defined satisfies $\Gamma_{ts} = \Gamma_{\mathbf{X}_{s,t}}$ and $\mathbf{X}_{s,t}\bullet\mathbf{X}_{t,u} = \mathbf{X}_{s,u}$. To this end, note that by definition of the structure group G, there exists $\gamma_{ts}\in\tilde{\mathcal{G}}_+\cong\mathcal{G}$ such that $\Gamma_{ts} = (\mathrm{id}\otimes\gamma_{ts})\Delta^+$. Let $\tilde{\mathbf{X}}_{s,t}\in\mathcal{G}$ be the element associated to γ_{ts} in the identification $\tilde{\mathcal{G}}_+\cong\mathcal{G}$, and we aim to show $\tilde{\mathbf{X}}_{s,t} = \mathbf{X}_{s,t}$. Indeed, from our identification $\mathcal{H}\leftrightarrow\tilde{\mathcal{T}}_+$, it follows that for all $\tau\in\mathcal{B}$

$$\langle\gamma_{ts},\mathcal{J}\dot\tau\rangle = \langle\tilde{\mathbf{X}}_{s,t},\tau\rangle.$$

On the other hand, we know that for all $\tau\in\mathcal{B}$

$$\langle\mathbf{X}_{s,t},\tau\rangle = (\Pi_s\mathcal{I}\dot\tau)(t) = (\Pi_t\Gamma_{ts}\mathcal{I}\dot\tau)(t) = (\Pi_t(\mathrm{id}\otimes\gamma_{ts})\Delta^+\mathcal{I}\dot\tau)(t) = \langle\gamma_{ts},\mathcal{J}\dot\tau\rangle,$$

where for the last equality we have used property (i) and the fact that

$$\Delta^+\mathcal{I}\dot\tau = 1\otimes\mathcal{J}\dot\tau + \sum\mathcal{I}(\dot\tau^{(1)})\otimes\dot\tau^{(2)},$$

where every term $\mathcal{I}(\dot\tau^{(1)})$ is of positive degree, and so $(\Pi_t\mathcal{I}(\dot\tau^{(1)}))(t) = 0$. This concludes the proof that $\Gamma_{ts} = \Gamma_{\mathbf{X}_{s,t}}$. To verify that $\mathbf{X}_{s,t}\bullet\mathbf{X}_{t,u} = \mathbf{X}_{s,u}$, we can now simply reorder the sequence of equalities (2.40). □

Following [BHZ19, Equation (6.18)] we introduce the *renormalization map* M_ℓ given by[12]

$$\mathrm{M}_\ell : \tilde{\mathcal{T}}\to\tilde{\mathcal{T}},\quad \tau\mapsto(\ell\otimes\mathrm{id})\,\Delta^-\tau,$$

for a given character $\ell\in\mathcal{G}_-\subset\mathcal{T}_-^*$. In our case, we have the fact that M_ℓ commutes with $\mathcal{I}$ (cf. end of Remark 2.6.5)

$$\mathrm{M}_\ell\mathcal{I} = \mathcal{I}\mathrm{M}_\ell, \tag{2.41}$$

which is readily verified by hand: $\mathcal{I}$ amounts to adding another edge to the root (thereby creating a new root), whereas M_ℓ amounts to extracting (negative) subtrees and mapping them to $\mathbb{R}$ (via ℓ). Clearly, the afore-mentioned edge (of degree 1) can not possibly be part of any singular subtree, hence the desired commutation.

This map acts on a model $\mathbf{\Pi} = (\Pi,\Gamma)$ and yields the *renormalised model* (see [BHZ19, Theorem 6.16]) given by

$$\Pi_s^{\mathrm{M}_\ell} := \Pi_s\mathrm{M}_\ell,\quad \Gamma_{t,s}^{\mathrm{M}_\ell} = \left(\mathrm{id}\otimes\gamma_{t,s}^{\mathrm{M}_\ell}\right)\Delta^+,\quad \gamma_{t,s}^{\mathrm{M}_\ell} = \gamma_{t,s}\mathrm{M}_\ell.$$

Recall from Section 2.3.3 the map $\delta:\mathcal{B}\to\mathcal{A}\otimes\mathcal{B}$, where $\mathcal{A}$ is the free commutative algebra generated by $\mathcal{B}$ (thought of as an isomorphic but different space to $\mathcal{H}$). Recall also the (vector space) isomorphism $\phi:\tau\mapsto\dot\tau$ as detailed in (2.38) with which we identify $\tilde{\mathcal{H}}\cong\mathcal{B}$. Let $\pi_-:\tilde{\mathcal{H}}\cong\mathcal{B}\to\mathcal{B}_-\cong\langle\mathcal{W}_-\rangle$ denote the projection onto terms of negative degree, which we extend multiplicatively to an algebra homomorphism $\pi_-:\mathcal{A}\to\mathcal{H}_-$. We now define the map

$$\delta^- = (\pi_-\otimes\mathrm{id})\delta : \tilde{\mathcal{H}}\to\mathcal{H}_-\otimes\tilde{\mathcal{H}}.$$

[12]While we deliberately used the same letter, do not confuse $\mathrm{M}_\ell:\tilde{\mathcal{T}}\to\tilde{\mathcal{T}}$ with $M_v:\mathcal{H}^*\to\mathcal{H}^*$.

For instance

$$\delta^{-} \bullet_0 = 1 \otimes \bullet_0,$$

whereas

$$\delta \bullet_0 = \bullet_0 \otimes \bullet_0 + 1 \otimes \bullet_0.$$

We are now ready to state the link between translation of branched rough paths and negative renormalization in the following two results.

Lemma 2.6.9. (i) *For all $\tau \in \mathcal{B}$ it holds that*

$$\Delta^{-} \dot{\tau} = \Delta^{-} \phi(\tau) = (\phi \otimes \phi)\, \delta^{-}(\tau).$$

(ii) *Let v be an element of $\mathcal{B}_{-}^{*}$ and let $\ell \in \mathcal{G}_{-}$ by the associated element in $\mathcal{G}_{-} \subset \mathcal{T}_{-}^{*}$, as was detailed in* (2.36)*. Then*

$$\mathrm{M}_{\ell} \dot{\tau} = \mathrm{M}_{\ell} \phi(\tau) = \phi(M_v^{*} \tau)$$

Proof. (i) Let us consider $[\tau]_{\bullet_i} \in \mathcal{B}$. We then have the following identities:

$$\Delta^{-} \phi([\tau]_{\bullet_i}) = \Delta^{-} \tau \Xi_i = \sum_{C = A \cdot B \subset \tau} \left(C \otimes (\mathcal{R}_C \tau)\Xi_i + A \cdot B \Xi_i \otimes \mathcal{R}_C \tau\right). \tag{2.42}$$

The sum is taken over all the couples (A, B) where A is a negative subforest of τ which does not include the root of τ and B is a subtree of τ at the root disjoint from A. In the sum in (2.42), the first term means that Ξ_i does not belong to the tree extracted at the root, while for the second term, Ξ_i belongs to the tree which comes from the product between Ξ_i and B giving a subtree of negative degree. One can derive the same identity for δ^{-}. We first rewrite δ^{-}:

$$\delta^{-} \tau = \sum_{A \subset \tau} A \otimes \tilde{\mathcal{R}}_A \tau,$$

where A is a subforest of τ and $\tilde{\mathcal{R}}_A \tau$ means that we contract the trees of A in τ and we leave a 0 decoration on their roots. Then the equivalent of (2.42) in that context is given by:

$$\delta^{-} [\tau]_{\bullet_i} = \sum_{\tilde{C} = \tilde{A} \cdot \tilde{B} \subset \tau} \left(\tilde{C} \otimes [\tilde{\mathcal{R}}_{\tilde{C}} \tau]_{\bullet_i} + \tilde{A} \cdot [\tilde{B}]_{\bullet_i} \otimes \tilde{\mathcal{R}}_{\tilde{A} \cdot [\tilde{B}]_{\bullet_i}} [\tau]_{\bullet_i}\right)$$

$$(\phi \otimes \phi)\, \delta^{-} [\tau]_{\bullet_i} = \sum_{\tilde{C} = \tilde{A} \cdot \tilde{B} \subset \tau} \left(\phi(\tilde{C}) \otimes (\tilde{\mathcal{R}}_{\tilde{C}} \tau)\Xi_i + \phi(\tilde{A}) \cdot \tilde{B} \Xi_i \otimes \phi\left(\tilde{\mathcal{R}}_{\tilde{A} \cdot [\tilde{B}]_{\bullet_i}} [\tau]_{\bullet_i}\right)\right).$$

Now we have the following identifications:

$$\phi(\tilde{C}) \leftrightarrow C, \quad \tilde{B}\Xi_i \leftrightarrow B\Xi_i, \quad \phi\left(\tilde{\mathcal{R}}_{\tilde{A} \cdot [\tilde{B}]_{\bullet_i}} [\tau]_{\bullet_i}\right) = \tilde{\mathcal{R}}_{\tilde{C}} \tau \leftrightarrow \mathcal{R}_C \tau, \quad (\tilde{\mathcal{R}}_{\tilde{C}} \tau)\Xi_i \leftrightarrow (\mathcal{R}_C \tau)\Xi_i,$$

which gives the result.

(ii) Recall that $\delta^{-}(\tau)$ has an image of the form "forest $\otimes$ tree", and that $\ell \circ \phi = v$ (which is a "dual" tree and multiplicative over forests). Also note that $M_v^{*} \tau = (v \otimes \mathrm{id})\, \delta = (v \otimes \mathrm{id})\, \delta^{-}$ whenever $v \in \mathcal{B}_{-}^{*}$ (which not true for general $v \in \mathcal{B}^{*}$), so that

$$\begin{aligned} \mathrm{M}_{\ell} \dot{\tau} &= (\ell \otimes \mathrm{id})\, \Delta^{-} \dot{\tau} \\ &= (\ell \otimes \mathrm{id})\, \Delta^{-} \phi(\tau) \\ &= (v \otimes \phi)\, \delta^{-}(\tau) \\ &= \phi\left((v \otimes \mathrm{id})\, \delta^{-}\right) \\ &= \phi(M_v^{*} \tau). \end{aligned}$$

□

Theorem 2.6.10. (i) *It holds that the restriction* $\Delta^- : \tilde{\mathcal{T}} \to \mathcal{T}_- \otimes \tilde{\mathcal{T}}$ *coincides with* $\delta^- : \tilde{\mathcal{H}} \to \mathcal{H}_- \otimes \tilde{\mathcal{H}}$*, where we have made the identifications* $\tilde{\mathcal{H}} \leftrightarrow \tilde{\mathcal{T}}$ *and* $\mathcal{H}_- \leftrightarrow \mathcal{T}_-$ *as above.*

(ii) *Let* v *be an element of* $\mathcal{B}_-^*$ *and let* $\ell \in \mathcal{G}_-$ *by the associated element in* $\mathcal{G}_- \subset \mathcal{T}_-^*$*, as was detailed in* (2.36)*. Then the following diagram commutes*

$$\begin{array}{ccc} \mathbf{X} & \longleftrightarrow & \mathbf{\Pi} \\ \downarrow & & \downarrow \\ M_v\mathbf{X} & \longleftrightarrow & \mathbf{\Pi}^{\mathrm{M}_\ell} \end{array}$$

(iii) *For* $v, v' \in \mathcal{B}_-$ *with associated characters* $\ell, \ell' \in \mathcal{G}_-$*, it holds that the character associated to* $v + v'$ *is* $\ell \circ \ell'$*, so that* $(\mathcal{B}_-, +) \cong (\mathcal{G}_-, \circ)$.

Remark 2.6.11. In view of the final statement of Theorem 2.5.1 part (ii), we see that the commuting diagram in (ii) holds upon replacing M_v by any algebra homomorphism $M : \mathcal{H}^* \to \mathcal{H}^*$ which preserves $\mathcal{B}^*$, leaves invariant every forest without a label 0, and satisfies $M\bullet_0 = \bullet_0 + v$.

Remark 2.6.12. The final statement (iii) effectively says that the renormalization group associated to branched rough paths is always abelian, despite the highly non-commutative nature of the Grossman-Larson Hopf algebra $\mathcal{H}^*$.

Proof of Theorem 2.6.10. Part (i) is a just of a reformulation of Lemma 2.6.9 (i). To verify part (ii), in view of Proposition 2.6.8, we only need to check that for all $\tau \in \mathcal{B}$

$$\Pi_s^{\mathrm{M}_l}\mathcal{I}\dot{\tau} = \langle M_v\mathbf{X}_{s,\cdot}, \tau\rangle = \langle \mathbf{X}_{s,\cdot}, M_v^*\tau\rangle\,.$$

The LHS can be rewritten as, thanks to (2.41) and Lemma 2.6.9 (ii)

$$\begin{aligned} \Pi_s^{\mathrm{M}_\ell}\mathcal{I}\dot{\tau} &= \Pi_s\mathrm{M}_\ell\mathcal{I}\dot{\tau} \\ &= \Pi_s\mathcal{I}\mathrm{M}_\ell\dot{\tau} \\ &= \Pi_s\mathcal{I}\phi\left(M_v^*\tau\right). \end{aligned}$$

Applying Proposition 2.6.8 with $\dot{\tau} = \phi\left(M_v\tau\right)$ then shows that

$$\Pi_s\mathcal{I}\phi\left(M_v\tau\right) = \langle \mathbf{X}_{s,\cdot}, M_v^*\tau\rangle$$

which is what we wanted to show.

Finally, to show (iii), we note that

$$\langle \ell \circ \ell', \tau\rangle = \langle \ell \otimes \ell', \Delta^-\tau\rangle = \langle \ell, \tau\rangle + \langle \ell', \tau\rangle, \quad \text{for all } \tau \in \mathcal{W}_-,$$

where the first equality follows by definition and the second from the fact that every element of $\mathcal{W}_-$ is primitive with respect to the coproduct Δ^-. Indeed from the Remark 2.6.3, we deduce that the coaction Δ^- maps every $\tau \in \mathcal{W}_-$ into $\tau \otimes \mathbf{1} + \sum_{(\tau)} \tau' \otimes \tau''$ such that τ'' is a tree of positive degree. However, Δ^- as coproduct on $\mathcal{T}_-$ (see (2.25)), will annihilate any term with τ'' of (strictly) positive degree. In particular then, $\Delta^-\tau = 1 \otimes \tau + \tau \otimes 1$ for all $\tau \in \mathcal{W}_-$, that is, any such τ is primitive. □

Chapter 3

Signatures of paths transformed by polynomial maps

This chapter is based on the article [CP20]. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Beiträge zur Algebra und Geometrie / Contributions to Algebra and Geometry, Signatures of paths transformed by polynomial maps, Laura Colmenarejo and Rosa Preiß, © The Managing Editors 2020.

3.1 Introduction

In the 1950s, K. T. Chen introduced the *iterated-integral signature* of a piecewise continuously differentiable path, which up to a natural equivalence relation, determines the initial path. In general, the signature of a path can be seen as a multidimensional time series. When the terminal time is fixed, the signature of a path can be seen as tensors and the calculation of the signature becomes a standard problem in data science. In [PSS19], M. Pfeffer, A. Seigal, and B. Sturmfels study the inverse problem: given partial information from a signature, can we recover the path? They consider signature tensors of order three under linear transformations and establish identifiability results and recovery algorithms for piecewise linear paths, polynomial paths, and generic dictionaries.

Coming from stochastic analysis, the signatures are becoming more relevant in other areas, such as algebraic geometry and combinatorics, and we would like to highlight some recent work. For instance, in [DR19], J. Diehl and J. Reizenstein offer a combinatorial approach to the understanding of invariants of multidimensional time series based on their signature. Another reference is [AFS19], in which C. Améndola, P. Friz, and B. Sturmfels look at the varieties of (expected) signatures of tensors for both deterministic and random paths, focusing on piecewise linear paths, polynomial paths and (mixtures of) Brownian motion (with drift). Answering one of their questions, in [Gal19], F. Galuppi looks at signature varieties of deterministic log-linear rough paths, which show surprising analogies with the Veronese variety.

In stochastic analysis, the study of the signatures of paths arises in the theory of rough paths, where [FV10, FH14] are textbook references. Iterated integrals and the non-commutative series that encode them have also arisen in a variety of contexts in geometry and arithmetic, including the work of R. Hain in [Hai02], M. Kapranov in [Kap09], and J. Balakrishnan in [Bal13]. The results we derive in this chapter have the potential for future applications in all of these contexts.

Let us now present our problem and our main two results, Theorems 3.1.1 and 3.1.2.

A piecewise continuously differentiable path X in $\mathbb{R}^d$ is a map defined by d piecewise smooth functions $X^i(t)$ in a parameter $t \in [0, L]$, for $i = 1, \ldots, d$. Its signature stores the collection of all the iterated integrals of the path X, which are of the form

$$\int_0^L \int_0^{r_n} \cdots \int_0^{r_2} dX_{r_1}^{i_1} \ldots dX_{r_n}^{i_n}, \tag{3.1}$$

where $X_r^i := X^i(r)$. The iterated integral (3.1) is a real number and it is associated to the sequence $(i_1, i_2 \ldots, i_n)$, for which the order is relevant. Therefore, we consider the signature $\sigma(X)$ as an element of $\mathrm{T}((\mathbb{R}^d))$, the space of formal power series in words in the alphabet $\{\mathtt{1}, \mathtt{2}, \ldots, \mathtt{d}\}$. This space becomes an algebra with the concatenation product, denoted by the symbol $\bullet$. Its algebraic dual, denoted by $\mathrm{T}(\mathbb{R}^d)$, is the space of non-commutative polynomials in the same set of words. It is a commutative algebra with the shuffle product, which is denoted by $\sqcup\!\sqcup$ and interleaves two words in all order-preserving ways, [Reu93, Section 1.4].

We also consider the following duality paring in $\mathrm{T}((\mathbb{R}^d)) \times \mathrm{T}(\mathbb{R}^d)$:

$$\left\langle \sum_{\mathtt{w} \in \mathcal{W}_d} a_{\mathtt{w}} \mathtt{w}, \mathtt{v} \right\rangle = a_{\mathtt{v}}, \tag{3.2}$$

where $\mathcal{W}_d$ denotes the set of words in the alphabet $\{\mathtt{1}, \ldots, \mathtt{d}\}$, together with the empty word $\mathtt{e}$.

Let X be a piecewise continuously differentiable path in $\mathbb{R}^d$ and $\sigma(X)$ be its signature. Consider a polynomial map p from $\mathbb{R}^d$ to $\mathbb{R}^m$. One can compute the image path $p(X)$ and ask for its signature, $\sigma(p(X))$. Then, the following question comes up:

How are both signatures, $\sigma(X)$ and $\sigma(p(X))$, related?

We approach this question from an algebraic point of view. We consider the dual map $p^* : \mathbb{R}[x_1, \ldots, x_m] \longrightarrow \mathbb{R}[x_1, \ldots, x_d]$, where both sets of variables are commutative. It is natural and common to embed the polynomial ring $\mathbb{R}[x_1, \ldots, x_m]$ into the tensor algebra $(\mathrm{T}(\mathbb{R}^m), \sqcup\!\sqcup)$. For that we identify the variable x_i with the letter $\mathtt{i}$ and we define the embedding, denoted by φ_m (or φ), by sending the monomial $x_{i_1} \cdots x_{i_l}$ to the shuffle product $\mathtt{i}_1 \sqcup\!\sqcup \cdots \sqcup\!\sqcup \mathtt{i}_l$, for $1 \leq i_1, \ldots, i_l \leq m$, and extending by linearity. By construction, this map is a morphism of commutative algebras, and it is injective but not surjective. For instance, for $t \geq 2$, $\varphi(x_1 \cdot x_2) = \mathtt{1} \sqcup\!\sqcup \mathtt{2} = \mathtt{12} + \mathtt{21}$ and there is no other way to obtain the words $\mathtt{12}$ and $\mathtt{21}$ as $\varphi_d(h)$, for any polynomial $h \in \mathbb{R}[x_1, \ldots, x_d]$. Therefore, we cannot find a polynomial in $\mathbb{R}[x_1, \ldots, x_d]$ with image $\mathtt{12}$.

Our first step is to define a map $M_p : (\mathrm{T}(\mathbb{R}^m), \sqcup\!\sqcup) \longrightarrow \big(\mathrm{T}(\mathbb{R}^d), \sqcup\!\sqcup\big)$, which is an algebra homomorphism and whose restriction to shuffles of letters is unique in the following sense.

Theorem 3.1.1. *There exists an algebra homomorphism M_p from $(\mathrm{T}(\mathbb{R}^m), \sqcup\!\sqcup)$ to $\big(\mathrm{T}(\mathbb{R}^d), \sqcup\!\sqcup\big)$ such that its restriction $M_p|_{Im(\varphi_m)}$ is the unique algebra homomorphism that makes the following* *diagram commute:*

$$
\begin{array}{ccc}
\mathbb{R}[x_1,\dots,x_m] & \xrightarrow{\quad p^* \quad} & \mathbb{R}[x_1,\dots,x_d] \\
\varphi_m \big\downarrow & & \big\downarrow \varphi_d \\
Im(\varphi_m) & \overset{M_p|_{Im(\varphi_m)}}{\dashrightarrow} & Im(\varphi_d) \\
\cap & & \cap \\
(\mathrm{T}(\mathbb{R}^m), \sqcup\!\sqcup) & & \big(\mathrm{T}(\mathbb{R}^d), \sqcup\!\sqcup\big)
\end{array}
\tag{3.3}
$$

The map M_p has some further interesting properties and is the key to relate the signature of a path with the signature of its transformation under a polynomial map.

Theorem 3.1.2. *Let $X : [0, L] \longrightarrow \mathbb{R}^d$ be a piecewise continuously differentiable path with $X(0) = 0$ and let $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a polynomial map with $p(0) = 0$. Then, for all $\mathtt{w} \in \mathrm{T}(\mathbb{R}^m)$,*

$$\langle \sigma(p(X)), \mathtt{w} \rangle = \langle \sigma(X), M_p(\mathtt{w}) \rangle .$$

Equivalently, $\sigma(p(X)) = M_p^* \left(\sigma(X)\right)$.

The chapter is organized as follows. In Section 3.2, we introduce briefly the framework of signatures of paths, as well as the basic notions that we need for our key combinatorial object, the shuffle algebra on words. We also include some known results such as the *shuffle identity* and *Chen's identity*, and one example of how to compute some coefficients in the signature of a particular path. In Section 3.3, we define the map M_p and prove several of its properties in Proposition 3.3.2. Moreover, we present the proof of our main theorems, Theorems 3.1.1 and 3.1.2. We also include a generalization of the last one for those paths that do not start at the origin, Corollary 3.3.5. In Section 3.3.1 we look at M_p as a half-shuffle homomorphism and give a generalization of Theorem 3.1.2 in terms of Zinbiel algebras. We furthermore point out how our work relates to a half-shuffle identity for signatures of paths, equation (3.8), which was already mentioned in earlier literature, e.g. [GK08, right after Equation (6)]. In Section 3.4, we also present two examples, a very particular one illustrating Theorem 3.1.2 and a more generic one showing how we can store $\sigma\left(p(X)\right)$ using a matrix that encodes the coefficients in $M_p(\mathtt{w})$. Moreover, we present some consequences where we look at polynomial paths (see Corollary 3.4.4), paths that lie in a given variety (see Corollary 3.4.6), and the product of signatures of two arbitrary paths (see Corollary 3.4.7). Finally, Section 3.5 is dedicated to applications and future work.

Acknowledgements

The authors would like to thank Bernd Sturmfels, Francesco Galuppi, Joscha Diehl, Mateusz Michałek, and Max Pfeffer for their fruitful conversations and observations. The collaboration between the co-authors would not have been possible without the financial support from the research institute MPI-MiS Leipzig (Germany). R.P. is currently supported by European Research Council through CoG-683164.

3.2 Signatures of paths and words

Given a piecewise continuously differentiable path $X : [0, L] \longrightarrow \mathbb{R}^d$, for any $i_1, \ldots, i_n \in \{1, 2, \ldots, d\}$ the following integral is classically well-defined

$$\int_0^L dX^{i_1} \ldots dX^{i_n} := \int_0^L \int_0^{r_n} \cdots \int_0^{r_2} dX^{i_1}_{r_1} \ldots dX^{i_n}_{r_n} = \int_0^L \int_0^{r_n} \cdots \int_0^{r_2} \dot{X}^{i_1}_{r_1} \ldots \dot{X}^{i_n}_{r_n} dr_1 \ldots dr_n.$$

We would like to store the collection of all these integrals.

Definition 3.2.1. *The* signature of X *is defined as the following formal power series*

$$\sigma(X) = \sum_{n \geq 0} \sum_{i_1 \ldots i_n} \underbrace{\int_0^L \int_0^{r_n} \cdots \int_0^{r_2} dX^{i_1}_{r_1} \ldots dX^{i_n}_{r_n}}_{\in \mathbb{R}} \cdot \mathtt{i}_1 \cdots \mathtt{i}_n \in \mathrm{T}((\mathbb{R}^d)).$$

As we mention in the introduction, $\mathrm{T}((\mathbb{R}^d))$ is the space of formal power series in words in the alphabet $\{\mathtt{1}, \ldots, \mathtt{d}\}$, and we denote by $\mathtt{e}$ the empty word. It is an algebra with the concatenation product, denoted by $\mathtt{w} \bullet \mathtt{v}$ (or simply $\mathtt{wv}$), which is well-defined since it respects the grading given by the number of letters appearing in each word. We also consider its algebraic dual $\mathrm{T}(\mathbb{R}^d)$, which is the set of polynomials in words in the same alphabet. The algebra $\mathrm{T}(\mathbb{R}^d)$ has the concatenation product, which is the same as for $\mathrm{T}((\mathbb{R}^d))$ if we multiply two finite power series. However, we consider $\mathrm{T}(\mathbb{R}^d)$ as an algebra with the shuffle product, which we define recursively as follows.

Definition 3.2.2. *Let* $\mathtt{w}$, $\mathtt{w}_1$ *and* $\mathtt{w}_2$ *be three words and* $\mathtt{a}$ *and* $\mathtt{b}$ *two letters. We define the* shuffle product *of two words recursively by*

$$\mathtt{e} \sqcup\!\sqcup \mathtt{w} = \mathtt{w} \sqcup\!\sqcup \mathtt{e} = \mathtt{w}, \quad \textit{and}$$
$$(\mathtt{w}_1 \bullet \mathtt{a}) \sqcup\!\sqcup (\mathtt{w}_2 \bullet \mathtt{b}) = (\mathtt{w}_1 \sqcup\!\sqcup (\mathtt{w}_2 \bullet \mathtt{b})) \bullet \mathtt{a} + ((\mathtt{w}_1 \bullet \mathtt{a}) \sqcup\!\sqcup \mathtt{w}_2) \bullet \mathtt{b}.$$

Note that in the shuffle product, we distinguish duplicated letters. For instance, for a letter $\mathtt{a}$, we have $\mathtt{a} \sqcup\!\sqcup \mathtt{a} = 2 \cdot \mathtt{aa}$. Notice that the concatenation is a non-commutative operation, whereas the shuffle product is commutative.

We also need a few notions on words. The *length of a word* $\mathtt{w}$ is denoted by $\ell(\mathtt{w})$ and counts the number of letters in $\mathtt{w}$. We also define $\ell\left(\sum_i \alpha_i \mathtt{w}_i\right) := \max_i\{\ell(\mathtt{w}_i)\}$, for any linear combination of words $\mathtt{w}_i$. Therefore, $\ell(\mathtt{w}_1 \bullet \mathtt{w}_2) = \ell(\mathtt{w}_1) + \ell(\mathtt{w}_2) = \ell(\mathtt{w}_1 \sqcup\!\sqcup \mathtt{w}_2)$. As an $\mathbb{R}$-vector space, $\mathrm{T}(\mathbb{R}^d)$ is graded by the length of the words:

$$\mathrm{T}(\mathbb{R}^d) = \bigoplus_{n \geq 0} \mathrm{T}^n(\mathbb{R}^d),$$

where $\mathrm{T}^n(\mathbb{R}^d)$ is the vector space spanned by the words of length n. We also denote by $\mathrm{T}^{\leq n}(\mathbb{R}^d)$ the partial direct sum $\bigoplus_{k \leq n} \mathrm{T}^k(\mathbb{R}^d)$. This notation extends to $\mathrm{T}((\mathbb{R}^d))$. In the same way, $\sigma^{(n)}(X)$ denotes the partial sum of $\sigma(X)$ for which all the appearing words have length exactly n. We are ready to prove the following result.

Proposition 3.2.3. *The shuffle product is associative.*

Proof. The associativity is clear for the empty word since $(\mathtt{e} \sqcup\!\sqcup \mathtt{e}) \sqcup\!\sqcup \mathtt{e} = \mathtt{e} = \mathtt{e} \sqcup\!\sqcup (\mathtt{e} \sqcup\!\sqcup \mathtt{e})$. Now, we proceed by induction. Assume that for any words $\mathtt{w}_1$, $\mathtt{v}_1$, and $\mathtt{u}_1$ such that $\ell(\mathtt{w}_1) + \ell(\mathtt{v}_1) + \ell(\mathtt{u}_1) = n$, for some $n \in \mathbb{N}_0$, we have that $(\mathtt{w}_1 \sqcup\!\sqcup \mathtt{v}_1) \sqcup\!\sqcup \mathtt{u}_1 = \mathtt{w}_1 \sqcup\!\sqcup (\mathtt{v}_1 \sqcup\!\sqcup \mathtt{u}_1)$. This is our inductive hypothesis.

Let $\mathtt{w}_2, \mathtt{v}_2, \mathtt{u}_2$ be arbitrary words with the property that $\ell(\mathtt{w}_2) + \ell(\mathtt{v}_2) + \ell(\mathtt{u}_2) = n + 1$. At least one of those words must thus be non-empty. If exactly two of the words are empty, both $(\mathtt{w}_2 \sqcup\!\sqcup \mathtt{v}_2) \sqcup\!\sqcup \mathtt{u}_2$ and $\mathtt{w}_2 \sqcup\!\sqcup (\mathtt{v}_2 \sqcup\!\sqcup \mathtt{u}_2)$ are obviously equal to the non-empty word. If exactly one of the words is empty, both $(\mathtt{w}_2 \sqcup\!\sqcup \mathtt{v}_2) \sqcup\!\sqcup \mathtt{u}_2$ and $\mathtt{w}_2 \sqcup\!\sqcup (\mathtt{v}_2 \sqcup\!\sqcup \mathtt{u}_2)$ are obviously equal to the shuffle product of the two non-empty words. In the remaining case, if $\mathtt{w}_2, \mathtt{v}_2, \mathtt{u}_2$ are all non-empty, there are words $\mathtt{w}, \mathtt{v}, \mathtt{u}$ and letters $\mathtt{i}, \mathtt{j}, \mathtt{k}$ such that $\mathtt{w}_2 = \mathtt{wi}$, $\mathtt{v}_2 = \mathtt{vj}$ and $\mathtt{u}_2 = \mathtt{uk}$. Then,

$$\begin{aligned}(\mathtt{w}_2 \sqcup\!\sqcup \mathtt{v}_2) \sqcup\!\sqcup \mathtt{u}_2 &= (\mathtt{wi} \sqcup\!\sqcup \mathtt{vj}) \sqcup\!\sqcup \mathtt{uk} = \big((\mathtt{w} \sqcup\!\sqcup \mathtt{vj}) \bullet \mathtt{i} + (\mathtt{wi} \sqcup\!\sqcup \mathtt{v}) \bullet \mathtt{j}\big) \sqcup\!\sqcup \mathtt{uk} \\ &= \big((\mathtt{w} \sqcup\!\sqcup \mathtt{vj}) \sqcup\!\sqcup \mathtt{uk}\big) \bullet \mathtt{i} + \big((\mathtt{wi} \sqcup\!\sqcup \mathtt{v}) \sqcup\!\sqcup \mathtt{uk}\big) \bullet \mathtt{j} + \Big(\big((\mathtt{w} \sqcup\!\sqcup \mathtt{vj})\mathtt{i} + (\mathtt{wi} \sqcup\!\sqcup \mathtt{v})\mathtt{j}\big) \sqcup\!\sqcup \mathtt{u}\Big) \bullet \mathtt{k} \\ &= \big((\mathtt{w} \sqcup\!\sqcup \mathtt{vj}) \sqcup\!\sqcup \mathtt{uk}\big) \bullet \mathtt{i} + \big((\mathtt{wi} \sqcup\!\sqcup \mathtt{v}) \sqcup\!\sqcup \mathtt{uk}\big) \bullet \mathtt{j} + \big((\mathtt{wi} \sqcup\!\sqcup \mathtt{vj}) \sqcup\!\sqcup \mathtt{u}\big) \bullet \mathtt{k}\end{aligned}$$

Analogously,

$$\mathtt{w}_2 \sqcup\!\sqcup (\mathtt{v}_2 \sqcup\!\sqcup \mathtt{u}_2) = \big(\mathtt{w} \sqcup\!\sqcup (\mathtt{vj} \sqcup\!\sqcup \mathtt{uk})\big) \bullet \mathtt{i} + \big(\mathtt{wi} \sqcup\!\sqcup (\mathtt{v} \sqcup\!\sqcup \mathtt{uk})\big) \bullet \mathtt{j} + \big(\mathtt{wi} \sqcup\!\sqcup (\mathtt{vj} \sqcup\!\sqcup \mathtt{u})\big) \bullet \mathtt{k}.$$

Thus, since

$$\ell(\mathtt{w}) + \ell(\mathtt{vj}) + \ell(\mathtt{uk}) = \ell(\mathtt{wi}) + \ell(\mathtt{v}) + \ell(\mathtt{uk}) = \ell(\mathtt{wi}) + \ell(\mathtt{vj}) + \ell(\mathtt{u}) = n,$$

we again get $(\mathtt{w}_2 \sqcup\!\sqcup \mathtt{v}_2) \sqcup\!\sqcup \mathtt{u}_2 = \mathtt{w}_2 \sqcup\!\sqcup (\mathtt{v}_2 \sqcup\!\sqcup \mathtt{u}_2)$ due to the induction hypothesis. □

Going back to the signatures, the dual pairing (3.2) in $\mathrm{T}((\mathbb{R}^d)) \times \mathrm{T}(\mathbb{R}^d)$ allows us to extract the coefficient of a word in the signature of a path in the following way:

$$\langle \sigma(X), \mathtt{i}_1 \mathtt{i}_2 \dots \mathtt{i}_n \rangle = \int_0^L dX^{i_1} \dots dX^{i_n}.$$

Both operations, the concatenation and the shuffle products, behave nicely with respect to the signature, as the following two known results describe. The first result, known as the *shuffle identity*, relates the signature of a path with the shuffle product.

Proposition 3.2.4 (Shuffle identity, [Ree58, Equation (2.5.2)])**.** *Let $X : [0, L] \longrightarrow \mathbb{R}^d$ be a piecewise continuously differentiable path. Then, for every $\mathtt{u}, \mathtt{v} \in \mathrm{T}(\mathbb{R}^d)$,*

$$\langle \sigma(X), \mathtt{u} \rangle \langle \sigma(X), \mathtt{v} \rangle = \langle \sigma(X), \mathtt{u} \sqcup\!\sqcup \mathtt{v} \rangle .$$

Another important result, known as *Chen's relation*, describes the signature when we concatenate paths. Let us see how the concatenation path is defined.

Definition 3.2.5. *Let $X, Y : [0, L] \longrightarrow \mathbb{R}^d$ be two piecewise continuously differentiable paths. We define the* concatenation *of X and Y as the path $X \sqcup Y : [0, 2L] \longrightarrow \mathbb{R}^d$ given by X on $[0, L]$ and by $Y_{\cdot - L} - Y_0 + X_L$ on $[L, 2L]$ (i.e. take Y, move it back to 0 and then move it to the end of X).*

The concatenation product interplays nicely with the concatenation of paths, as the following proposition shows.

Proposition 3.2.6 (Chen's identity, [Che54, Theorem 3.1][Che57, Theorem 2.1])**.** *Let $X, Y : [0, L] \longrightarrow \mathbb{R}^d$ be two piecewise continuously differentiable paths and consider their concatenation $X \sqcup Y : [0, 2L] \longrightarrow \mathbb{R}^d$. Then, the signature of the concatenation is the concatenation product of the signatures:*

$$\sigma(X \sqcup Y) = \sigma(X) \bullet \sigma(Y).$$

We finish this section with an example on how to compute the first terms of the signature of a path.

Example 3.2.7. *Consider the path $X : [0,1] \longrightarrow \mathbb{R}^2$ given by $X^1(t) = t$ and $X^2(t) = t^2$. We compute a few terms of its signature.*

$$\langle\sigma(X), \mathtt{1}\rangle = \int_0^1 dX^1_{r_1} = \int_0^1 1dt = 1 \qquad \langle\sigma(X), \mathtt{2}\rangle = \int_0^1 dX^2_{r_1} = \int_0^1 2tdt = 1$$

$$\langle\sigma(X), \mathtt{11}\rangle = \int_0^1 \int_0^{r_2} dX^1_{r_1} dX^1_{r_2} = \int_0^1 r_2 dX^1_{r_2} = \int_0^1 r_2 dr_2 = \frac{1}{2}$$

$$\langle\sigma(X), \mathtt{12}\rangle = \int_0^1 \int_0^{r_2} dX^1_{r_1} dX^2_{r_2} = \int_0^1 r_2 dX^2_{r_2} = \int_0^1 2r_2^2 dr_2 = \frac{2}{3}$$

$$\langle\sigma(X), \mathtt{21}\rangle = \int_0^1 \int_0^{r_2} dX^2_{r_1} dX^1_{r_2} = \int_0^1 r_2^2 dX^1_{r_2} = \int_0^1 r_2^2 dr_2 = \frac{1}{3}$$

$$\langle\sigma(X), \mathtt{22}\rangle = \int_0^1 \int_0^{r_2} dX^2_{r_1} dX^2_{r_2} = \int_0^1 r_2^2 dX^2_{r_2} = \int_0^1 2r_2^3 dr_2 = \frac{2}{4} = \frac{1}{2}$$

$$\langle\sigma(X), \mathtt{222}\rangle = \int_0^1 \int_0^{r_3} \int_0^{r_2} dX^2_{r_1} dX^2_{r_2} dX^2_{r_3} = \int_0^1 \int_0^{r_3} r_2^2 dX^2_{r_2} dX^2_{r_3} = \int_0^1 \int_0^{r_3} 2r_2^3 dr_2 dX^2_{r_3} = \int_0^1 \frac{r_3^4}{2} dX^2_{r_3} = \int_0^1 r_3^5 dr_3 = \left.\frac{r_3^6}{6}\right|_0^1 = \frac{1}{6}$$

Therefore, the signature of X is of the form

$$\sigma(X) = \mathtt{1} + \mathtt{2} + \frac{1}{2} \cdot (\mathtt{11} + \mathtt{22}) + \frac{1}{3} \cdot (2 \cdot \mathtt{12} + \mathtt{21}) + \frac{1}{6} \cdot \mathtt{222} + \dots$$

3.3 Signatures under the action of polynomial maps

Let $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a polynomial map given by the polynomials $p_i(x_1, x_2, \dots, x_d)$, for $i = 1, 2, \dots, m$, with the property that $p(0) = 0$. The *degree of the polynomial map* p is the maximum of the degree of the polynomials that define it, $\deg(p) = \max_i \deg(p_i)$. Moreover, we say that a polynomial map is *homogeneous* if the polynomials p_i are homogeneous of the same degree. Finally, we denote by J_p the Jacobian matrix of format $m \times d$ with entries $J_p^{ij} = \partial_j p_i$, for $i \in \{1, 2, \dots, m\}$ and $j \in \{1, 2, \dots, d\}$.

Recall the algebra homomorphism φ_d defined by:

$$\begin{array}{rrcl} \varphi_d : & \mathbb{R}[x_1, x_2, \dots, x_d] & \longrightarrow & (\mathrm{T}(\mathbb{R}^d), ⧢) \\ & x_i & \longmapsto & \mathtt{i} \\ & x_{i_1} \cdots x_{i_l} & \longmapsto & \mathtt{i}_1 ⧢ \cdots ⧢ \mathtt{i}_l \end{array}$$

As we mention above, by the properties of the shuffle product in $\mathrm{T}(\mathbb{R}^d)$, this map is injective but is not surjective.

We now consider maps M_p, φ, and p^* which complete the diagram (3.3) that arises in our main question in the following way:

$$\begin{array}{ccc} \mathbb{R}[x_1, \dots, x_m] & \xrightarrow{\;p^*\;} & \mathbb{R}[x_1, \dots, x_d] \\ \downarrow{\scriptstyle \varphi_m} & & \downarrow{\scriptstyle \varphi_d} \\ (\mathrm{T}(\mathbb{R}^m), ⧢) & \overset{\exists M_p}{\dashrightarrow} & (\mathrm{T}(\mathbb{R}^d), ⧢) \end{array}$$

Notice that the map M_p is unique when restricted to the image of φ_m, but not the full M_p on the tensor algebra $\mathrm{T}(\mathbb{R}^m)$.

Definition 3.3.1. *For any polynomial map $p: \mathbb{R}^d \to \mathbb{R}^m$ such that $p(0) = 0$, let $k_p^{ij} = \varphi_d\left(J_p^{ij}\right) \in \mathrm{T}(\mathbb{R}^d)$, where J_p^{ij} is the (i,j)-entry of the Jacobian matrix of p. We define the map $M_p : \mathrm{T}(\mathbb{R}^m) \longrightarrow \mathrm{T}(\mathbb{R}^d)$ recursively as follows:*

$$M_p(\mathtt{e}) = \mathtt{e}, \text{ for } \mathtt{e} \text{ the empty word, and}$$
$$M_p(\mathtt{wi}) = \sum_{j=1}^{d} \left(M_p(\mathtt{w}) \sqcup\!\sqcup\, k_p^{ij}\right) \bullet \mathtt{j}, \text{ for any word } \mathtt{w} \text{ and any letter } \mathtt{i} \in \{\mathtt{1}\ldots,\mathtt{m}\}.$$

Note: Do not confuse M_p with the map M_v from Definition 2.3.8, they are different in nature and not even defined on the same space!

The following result summarizes a few properties of the map M_p. We will use these properties to show that the map that M_p as we construct it restricts according to what we need.

Proposition 3.3.2. *Consider two polynomial maps $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ and $q : \mathbb{R}^m \longrightarrow \mathbb{R}^s$, with $p(0) = 0$ and $q(0) = 0$, and the algebra homomorphisms φ_m and φ_d. Then, we have the following list of properties:*

1. *$M_p : (\mathrm{T}(\mathbb{R}^m), \sqcup\!\sqcup) \longrightarrow (\mathrm{T}(\mathbb{R}^d), \sqcup\!\sqcup)$ is an algebra homomorphism.*
2. *For $i = 1, \ldots, s$, $M_p\left(\varphi_m(q_i)\right) = \varphi_d(q_i \circ p)$, where the q_i's are the polynomials defining the polynomial map q.*
3. *The coefficients in Definition 3.3.1 satisfy that $k_{q\circ p}^{ij} = \displaystyle\sum_{l=1}^{m} M_p(k_q^{il}) \sqcup\!\sqcup\, k_p^{lj}$.*
4. *The map M_p behaves well with respect to composition of polynomial maps. That is, $M_{q\circ p} = M_p M_q$.*
5. *If p is a polynomial map of degree n, then $M_p\left(\mathrm{T}^k(\mathbb{R}^m)\right) \subseteq \mathrm{T}^{\leq nk}(\mathbb{R}^d)$.*
6. *If p is an homogeneous polynomial map of degree n, $M_p\left(\mathrm{T}^k(\mathbb{R}^m)\right) \subseteq \mathrm{T}^{nk}(\mathbb{R}^d)$*

Proof. 1. We need to show that, for any words $\mathtt{w}_1$ and $\mathtt{w}_2$ in $\mathrm{T}(\mathbb{R}^m)$,

$$M_p(\mathtt{w}_1 \sqcup\!\sqcup\, \mathtt{w}_2) = M_p(\mathtt{w}_1) \sqcup\!\sqcup\, M_p(\mathtt{w}_2).$$

We proceed by induction on $\ell(\mathtt{w}_1) + \ell(\mathtt{w}_2)$. For $\ell(\mathtt{w}_1) + \ell(\mathtt{w}_2) \leq 1$, at least one of the two words is the empty word $\mathtt{e}$, and so we assume that $\mathtt{w}_2 = \mathtt{e}$. Therefore,

$$M_p(\mathtt{w}_1 \sqcup\!\sqcup\, \mathtt{w}_2) = M_p(\mathtt{w}_1 \sqcup\!\sqcup\, \mathtt{e}) = M_p(\mathtt{w}_1) = M_p(\mathtt{w}_1) \sqcup\!\sqcup\, \mathtt{e} = M_p(\mathtt{w}_1) \sqcup\!\sqcup\, M_p(\mathtt{w}_2).$$

Assume now that the statement is true for any pair of words with sum of lengths at most $n-1$. Let $\mathtt{u}$ and $\mathtt{v}$ be two words such that $\ell(\mathtt{u}) + \ell(\mathtt{v}) = n - 1$ and $\mathtt{a}$ and $\mathtt{b}$ two arbitrary letters. Then, using Definitions 3.2.2 and 3.3.1, and the inductive hypothesis (IH),

$$M_p(\mathtt{ua} \sqcup\!\sqcup \mathtt{vb}) \overset{\text{Def.}}{\underset{3.2.2}{=}} M_p\left((\mathtt{u} \sqcup\!\sqcup \mathtt{vb}) \bullet \mathtt{a} + (\mathtt{ua} \sqcup\!\sqcup \mathtt{v}) \bullet \mathtt{b}\right) \overset{\text{Def.}}{\underset{3.3.1}{=}}$$
$$\sum_{i=1}^{d} \left[M_p(\mathtt{u} \sqcup\!\sqcup \mathtt{vb}) \sqcup\!\sqcup k_p^{ai} + M_p(\mathtt{ua} \sqcup\!\sqcup \mathtt{v}) \sqcup\!\sqcup k_p^{bi}\right] \bullet \mathtt{i} \overset{\text{Def. 3.3.1}}{\underset{(IH)}{=}}$$
$$\sum_{i=1}^{d}\sum_{j=1}^{d} \left[M_p(\mathtt{u}) \sqcup\!\sqcup \left((M_p(\mathtt{v}) \sqcup\!\sqcup k_p^{bj}) \bullet \mathtt{j}\right) \sqcup\!\sqcup k_p^{ai}\right] \bullet \mathtt{i} +$$
$$\sum_{i=1}^{d}\sum_{j=1}^{d} \left[\left((M_p(\mathtt{u}) \sqcup\!\sqcup k_p^{aj}) \bullet \mathtt{j}\right) \sqcup\!\sqcup M_p(\mathtt{v}) \sqcup\!\sqcup k_p^{bi}\right] \bullet \mathtt{i} \overset{\text{Def.}}{\underset{3.2.2}{=}}$$
$$\left(\sum_{i=1}^{d} (M_p(\mathtt{u}) \sqcup\!\sqcup k_p^{ai}) \bullet \mathtt{i}\right) \sqcup\!\sqcup \left(\sum_{i=1}^{d} (M_p(\mathtt{v}) \sqcup\!\sqcup k_p^{bi}) \bullet \mathtt{i}\right) \overset{\text{Def.}}{\underset{3.3.1}{=}}$$
$$M_p(\mathtt{ua}) \sqcup\!\sqcup M_p(\mathtt{vb}).$$

2. Assume that $M_p(\mathtt{i}) = \varphi_d(p_i)$, for all i. Then, for the monomial $h(x_1, \ldots, x_m) = x_1^{n_1} \cdot x_2^{n_2} \cdots x_m^{n_m}$, $\varphi_m(h) = \mathtt{1}^{\sqcup\!\sqcup n_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \mathtt{m}^{\sqcup\!\sqcup n_m}$. Therefore, by the property (1),

$$\begin{aligned} M_p(\varphi_m(h)) &= M_p(\mathtt{1})^{\sqcup\!\sqcup n_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup M_p(\mathtt{m})^{\sqcup\!\sqcup n_m} = \varphi_d(p_1)^{\sqcup\!\sqcup n_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \varphi_d(p_m)^{\sqcup\!\sqcup n_m} \\ &= \varphi_d(p_1^{n_1} \cdots p_m^{n_m}) = \varphi_d(h \circ p), \end{aligned}$$

and the property (2) follows by linearity.

We prove now the claim $M_p(\mathtt{i}) = \varphi_d(p_i)$, for all i. Since the two maps $p \longmapsto M_p(\mathtt{i})$ and $p \longmapsto \varphi_d(p_i)$ are linear, it is enough to prove the claim for the case when p_i is a monomial of the form $p_i = x_1^{n_1} \cdot x_2^{n_2} \cdots x_d^{n_d}$, with at least one of the n_i's non-zero. In this case,

$$\begin{aligned} M_p(\mathtt{i}) &= \sum_{j=1}^{d} \left(M_p(\mathtt{e}) \sqcup\!\sqcup k_p^{ij}\right) \bullet \mathtt{j} = \sum_{j=1}^{d} k_p^{ij} \bullet \mathtt{j} \\ &= \sum_{\substack{j=1 \\ n_j \neq 0}}^{d} n_j \left(\mathtt{1}^{\sqcup\!\sqcup (n_1 - \delta_{1j})} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \mathtt{d}^{\sqcup\!\sqcup (n_d - \delta_{dj})}\right) \bullet \mathtt{j} = \mathtt{1}^{\sqcup\!\sqcup n_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \mathtt{d}^{\sqcup\!\sqcup n_d} = \varphi_d(p_i), \end{aligned}$$

where $*$ follows by applying enough iterations of the recursive definition of the shuffle product and the fact that $\mathtt{i}^{\sqcup\!\sqcup n+1} = (n+1)\mathtt{i}^{\sqcup\!\sqcup n} \bullet \mathtt{i}$.

3. By the *chain rule*, $J_{q \circ p}^{ij} = \sum_{l=1}^{m} \left(J_q^{il} \circ p\right) \cdot J_p^{lj}$. By (2), we have that $\varphi_d(J_q^{il} \circ p) = M_p\left(\varphi_m(J_q^{il})\right) = M_p(k_q^{il})$. Thus,

$$k_{q \circ p}^{ij} = \varphi_d\left(J_{q \circ p}^{ij}\right) = \varphi_d\left(\sum_{l=1}^{m} \left(J_q^{il} \circ p\right) \cdot J_p^{lj}\right) = \sum_{l=1}^{m} \varphi_d\left(J_q^{il} \circ p\right) \sqcup\!\sqcup \varphi_d(J_p^{lj}) = \sum_{l=1}^{m} M_p(k_q^{il}) \sqcup\!\sqcup k_p^{lj}.$$

4. We proceed by induction on $\ell(\mathtt{w})$. For a letter $\mathtt{i}$,

$$M_p \circ M_q(\mathtt{i}) = M_p(M_q(\mathtt{i})) = M_p\left(\sum_{j=1}^{m} k_q^{ij} \bullet \mathtt{j}\right) = \sum_{j=1}^{m} M_p\left(k_q^{ij} \bullet \mathtt{j}\right) =$$
$$\sum_{j=1}^{m}\left(\sum_{l=1}^{d} M_p(k_q^{ij} \sqcup\!\sqcup k_p^{jl}\right) \bullet \mathtt{l} = \sum_{l=1}^{d}\left(\sum_{j=1}^{m} M_p(k_q^{ij} \sqcup\!\sqcup k_p^{jl}\right) \bullet \mathtt{l} \overset{(3)}{=} \sum_{l=1}^{d} k_{q\circ p}^{il} \bullet \mathtt{l} = M_{q\circ p}(\mathtt{i}).$$

Now, we assume that the statement is true for all the words of length at most n and we refer to this as (IH). Let $\mathtt{w}$ be one of these words and $\mathtt{i}$ any letter. Then,

$$M_p \circ M_q(\mathtt{wi}) = M_p\left(M_q(\mathtt{wi})\right) = M_p\left(\left(\sum_{j=1}^{m} M_q(\mathtt{w}) \sqcup\!\sqcup k_q^{ij}\right) \bullet \mathtt{j}\right) =$$
$$\sum_{j=1}^{m} M_p\left((M_q(\mathtt{w}) \sqcup\!\sqcup k_q^{ij}) \bullet \mathtt{j}\right) = \sum_{j=1}^{m}\sum_{l=1}^{d}\left[M_p\left(M_q(\mathtt{w}) \sqcup\!\sqcup k_q^{ij}\right) \sqcup\!\sqcup k_p^{jl}\right] \bullet \mathtt{l} \overset{(1)}{=}$$
$$\sum_{j=1}^{m}\sum_{l=1}^{d}\left[M_p(M_q(\mathtt{w})) \sqcup\!\sqcup M_p(k_q^{ij}) \sqcup\!\sqcup k_p^{jl}\right] \bullet \mathtt{l} \overset{(\mathrm{IH})}{=} \sum_{j=1}^{m}\sum_{l=1}^{d}\left[M_{q\circ p}(\mathtt{w})) \sqcup\!\sqcup M_p(k_q^{ij}) \sqcup\!\sqcup k_p^{jl}\right] \bullet \mathtt{l} =$$
$$\sum_{l=1}^{d}\left[M_{q\circ p}(\mathtt{w})) \sqcup\!\sqcup \left(\sum_{j=1}^{m} M_p(k_q^{ij}) \sqcup\!\sqcup k_p^{jl}\right)\right] \bullet \mathtt{l} \overset{(3)}{=} \sum_{l=1}^{d}\left[M_{q\circ p}(\mathtt{w}) \sqcup\!\sqcup k_{q\circ p}^{il}\right] \bullet \mathtt{l} = M_{q\circ p}(\mathtt{wi}).$$

5. We start by noticing that since the polynomial map p has degree n, then $\deg(p_i) \leq n$, for all i. Thus, $\deg(J_p^{ij}) \leq n-1$ and $\ell\left(\varphi_d(J_p^{ij})\right) \leq n-1$, for all i and j.

   Now, we proceed by induction on k. For $k = 1$, $\mathrm{T}^1(\mathbb{R}^m)$ is the set of letters $\{\mathtt{1}, \ldots, \mathtt{m}\}$. Since for a letter $\mathtt{i}$ in this set $M_p(\mathtt{i}) = \sum_{j=1}^{d} k_p^{ij} \bullet \mathtt{j}$, then $\ell(M_p(\mathtt{i})) = \max_j \left\{\ell(k_p^{ij} \bullet \mathtt{j})\right\} \leq n$. Thus, $M_p(\mathtt{i}) \in \mathrm{T}^{\leq n}(\mathbb{R}^d)$.

   Assume that the statement is true for k. Any word $\mathtt{w}' \in \mathrm{T}^{k+1}(\mathbb{R}^m)$ can be written as $\mathtt{w}' = \mathtt{w} \bullet \mathtt{i}$, with $\mathtt{w} \in \mathrm{T}^k(\mathbb{R}^m)$ and $\mathtt{i}$ a letter. We analyze the length of $M_p(\mathtt{w} \bullet \mathtt{i})$. Since $M_p(\mathtt{w} \bullet \mathtt{i}) = \sum_{j=1}^{d}\left(M_p(\mathtt{w}) \sqcup\!\sqcup k_p^{ij}\right) \bullet \mathtt{j}$, it is enough to upper bound the length of the terms appearing in the sum. By the inductive hypothesis, $\ell(M_p(\mathtt{w})) \leq nk$, and since $\ell(k_p^{ij}) \leq n-1$, $\ell(M_p(\mathtt{w}) \sqcup\!\sqcup k_p^{ij}) \leq nk + n - 1$. Therefore, $\ell(M_p(\mathtt{w} \bullet \mathtt{i})) \leq nk + n - 1 + 1 = n(k+1)$.

6. In this case, since p is homogeneous of degree n, then $\deg(p_i) = n$, for all i. Moreover, $\deg(J_p^{ij}) = n-1$, if the variable x_j appears in p_i, or zero, otherwise.

   We proceed by induction on k. For $k = 1$, let $\mathtt{i}$ be a letter in $\{\mathtt{1}, \ldots, \mathtt{m}\}$. Then, $M_p(\mathtt{i}) = \sum_{j=1}^{d} k_p^{ij} \bullet \mathtt{j}$. This sum contains only terms $k_p^{ij} \bullet \mathtt{j}$, which has length exactly n, otherwise k_p^{ij} is zero according to our observation about the Jacobian entries above. Therefore, the statement follows.

Now, assume the statement is true for k. Let $\mathtt{w} \in \mathrm{T}^k(\mathbb{R}^m)$ be a word and $\mathtt{i}$ a letter. In this case, $M_p(\mathtt{w} \bullet \mathtt{i}) = \sum_{j=1}^{d} \left(M_p(\mathtt{w}) \sqcup\!\sqcup\, k_p^{ij}\right) \bullet \mathtt{j}$. Again, the terms appearing in this sum have length $nk + n - 1 + 1 = n(k+1)$, which concludes the proof.

□

Once we have these properties, we recall Theorem 3.1.1 and prove it.

Theorem 3.3.3. (Theorem 3.1.1) *There exists an algebra homomorphism* $M_p : (\mathrm{T}(\mathbb{R}^m), \sqcup\!\sqcup) \longrightarrow \left(\mathrm{T}(\mathbb{R}^d), \sqcup\!\sqcup\right)$ *such that its restriction* $M_p|_{Im(\varphi_m)}$ *is the unique algebra homomorphism that makes the following diagram commute:*

$$
\begin{array}{ccc}
\mathbb{R}[x_1, \dots, x_m] & \xrightarrow{\quad p^* \quad} & \mathbb{R}[x_1, \dots, x_d] \\
\Big\downarrow {\scriptstyle \varphi_m} & & \Big\downarrow {\scriptstyle \varphi_d} \\
Im(\varphi_m) & \xrightarrow{\exists!\, M_p|_{Im(\varphi_m)}} & Im(\varphi_d)
\end{array}
$$

Proof. By (1) in Proposition 3.3.2, the restriction of M_p to the image $Im(\varphi_m)$ is an algebra homomorphism. Moreover, due to (2), we have that $M_p\left(Im(\varphi_m)\right) \subseteq Im(\varphi_d)$. Since we restrict to their images, φ_m and φ_d are isomorphisms and the map $M_p|_{Im(\varphi_m)}$ is the unique one making the diagram commute. □

Let us see now the answer to our main question, which is stated as Theorem 3.1.2.

Theorem 3.3.4. (Theorem 3.1.2) *Let $X : [0, L] \longrightarrow \mathbb{R}^d$ be a piecewise continuously differentiable path with $X_0 = 0$ and let $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a polynomial map with $p(0) = 0$. Then, for all $\mathtt{w} \in \mathrm{T}(\mathbb{R}^m)$,*

$$\langle \sigma(p(X)), \mathtt{w} \rangle = \langle \sigma(X), M_p(\mathtt{w}) \rangle .$$

Equivalently, $\sigma(p(X)) = M_p^*\left(\sigma(X)\right)$.

Proof. Let $Y = p(X)$. Then, $\dot{Y}_s = \sum_{j=1}^{d} J_p^{ij}(X_s) \cdot \dot{X}_s^j$, for almost every $s \in [0, L]$. Notice that the entries of the Jacobian matrix can be seen as coefficients of the signature of X,

$$J_p^{ij}(X_t) = \left\langle \sigma\left(X|_{[0,t]} \right), k_p^{ij} \right\rangle . \tag{3.4}$$

We proceed by induction on the length of the word $\mathtt{w}$. For a letter $\mathtt{i}$, we have that

$$
\begin{aligned}
\langle \sigma(Y), \mathtt{i} \rangle = \int_0^L dY_t^i = \int_0^L \dot{Y}_t^i dt = \sum_{j=1}^{d} \int_0^L J_p^{ij}(X_t) \dot{X}_t^j dt = \\
\sum_{j=1}^{d} \int_0^L J_p^{ij}(X_t) dX_t^j \overset{(3.4)}{=} \sum_{j=1}^{d} \int_0^L \left\langle \sigma\left(X|_{[0,t]} \right), k_p^{ij} \right\rangle dX_t^j = \\
\sum_{j=1}^{d} \left\langle \sigma(X), k_p^{ij} \bullet \mathtt{j} \right\rangle = \langle \sigma(X), M_p(\mathtt{i}) \rangle .
\end{aligned}
$$

Now, assume that the statement is true for all the words of length at most n. Let $\mathtt{w}$ be any of these words and $\mathtt{i}$ any letter. By the definition of the signature,

$$\langle \sigma(Y), \mathtt{wi} \rangle = \int_0^L \left\langle \sigma\left(Y|_{[0,t]} \right), \mathtt{w} \right\rangle dY_t^i = \int_0^L \left\langle \sigma\left(Y|_{[0,t]} \right), \mathtt{w} \right\rangle \dot{Y}_t^i dt =$$
$$\sum_{j=1}^d \int_0^L \left\langle \sigma\left(Y|_{[0,t]} \right), \mathtt{w} \right\rangle J_p^{ij}(X_t) \dot{X}_t^j dt = \sum_{j=1}^d \int_0^L \left\langle \sigma\left(Y|_{[0,t]} \right), \mathtt{w} \right\rangle J_p^{ij}(X_t) dX_t^j \overset{(3.4)}{=}$$
$$\sum_{j=1}^d \int_0^L \left\langle \sigma\left(Y|_{[0,t]} \right), \mathtt{w} \right\rangle \left\langle \sigma\left(X|_{[0,t]} \right), k_p^{ij} \right\rangle dX_t^j. \quad (3.5)$$

Now, apply the inductive hypothesis to $\left\langle \sigma\left(Y|_{[0,t]} \right), \mathtt{w} \right\rangle$ in (3.5), and then by the shuffle identity, Proposition 3.2.4, and the definition of the signature,

$$\langle \sigma(Y), \mathtt{wi} \rangle = \sum_{j=1}^d \int_0^L \left\langle \sigma\left(X|_{[0,t]} \right), M_p(\mathtt{w}) \right\rangle \left\langle \sigma\left(X|_{[0,t]} \right), k_p^{ij} \right\rangle dX_t^j =$$
$$\sum_{j=1}^d \int_0^L \left\langle \sigma\left(X|_{[0,t]} \right), M_p(\mathtt{w}) \sqcup\!\sqcup\, k_p^{ij} \right\rangle dX_t^j = \sum_{j=1}^d \left\langle \sigma(X_t), \left(M_p(\mathtt{w}) \sqcup\!\sqcup\, k_p^{ij} \right) \bullet \mathtt{j} \right\rangle =$$
$$\langle \sigma(X), M_p(\mathtt{wi}) \rangle .$$

□

We finish this section with a generalization of Theorem 3.1.2 to polynomial maps that do not satisfy the condition $p(0) = 0$ and paths that do not start at the origin.

Corollary 3.3.5. *Let $X : [0, L] \longrightarrow \mathbb{R}^d$ be a piecewise continuously differentiable path and let $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a polynomial map. Consider the map $\tilde{p}$ given by $\tilde{p}(y) = p(y + X_0) - p(X_0)$. Then, for all $\mathtt{w} \in \mathrm{T}(\mathbb{R}^m)$,*

$$\langle \sigma(p(X)), \mathtt{w} \rangle = \langle \sigma(X), M_{\tilde{p}}(\mathtt{w}) \rangle .$$

Proof. The statement follows using the same argument as in the proof of Theorem 3.1.2 if we take into account that in this case,

$$J_p^{ij}(X_t) = J_{\tilde{p}}^{ij}(X_t - X_0) = \left\langle \sigma\left(X|_{[0,t]} \right), k_{\tilde{p}}^{ij} \right\rangle .$$

□

3.3.1 M_p as a half-shuffle homomorphism

The shuffle product can be seen as the symmetrization of the right half-shuffle, which we define in the following way. Let $\mathrm{T}^{\geq 1}(\mathbb{R}^d) = \bigoplus_{n \geq 1} \mathrm{T}^n(\mathbb{R}^d)$ denote the vector space spanned by the non-empty words built from d letters.

Definition 3.3.6 ([Sch59, Section IV, Equation (S2)],[FP13, Definition 2.1],cf. also [EM53, Sect. 18])**.** *The* right half-shuffle $\succ : \mathrm{T}^{\geq 1}(\mathbb{R}^d) \times \mathrm{T}^{\geq 1}(\mathbb{R}^d) \to \mathrm{T}^{\geq 1}(\mathbb{R}^d)$ *is recursively given on words as*

$$\mathtt{w} \succ \mathtt{i} := \mathtt{wi},$$
$$\mathtt{w} \succ \mathtt{vi} := (\mathtt{w} \succ \mathtt{v} + \mathtt{v} \succ \mathtt{w}) \bullet \mathtt{i},$$

where $\mathtt{w}, \mathtt{v}$ are words and $\mathtt{i}$ is a letter.

Therefore, for any non-empty words $\mathtt{w}, \mathtt{v}$

$$\mathtt{w} \sqcup\!\sqcup \mathtt{v} = \mathtt{w} \succ \mathtt{v} + \mathtt{v} \succ \mathtt{w}.$$

Indeed, for non-empty words $\mathtt{w}, \mathtt{v}$ and letters $\mathtt{i}, \mathtt{j}$, we have

$$\begin{aligned} \mathtt{i} \succ \mathtt{j} + \mathtt{j} \succ \mathtt{i} &= \mathtt{ij} + \mathtt{ji}, \\ \mathtt{wi} \succ \mathtt{j} + \mathtt{j} \succ \mathtt{wi} &= \mathtt{wij} + (\mathtt{j} \succ \mathtt{w} + \mathtt{w} \succ \mathtt{j}) \bullet \mathtt{i}, \quad \text{and} \\ \mathtt{wi} \succ \mathtt{vj} + \mathtt{vj} \succ \mathtt{wi} &= (\mathtt{wi} \succ \mathtt{v} + \mathtt{v} \succ \mathtt{wi}) \bullet \mathtt{j} + (\mathtt{vj} \succ \mathtt{w} + \mathtt{w} \succ \mathtt{vj}) \bullet \mathtt{i}, \end{aligned}$$

in accordance with Definition 3.2.2. Thus, the second equation in Definition 3.3.6 can be rewritten as

$$\mathtt{w} \succ \mathtt{vi} = (\mathtt{w} \sqcup\!\sqcup \mathtt{v}) \bullet \mathtt{i}. \tag{3.6}$$

It turns out that the right half-shuffle is an example of a more general type of algebras, the Zinbiel algebras.

Definition 3.3.7 ([EM53, Equation (18.7)][1], [Sch59, Section IV, Equation (S0)], [FP13, Equation (1.4)])**.** *A (right)* Zinbiel algebra *is a vector space Z together with a bilinear map $\succcurlyeq\, : Z \times Z \to Z$ such that, for all $a, b, c \in Z$,*

$$a \succcurlyeq (b \succcurlyeq c) = (a \succcurlyeq b + b \succcurlyeq a) \succcurlyeq c.$$

We include here the proof of the next result since it is interesting.

Theorem 3.3.8 ([Sch59, Section IV, Equation (S0)],[Lod95, Proposition 1.8])**.** $(\mathrm{T}^{\geq 1}(\mathbb{R}^d), \succ)$ *is a Zinbiel algebra, i.e. for any non-empty words* $\mathtt{w}$, $\mathtt{v}$, *and* $\mathtt{u}$,

$$\mathtt{w} \succ (\mathtt{v} \succ \mathtt{u}) = (\mathtt{w} \sqcup\!\sqcup \mathtt{v}) \succ \mathtt{u}. \tag{3.7}$$

Proof. Let $\mathtt{w}$, $\mathtt{v}$, and $\mathtt{u}$ be non-empty words and $\mathtt{i}$ be an arbitrary letter. By the definition of the half-shuffle and Equation (3.6), we have

$$\mathtt{w} \succ (\mathtt{v} \succ \mathtt{i}) = \mathtt{w} \succ \mathtt{vi} = (\mathtt{w} \sqcup\!\sqcup \mathtt{v}) \bullet \mathtt{i} = (\mathtt{w} \sqcup\!\sqcup \mathtt{v}) \succ \mathtt{i}.$$

Using Equation (3.6) and associativity of the shuffle product, Proposition 3.2.3, we obtain that

$$\mathtt{w} \succ (\mathtt{v} \succ \mathtt{ui}) = \mathtt{w} \succ \big((\mathtt{v} \sqcup\!\sqcup \mathtt{u}) \bullet \mathtt{i}\big) = \big(\mathtt{w} \sqcup\!\sqcup (\mathtt{v} \sqcup\!\sqcup \mathtt{u})\big) \bullet \mathtt{i} = \big((\mathtt{w} \sqcup\!\sqcup \mathtt{v}) \sqcup\!\sqcup \mathtt{u}\big) \bullet \mathtt{i} = (\mathtt{w} \sqcup\!\sqcup \mathtt{v}) \succ \mathtt{ui}.$$

□

In fact, it is known that $(\mathrm{T}^{\geq 1}(\mathbb{R}^d), \succ)$ is free in this case.

Theorem 3.3.9 ([Sch59, Section IV, page 19][2][Lod95, Proposition 1.8])**.** *Indeed,* $(\mathrm{T}^{\geq 1}(\mathbb{R}^d), \succ)$ *is the free Zinbiel algebra over* $\mathbb{R}^d$.

This means that for any Zinbiel algebra $(Z, \succcurlyeq)$ and any linear map $B : \mathbb{R}^d \to Z$, there is a unique homomorphism $\Lambda_B : (\mathrm{T}^{\geq 1}(\mathbb{R}^d), \succ) \to (Z, \succcurlyeq)$ such that $B = \Lambda_B \circ \iota$, where $\iota : \mathbb{R}^d \to \mathrm{T}^{\geq 1}(\mathbb{R}^d)$ is the canonical embedding. This is known as the *universal property of the free Zinbiel algebra* and is described in Diagram 3.3.1. We call Λ_B the *unique extension of B to a Zinbiel homomorphism.*

[1] Eilenberg and Mac Lane in [EM53, Section 18] discovered what we call today a ??, i.e. one has $(x \downarrow y) \downarrow z = x \downarrow (y \downarrow z + (-1)^{|y|+|z|} z \downarrow y)$ for homogeneous elements x, y.

[2] Note that Schützenberger in [Sch59, Section IV] treated the shuffle algebra as a $\mathbb{Z}$-module, and so his observation there was, stated in modern terms, the fact that the $\mathbb{Z}$ shuffle algebra is the free $\mathbb{Z}$ Zinbiel algebra.

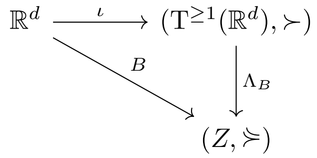


Diagram 3.3.1: Universal property of the free Zinbiel algebra

Proof. (Based on the proof of [Lod95, Proposition 1.8].) Define the linear map $\Lambda_B : \mathrm{T}^{\geq 1}(\mathbb{R}^d) \to Z$ recursively by

$$\Lambda_B\mathtt{i} := L\mathtt{i}, \qquad \Lambda_B\mathtt{vi} := \Lambda_B\mathtt{v} \succcurlyeq \Lambda_B\mathtt{i},$$

where we identified $\mathbb{R}^d$ with the letters in $\mathrm{T}^{\geq 1}(\mathbb{R}^d)$. Since $\mathtt{vi} = \mathtt{v} \succ \mathtt{i}$, this is the only candidate for a map with the desired properties. It remains to show that it is indeed a homomorphism of Zinbiel algebras.

By definition, it holds that $\Lambda_B\mathtt{ij} = \Lambda_B\mathtt{i} \succcurlyeq \Lambda_B\mathtt{j}$. Thus, assume that $\Lambda_B\mathtt{xy} = \Lambda_B\mathtt{x} \succcurlyeq \Lambda_B\mathtt{y}$ holds for all nonempty words $\mathtt{x}$, $\mathtt{y}$ such that $|\mathtt{x}| + |\mathtt{y}| = n$. Then, for all nonempty words $\mathtt{w}$ and $\mathtt{v}$ such that $|\mathtt{w}| + |\mathtt{v}| = n$, we have

$$\begin{aligned}\Lambda_B\mathtt{w} \succcurlyeq \Lambda_B\mathtt{vi} &= \Lambda_B\mathtt{w} \succcurlyeq (\Lambda_B\mathtt{v} \succcurlyeq \Lambda_B\mathtt{i}) = (\Lambda_B\mathtt{w} \succcurlyeq \Lambda_B\mathtt{v} + \Lambda_B\mathtt{v} \succcurlyeq \Lambda_B\mathtt{w}) \succcurlyeq \Lambda_B\mathtt{i} \\ &= \Lambda_B(\mathtt{w} \succ \mathtt{v} + \mathtt{v} \succ \mathtt{w}) \succcurlyeq \Lambda_B\mathtt{i} = \Lambda_B((\mathtt{w} \succ \mathtt{v} + \mathtt{v} \succ \mathtt{w}) \succ \mathtt{i}) = \Lambda_B(\mathtt{w} \succ (\mathtt{v} \succ \mathtt{i})) \\ &= \Lambda_B(\mathtt{w} \succ \mathtt{vi}),\end{aligned}$$

and $\Lambda_B\mathtt{ui} := \Lambda_B\mathtt{u} \succcurlyeq \Lambda_B\mathtt{i}$, for all nonempty words $\mathtt{u}$ with $|\mathtt{u}| = n$ by definition. The claim follows by induction over n. □

The following result describes the relation between the map M_p and the half-shuffle.

Theorem 3.3.10. *The restriction of M_p to $\mathrm{T}^{\geq 1}(\mathbb{R}^d)$, denoted $M_p|_{\mathrm{T}^{\geq 1}(\mathbb{R}^d)}$, is the unique half-shuffle homomorphism such that $M_p(\mathtt{i}) = \varphi_d(p_i)$.*

Proof. We have $M_p(\mathtt{i}) = \varphi(p_i)$ by Proposition 3.3.2 (3). Then, using (3.6) and the definition of M_p, we get

$$M_p(\mathtt{wi}) = \sum_{j=1}^{d} \left(M_p(\mathtt{w}) \sqcup\!\sqcup\, k_p^{ij}\right) \bullet \mathtt{j} = M_p(\mathtt{w}) \succ \left(\sum_{j=1}^{d} k_p^{ij} \bullet \mathtt{j}\right) = M_p(\mathtt{w}) \succ M_p(\mathtt{i}),$$

and thus the statement follows immediately from the proof of Theorem 3.3.9. □

We finish this section with a generalization of Theorem 3.1.2 in terms of Zinbiel algebras stated in the following result.

Theorem 3.3.11. *Let $X : [0, L] \longrightarrow \mathbb{R}^d$ be a piecewise continuously differentiable path and $B : \mathbb{R}^m \to (\mathrm{T}^{\geq 1}(\mathbb{R}^d), \succ)$ be a linear map. Then, the signature of the path*

$$Y : [0, L] \longrightarrow \mathbb{R}^m, \quad Y_t^i := \left\langle \sigma\left(X|_{[0,t]}\right), B\mathtt{i} \right\rangle,$$

is a linear transformation of the signature of X, namely

$$\langle \sigma(Y), \mathtt{w} \rangle = \langle \sigma(X), \Lambda_B\mathtt{w} \rangle,$$

where Λ_B is the unique extension of B to a Zinbiel homomorphism.

Remark 3.3.12. To see that this result indeed implies Theorem 3.1.2, consider a piecewise continuously differentiable path $X : [0, L] \longrightarrow \mathbb{R}^d$ with $X_0 = 0$ and a polynomial map $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ with $p(0) = 0$. Let then $B : \mathbb{R}^m \to (\mathrm{T}^{\geq 1}(\mathbb{R}^d), \succ)$ be the linear map given by $B\mathtt{i} = \varphi_d(p_i)$ and

$$Y : [0, L] \longrightarrow \mathbb{R}^m, \quad Y_t^i := \left\langle \sigma\left(X|_{[0,t]}\right), B\mathtt{i}\right\rangle$$

as in the previous result. Since

$$\langle \sigma\left(X|_{[0,t]}\right), \varphi_d(x_i)\rangle = \langle \sigma\left(X|_{[0,t]}\right), \mathtt{i}\rangle = X_t^i - X_0^i = X_t^i,$$

and $q \longmapsto \langle \sigma\left(X|_{[0,t]}\right), \varphi_d(q)\rangle$ is an algebra homomorphism due to the shuffle identity Proposition 3.2.4, we have that

$$\left\langle \sigma\left(X|_{[0,t]}\right), B\mathtt{i}\right\rangle = \left\langle \sigma\left(X|_{[0,t]}\right), \varphi_d(p_i)\right\rangle = p_i(X),$$

and thus $Y = p(X)$. Note Theorem 3.3.10 implies $\Lambda_B = M_p$, and thus the statement of Theorem 3.3.11 in this case reads as

$$\langle \sigma(p(X)), \mathtt{w}\rangle = \langle \sigma(X), M_p\mathtt{w}\rangle,$$

which is exactly the relation we get from Theorem 3.1.2.

In Example 3.3.14, we will look at a situation which is not covered by Theorem 3.1.2, yet does fall under the scope of Theorem 3.3.11, meaning that the latter is indeed more general.

Before proving this result, we introduce some notation. We denote by X_{0s}^z the coefficient of z in the signature of the path X restricted to the interval $[0, s]$. That is,

$$X_{0s}^z := \langle \sigma(X\restriction_{[0,s]}), z\rangle.$$

Note that $X_{0s}^{\mathtt{i}} = X_s^i - X_0^i$. Then, the path Y introduced in Theorem 3.3.11 is given by $Y_s^{\mathtt{i}} = X_{0s}^{B\mathtt{i}}$. Moreover, for any letter $\mathtt{i}$, we define the maps $\mathrm{T}_{\mathtt{i}}^-, \mathrm{T}_{\mathtt{i}}^+ : \mathrm{T}(\mathbb{R}^d) \to \mathrm{T}(\mathbb{R}^d)$ to be the unique linear maps given recursively by $\mathrm{T}_{\mathtt{i}}^+\,\mathtt{w} = \mathtt{wi}$ and by $\mathrm{T}_{\mathtt{i}}^-\,\mathtt{wj} = \delta_{\mathtt{ij}}\mathtt{w}$ with $\mathrm{T}_{\mathtt{i}}^-\,\mathtt{e} = 0$, respectively, for any word $\mathtt{w}$.

These two maps allow us to define the right half-shuffle as shown in the following technical result.

Lemma 3.3.13. *For any $x, y \in \mathrm{T}^{\geq 1}(\mathbb{R}^d)$, we have* $x \succ y = \sum_{\mathtt{i}=1}^{\mathtt{d}} \mathrm{T}_{\mathtt{i}}^+(x \sqcup\!\sqcup \mathrm{T}_{\mathtt{i}}^-\, y)$.

Proof. This is just a reformulation of (3.6) in the following way. For any word $\mathtt{v}$, any non-empty word $\mathtt{w}$, and any letter $\mathtt{j}$, we have that

$$\sum_{\mathtt{i}=1}^{\mathtt{d}} \mathrm{T}_{\mathtt{i}}^+(\mathtt{w} \sqcup\!\sqcup \mathrm{T}_{\mathtt{i}}^-\,\mathtt{vj}) = \mathrm{T}_{\mathtt{j}}^+(\mathtt{w} \sqcup\!\sqcup \mathtt{v}) = (\mathtt{w} \sqcup\!\sqcup \mathtt{v}) \bullet \mathtt{j} = \mathtt{w} \succ \mathtt{vj},$$

where the last equality follows from (3.6). Then, the general statement for any $x, y \in \mathrm{T}^{\geq 1}(\mathbb{R}^d)$ follows from (bi)linearity. □

Now we are ready to prove Theorem 3.3.11.

Proof of Theorem 3.3.11. For better readability, we put $\Lambda := \Lambda_B$. First note that by the definition of the signature and the fact that X is continuously differentiable almost everywhere, we have

$$Y_s^{\mathtt{i}} = X_{0s}^{B\mathtt{i}} = X_{0s}^{\Lambda\mathtt{i}} = \sum_{i=1}^{d} \int_0^s X_{0t}^{\mathrm{T}_{\mathtt{i}}^-\,\Lambda\mathtt{i}} \mathrm{d}X_t^i = \sum_{i=1}^{d} \int_0^s X_{0t}^{\mathrm{T}_{\mathtt{i}}^-\,\Lambda\mathtt{i}} \dot{X}_t^i \mathrm{d}t$$

and thus, for almost all $s \in [0, L]$,

$$\dot{Y}_s^{\mathtt{i}} = \sum_{i=1}^{d} X_{0t}^{\mathrm{T}_{\mathtt{i}}^{-}\ \Lambda\mathtt{i}} \dot{X}_t^i.$$

Following an inductive argument, assume now that $X_{0s}^{\Lambda\mathtt{w}} = Y_{0s}^{\mathtt{w}}$ holds for some word $\mathtt{w}$. Then,

$$\begin{aligned} Y_{0s}^{\mathtt{wi}} &= \int_0^s Y_{0t}^{\mathtt{w}} \mathrm{d}Y_t^i = \int_0^s Y_{0t}^{\mathtt{w}} \dot{Y}_t^i \mathrm{d}t = \sum_{l=1}^{d} \int_0^s X_{0t}^{\Lambda\mathtt{w}} X_{0t}^{\mathrm{T}_l^{-}\ \Lambda\mathtt{i}} \dot{X}_t^l \mathrm{d}t \\ &= \int_0^s X_{0t}^{\sum_{\mathtt{l}=1}^{\mathtt{d}} \Lambda\mathtt{w} \sqcup\!\sqcup \mathrm{T}_{\mathtt{l}}^{-}\ \Lambda\mathtt{i}} \mathrm{d}X_{0t}^{\mathtt{l}} = X_{0t}^{\sum_{\mathtt{l}=1}^{\mathtt{d}} \mathrm{T}_{\mathtt{l}}^{+}(\Lambda\mathtt{w} \sqcup\!\sqcup \mathrm{T}_{\mathtt{l}}^{-}\ \Lambda\mathtt{i})} = X_{0t}^{\Lambda\mathtt{w} \succ \Lambda\mathtt{i}} = X_{0t}^{\Lambda\mathtt{wi}}, \end{aligned}$$

where we used Lemma 3.3.13 and the fact that Λ is a homomorphism of Zinbiel algebras. □

Theorem 3.3.11 can also be directly shown using

$$\int_0^s X_{0t}^x \mathrm{d}X_{0t}^y = X_{0s}^{x \succ y} \tag{3.8}$$

for any $x, y \in \mathrm{T}^{\geq 1}(\mathbb{R}^d)$, a relation which is quite fundamental for an algebraic understanding of the signature and was mentioned already for example in [GK08] right after Equation (6). Conversely, starting from Theorem 3.3.11, equation (3.8) is immediate with the choice $B\mathtt{1} = x, B\mathtt{2} = y$.

The following example illustrates the results presented in this section.

Example 3.3.14. *For a given path $X : [0, L] \to \mathbb{R}^3$, let*

$$Y = \big(\mathrm{Area}(X^2, X^3), \mathrm{Area}(X^3, X^1), \mathrm{Area}(X^1, X^2)\big)$$

denote the 'area path' of X, where $\mathrm{Area}(X^i, X^j)_t = \int_0^t \int_0^s \mathrm{d}X_u^i \mathrm{d}X_s^j - \int_0^t \int_0^s \mathrm{d}X_u^j \mathrm{d}X_s^i$. Let us compute $\langle \sigma(Y), \mathtt{12} - \mathtt{21} \rangle$ in the special case that $X : [0, 1] \to \mathbb{R}^3$, $X(t) = (t, t^2, t^3)$.

$$\mathrm{Area}(X^i, X^j)_t = \int_0^t \int_0^s \mathrm{d}X_u^i \mathrm{d}X_s^j - \int_0^t \int_0^s \mathrm{d}X_u^j \mathrm{d}X_s^i.$$

Let us compute $\langle \sigma(Y), \mathtt{12} - \mathtt{21} \rangle$ in the special case that $X : [0, 1] \to \mathbb{R}^3$, with $X(t) = (t, t^2, t^3)$.

First of all, we have

$$Y_s^i = \big\langle \sigma(X|_{[0,T]}), B\mathtt{i} \big\rangle,$$

where B denotes the linear map from $\mathbb{R}^3$ (interpreted as the vector space spanned by the letters $\mathtt{1}, \mathtt{2}, \mathtt{3}$) to $\mathrm{T}(\mathbb{R}^3)$ given by

$$B\mathtt{1} = \mathtt{23} - \mathtt{32}, \quad B\mathtt{2} = \mathtt{31} - \mathtt{13}, \quad B\mathtt{3} = \mathtt{12} - \mathtt{21}.$$

Theorem 3.3.11 thus applies, and in order to see how $\langle \sigma(Y), \mathtt{12} - \mathtt{21} \rangle$ can be expressed in terms of iterated integrals of X, we only need to compute $\Lambda_B(\mathtt{12} - \mathtt{21}) = B\mathtt{1} \succ B\mathtt{2} - B\mathtt{2} \succ B\mathtt{1}$. To this end,

$$\begin{aligned} (\mathtt{23} - \mathtt{32}) \succ (\mathtt{31} - \mathtt{13}) &= 2 \cdot \mathtt{2331} - 2 \cdot \mathtt{3321} - \mathtt{2313} - \mathtt{2133} - \mathtt{1233} + \mathtt{3213} + \mathtt{3123} + \mathtt{1323}, \\ (\mathtt{31} - \mathtt{13}) \succ (\mathtt{23} - \mathtt{32}) &= \mathtt{3123} + \mathtt{3213} + \mathtt{2313} - \mathtt{1323} - \mathtt{1233} - \mathtt{2133} - \mathtt{3132} - 2 \cdot \mathtt{3312} \\ &\quad + 2 \cdot \mathtt{1332} + \mathtt{3132}. \end{aligned}$$

Thus, $\Lambda_B(12-21) = 2\cdot(-1323+1332+2313-2331+3312-3321)$. *Since in our special case of* $X(t)=(t,t^2,t^3)$ *it holds that [DR19, Remark 3.30][AFS19, Example 2.2]*

$$\int_0^1\int_0^u\int_0^t\int_0^s \mathrm{d}X_r^i\mathrm{d}X_s^j\mathrm{d}X_t^k\mathrm{d}X_u^l = \frac{j\cdot k\cdot l}{(j+i)(k+j+i)(l+k+j+i)},$$

we finally obtain that

$$\begin{aligned}\langle\sigma(Y), 12-21\rangle &= 2\cdot\left(-\frac{3\cdot2\cdot3}{4\cdot6\cdot9}+\frac{3\cdot3\cdot2}{4\cdot7\cdot9}+\frac{3\cdot1\cdot3}{5\cdot6\cdot9}-\frac{3\cdot3\cdot1}{5\cdot8\cdot9}+\frac{3\cdot1\cdot2}{6\cdot7\cdot9}-\frac{3\cdot2\cdot1}{6\cdot8\cdot9}\right)\\ &= -\frac{1}{315}\approx -0.00317.\end{aligned}$$

We hereby computed the value of a particular area of areas *of the original path* X*, namely* $\mathrm{Area}(\mathrm{Area}(X^2,X^3),\mathrm{Area}(X^3,X^1))_1$. *See [DLPR21] and Chapter 4 of this thesis for more on the algebraic theory behind* areas of areas*.*

3.4 Examples and consequences

Let us start with an *easy example* to illustrate how Theorem 3.1.2 works, and also the property (5) in Proposition 3.3.2.

Example 3.4.1. *Consider the polynomial map* $p:\mathbb{R}^2\longrightarrow\mathbb{R}^3$ *given by the polynomials* $p_1=x^2$, $p_2=y^3$ *and* $p_3=x-y$. *Consider also the path in Example 3.2.7,* $X:[0,1]\longrightarrow\mathbb{R}^2$ *given by* $X^1(t)=t$ *and* $X^2(t)=t^2$. *We want to compute a few terms in the signature of the path* $p(X)$.

We start computing the Jacobian matrix and its image under φ*:*

$$\left(J_p^{ij}\right)_{i,j} = \begin{pmatrix}2x & 0\\ 0 & 3y^2\\ 1 & -1\end{pmatrix} \qquad \text{and} \qquad \left(k_p^{ij}\right)_{i,j} = \left(\varphi\left(J_p^{ij}\right)\right)_{i,j} = \begin{pmatrix}2\cdot 1 & 0\\ 0 & 6\cdot 22\\ \mathtt{e} & -\mathtt{e}\end{pmatrix}.$$

Notice that $\varphi(3y^2) = 3\cdot 2 \sqcup\!\sqcup 2 = 6\cdot 22$. *We use Definition 3.3.1 to compute the image of a few words:*

$$\begin{aligned}&M_p(1) = 2\cdot 11, \qquad M_p(2) = 6\cdot 222, \qquad M_p(3) = 1-2,\\ &M_p(33) = (1-2)\bullet 1 - (1-2)\bullet 2 = 11+22-12-21.\end{aligned}$$

We observe that for any word $\mathtt{w}$ *above,* $\ell(M_p(\mathtt{w}))\leq 3\cdot\ell(\mathtt{w})$. *This is due to property* (5) *in Proposition 3.3.2 since* $\deg(p)=3$. *Now, applying Theorem 3.1.2 and looking at the signature terms computed in Example 3.2.7, we obtain that*

$$\begin{aligned}\langle\sigma(p(X)), 1\rangle &= \langle\sigma(X), M_p(1)\rangle = \langle\sigma(X), 2\cdot 11\rangle = 2\cdot\frac{1}{2} = 1,\\ \langle\sigma(p(X)), 2\rangle &= \langle\sigma(X), M_p(2)\rangle = \langle\sigma(X), 6\cdot 222\rangle = 6\cdot\frac{1}{6} = 1,\\ \langle\sigma(p(X)), 3\rangle &= \langle\sigma(X), M_p(3)\rangle = \langle\sigma(X), 1-2\rangle = 1-1 = 0, \text{ and}\\ \langle\sigma(p(X)), 33\rangle &= \langle\sigma(X), M_p(33)\rangle = \langle\sigma(X), 11+22-12-21\rangle = 0.\end{aligned}$$

This second example is more generic and shows the property (6) in Proposition 3.3.2.

Example 3.4.2. *Let X be any piecewise continuously differentiable path in $\mathbb{R}^2$. Consider the polynomial map $p : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ given by $p(x, y) = (x^2, xy, y^2)$, and fix $k = 2$.*

By Theorem 3.1.2, the coefficient of $\mathtt{w}$ in $\sigma(p(X))$ is given by the coefficient of $M_p(\mathtt{w})$ in $\sigma(X)$. Moreover, by Proposition 3.3.2 (6), the words appearing in $M_p(\mathtt{w})$ have length exactly 4. One way of storing $\sigma^{(2)}(p(X))$ is using a matrix that encodes the coefficients in $M_p(\mathtt{w})$,

$$\underset{\sigma^{(2)}(p(X))}{\begin{bmatrix} \mathtt{11} \\ \mathtt{12} \\ \mathtt{13} \\ \mathtt{21} \\ \mathtt{22} \\ \mathtt{23} \\ \mathtt{31} \\ \mathtt{32} \\ \mathtt{33} \end{bmatrix}} \rightsquigarrow \begin{bmatrix} 12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & 2 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 4 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 4 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 2 & 2 & 0 & 0 & 2 & 2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 4 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 2 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \end{bmatrix} \overset{\sigma^{(4)}(X)}{\begin{bmatrix} \mathtt{1111} \\ \mathtt{1112} \\ \mathtt{1121} \\ \mathtt{1122} \\ \mathtt{1211} \\ \mathtt{1212} \\ \mathtt{1221} \\ \mathtt{1222} \\ \mathtt{2111} \\ \mathtt{2112} \\ \mathtt{2121} \\ \mathtt{2122} \\ \mathtt{2211} \\ \mathtt{2212} \\ \mathtt{2221} \\ \mathtt{2222} \end{bmatrix}}.$$

More generally, by Proposition 3.3.2 (6), for any word $\mathtt{w}$ in the alphabet $\{\mathtt{1}, \mathtt{2}, \mathtt{3}\}$ with $\ell(\mathtt{w}) = k$, $M_p(\mathtt{w})$ is a sum of words in the alphabet $\{\mathtt{1}, \mathtt{2}\}$ of length $2k$. Applying Theorem 3.1.2, the information of $\sigma^{(k)}(p(X))$ can be stored in terms of $\sigma^{(2k)}(X)$. In fact, there exists a matrix M of format $3^k \times 2^{2k}$ that describes the change of coordinates *in the following sense. Fix an order on the words, for instance,* lexicographic order *as above. The coefficients of $\sigma^{(k)}(p(X))$ and the coefficients of $\sigma^{(2k)}(X)$ are related as*

$$\underset{\sigma^{(k)}(p(X))}{\begin{bmatrix} \overbrace{\mathtt{11\ldots1}}^{k} \\ \mathtt{1\ldots12} \\ \vdots \\ \mathtt{3\ldots32} \\ \underbrace{\mathtt{3\ldots33}}_{k} \end{bmatrix}} = M \underset{\sigma^{(2k)}(X)}{\begin{bmatrix} \overbrace{\mathtt{11\ldots1}}^{2k} \\ \mathtt{1\ldots12} \\ \vdots \\ \mathtt{2\ldots21} \\ \underbrace{\mathtt{2\ldots22}}_{2k} \end{bmatrix}},$$

where the row indexed by $\mathtt{w}$ in M is given by the coefficients of $M_p(\mathtt{w})$. Moreover, we want to point out the following properties of this homomorphism M_p:

- *For* $\mathtt{w} = \underbrace{\mathtt{1\ldots1}}_{k \text{ times}}$, $M_p(\mathtt{w}) = \dfrac{(2k)!}{k!} \cdot \mathtt{w} \bullet \mathtt{w}$.
- *For* $\mathtt{w} = \underbrace{\mathtt{2\ldots2}}_{k \text{ times}}$, $M_p(\mathtt{w}) = k! \cdot \underbrace{\mathtt{1\ldots1}}_{k \text{ times}} \sqcup\!\sqcup \underbrace{\mathtt{2\ldots2}}_{k \text{ times}}$.

- *For* $\mathtt{w} = \underbrace{\mathtt{3}\ldots\mathtt{3}}_{k\ times}$, $M_p(\mathtt{w}) = \frac{(2k)!}{k!}\underbrace{\mathtt{2}\ldots\mathtt{2}}_{2k\ times}$.

The rest of this section is dedicated to analysing the consequences of Theorems 3.1.1 and 3.1.2 in several particular cases. We start looking at the case in which X is itself a polynomial map and we need the following definition.

Definition 3.4.3. *For an element* $\mathtt{a} \in \mathrm{T}((\mathbb{R}^d))$*, with zero coefficient for the empty word* $\mathtt{e}$*, we define the* concatenation product exponential *of* $\mathtt{a}$ *as*

$$\exp_\bullet(\mathtt{a}) := \sum_{n\geq 0} \frac{\mathtt{a}^{\bullet n}}{n!}.$$

More information on this exponential map and its inverse, the *logarithm*, can be found in [Reu93, Section 3.1].

Corollary 3.4.4. *Let* $X : [0, L] \longrightarrow \mathbb{R}^d$ *be a polynomial map, with* $L \in \mathbb{R}$*,* $L \geq 1$*. Then, for any* $\mathtt{w} \in \mathrm{T}(\mathbb{R}^d)$*,*

$$\langle \sigma(X), \mathtt{w}\rangle = \langle \exp_\bullet(L \cdot \mathtt{1}), M_{\tilde{X}}(\mathtt{w})\rangle\,,$$

where $\tilde{X}(y) = X(y) - X_0$*. Equivalently,* $\sigma(X) = M^*_{\tilde{X}}(\exp_\bullet(L \cdot \mathtt{1}))$*.*

Proof. Let $Y : [0, L] \longrightarrow \mathbb{R}$ be the path given by $Y(t) = t$. Then, $\sigma(Y) = \exp_\bullet(L \cdot \mathtt{1})$. The statement follows by Corollary 3.3.5 applied to the path Y and the polynomial map X. □

The next result looks at the case when $M^*_p(\sigma(X)) = 0$ from the perspective of the polynomial map and of the piecewise continuously differentiable path. We first introduce two concepts. Given a polynomial map p, we define the *ideal generated by* p as the ideal I_p generated by the polynomials that define the map p, i.e. $I_p = \langle p_1, \ldots, p_m\rangle$. Moreover, we define a *tree-like path* as a path X such that $\sigma(X) = 0$. This definition is the characterization obtained for bounded variation paths by B. M. Hambly and T. J. Lyons in [HL10, Theorem 4], while a more general topological definition can be found in [BGLY16, Definition 1.1] and in Section 1.1 of this thesis.

Remark 3.4.5. A very simple example of a tree-like path is a concatenation of paths $A\sqcup B\sqcup C\sqcup D$ such that the paths A and B (resp. C and D) are of the same shape, but parametrized in the opposite direction. When we compute the integrals on such a path, we get cancellations and the signature of the path does not see the $A \sqcup B \sqcup C \sqcup D$ loop, i.e. $\sigma(A \sqcup B \sqcup C \sqcup D) = \mathbf{e}$, the empty word.

The following result describes the situation in which the (image of the) path lies in the zeros of the ideal I_p, for some polynomial map p.

Corollary 3.4.6. *Let* $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ *be a polynomial map with associated ideal* I_p*. We define the polynomial map* $\tilde{p}(y) = p(y + X_0) - p(X_0)$*, for which* $\tilde{p}(0) = 0$*.*

- *If* $X : [0, L] \longrightarrow \mathbb{R}^d$ *is a piecewise continuously differentiable path such that* $X(t) \in \mathcal{V}(I_p)$ *for all* $t \in [0, L]$*, then* $M^*_{\tilde{p}}(\sigma(X)) = \mathbf{e}$*.*

- *Conversely, if* $M^*_{\tilde{p}}(\sigma(X)) = \mathbf{e}$*, for some piecewise continuously differentiable path* $X : [0, L] \longrightarrow \mathbb{R}^d$*, then* $p(X)$ *is tree-like.*

The last consequence is that the dual map M^*_p behaves nicely with respect to the concatenation of signatures.

Corollary 3.4.7. *Let* $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ *be a polynomial map with* $p(0) = 0$ *and let* $X, Y : [0, L] \longrightarrow \mathbb{R}^d$ *be two piecewise continuously differentiable paths with* $X_0 = Y_0 = 0$. *Then,*

$$M_p^*(\sigma(X) \bullet \sigma(Y)) = M_p^*(\sigma(X)) \bullet M_q^*(\sigma(Y)),$$

where $q(y) = p(y + X_L - X_0) - p(X_L - X_0)$.

3.5 Applications and future work

The results presented in this chapter solve an algebraic question that arises from the signatures of paths, an object commonly studied in stochastic analysis. There are several interesting problems that do not fit on the algebraic flavour of this chapter. We summarize them in the following list.

(A) In comparison with the results presented in [PSS19], we would like to explore a non-linear version of their approach using dictionaries. The idea is that if we have a family of generic paths, χ, for which we know the signature and a polynomial map p, then Theorem 3.1.2 allows us to compute the signature of all the paths in $p(\chi)$.

For instance, Example 3.4.2 shows that we can compute the signature of $p(X)$ for any piecewise continuously differentiable path X by multiplying the signature of X by a matrix at each level. Therefore, we have the following question:

Is it possible to understand $\sigma(p(X))$ *in terms of* $\sigma(X)$ *and* M_p *in the language of tensors?*

(B) Another line of future research is focused on the map M_p. Since it is defined from a polynomial map without involving any piecewise continuously differentiable path and gives us the commuting diagram (3.3), we intuit that it is worth to look for more interesting properties. For that, we should look to the *big picture* involving the Hopf algebra structures, as well as other constructions.

In this direction, at the end of Example 3.4.2 we describe combinatorially $M_p(\mathtt{w})$ for some particular words $\mathtt{w}$. A more general question would be the following:

For which words $\mathtt{w}$ and polynomial maps p is there a *non-recursive* combinatorial formula for $M_p(\mathtt{w})$?

Answering this question could be very useful from the computational perspective.

Chapter 4

Areas of areas generate the shuffle algebra

This chapter is based on [DLPR21]. Changes and additions to the material have been made by the author of this thesis towards the version presented in this chapter.

4.1 Introduction

We give a concise introduction here and spell out the notation more fully in the next section. The shuffle algebra $T(\mathbb{R}^d)$ over d letters is the vector space spanned by words in the letters $\mathtt{1}, \ldots, \mathtt{d}$ with the commutative shuffle product. This is a free commutative algebra over the Lyndon words. Put differently, it can be viewed as a polynomial algebra in new commuting variables x_w, where w ranges over all Lyndon words. That is, as commutative algebras,

$$\mathbb{R}[x_w : \ w \text{ is Lyndon}] \cong T(\mathbb{R}^d).$$

The isomorphism is given by $x_w \mapsto w \in T(\mathbb{R}^d)$. There are many more (free) generators known: any basis for the Lie algebra, coordinates of the first kind, ... (compare Corollary 4.4.5).

The relevance for iterated integrals is as follows. Let $X : [0,T] \to \mathbb{R}^d$ be a (piecewise smooth) curve and let $f_i \in T(\mathbb{R}^d), i \in I$, be a generating set of the shuffle algebra. Then: any term in the iterated-integrals signature $S(X)_{0,T}$ (introduced in [Che54] under the name *exponential homomorphism*, see also e.g. [LCL07, Chapter 2]) is a polynomial in the real numbers

$$\Big\langle f_i, S(X)_{0,T} \Big\rangle, i \in I.$$

Indeed, by assumption, any word w can be written as

$$w = P_{\sqcup\!\sqcup}(f_i : i \in I),$$

where $P_{\sqcup\!\sqcup}$ is some shuffle polynomial in finitely many of the f_i. By the shuffle identity we then get

$$\Big\langle w, S(X)_{0,T} \Big\rangle = \Big\langle P_{\sqcup\!\sqcup}(f_i : i \in I), S(X)_{0,T} \Big\rangle = P\left(\Big\langle f_i, S(X)_{0,T} \Big\rangle : i \in I\right),$$

where P is the corresponding polynomial expression in the real numbers $\langle f_i, S(X)_{0,T}\rangle$, $i \in I$. The latter numbers then contain all the information of the iterated-integrals signature, since every iterated integral is a polynomial expression in them.

We are interested in whether there is a shuffle generating set in terms of "areas of areas". Define the following bilinear operation on $T(\mathbb{R}^d)$

$$\mathsf{area}(x, y) := x \succ y - y \succ x,$$

where $\succ$ denotes the half-shuffle. For $v, w \in T(\mathbb{R}^d)$, let

$$V_t = \Big\langle v, S(X)_{0,t}\Big\rangle \qquad \text{and} \qquad W_t = \Big\langle w, S(X)_{0,t}\Big\rangle,$$

and define

$$\mathsf{Area}(V, W)_t := \int_0^t \int_0^s dV_r dW_s - \int_0^t \int_0^s dW_r dV_s.$$

We then have

$$\mathsf{Area}(V, W)_t = \Big\langle \mathsf{area}(v, w), S(X)_{0,t}\Big\rangle.$$

Our naming of area and Area stems from the fact that $\mathsf{Area}(V, W)$ is (two times) the *signed area* (see Figure 4.1) enclosed by the two-dimensional curve (V, W), see [LY06, Proposition 1 (2)] and the remark on geometric interpretation at the beginning of [DR19, Section 3.2][1] See [CK16, Section 1.2.4 and Sections 2.1.2-2.1.4] for a detailed discussion of the signed area with lots of illustrations. Note that the antisymmetrization $\mathsf{Area}(V, W)_t$ of the Riemann-Stieltjes integral (where the Riemann-Stieltjes integral forms a Zinbiel algebra on a suitable space of functions V with $V_0 = 0$) has already been looked at as an algebraic operation by Rocha in 2003 in [Roc03b, Equation (7)], in [Roc03a, Equation (6.11)] and in 2005 in [Roc05, Equation (2.4)]. It was even already noted by Rocha in [Roc03b, Section 3, page 321] and [Roc05, Section 2, page 3] that the operation Area except being antisymmetric does not satisfy any additional identity of order three.

The following question is inspired by a remark made by T.L. during a talk in 2011:

Is repeated application of the Area *operator enough to get the whole signature of a path* X*?*

For $d = 2$ and the first two levels, this is quickly verified. We start with the increments themselves, which we assume to be given (we think of them as "0-th order" areas), which are $\int dX^1$ and $\int dX^2$. Then we can write, using integration-by-parts,

$$\begin{aligned}
\iint dX^1 dX^1 &= \frac{1}{2} \int dX^1 \cdot \int dX^1 \\
\iint dX^2 dX^2 &= \frac{1}{2} \int dX^2 \cdot \int dX^2 \\
\iint dX^1 dX^2 &= \frac{1}{2}\Big(\iint dX^1 dX^2 - \iint dX^2 dX^1 + \int dX^1 \cdot \int dX^2 \Big) \\
\iint dX^2 dX^1 &= \frac{1}{2}\Big(- \Big(\iint dX^1 dX^2 - \iint dX^2 dX^1 \Big) + \int dX^1 \cdot \int dX^2 \Big).
\end{aligned}$$

[1] thanks to Joscha Diehl for pointing the author of this thesis to the exact references here.

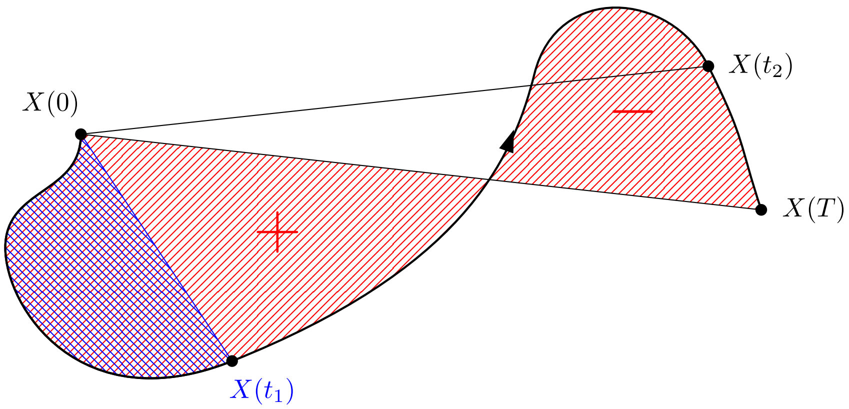


Figure 4.1: The signed area of a curve X, shown at points $t = t_1$ (shaded blue) and at $t = T$ (shaded red).

and hence get all iterated integrals up to order 2.

Products of integrals become, on the algebra side, $\sqcup\!\sqcup$-products. This reads as

$$\mathtt{11} = \frac{1}{2}\mathtt{1} \sqcup\!\sqcup \mathtt{1} \qquad \mathtt{22} = \frac{1}{2}\mathtt{2} \sqcup\!\sqcup \mathtt{2}$$

$$\mathtt{12} = \frac{1}{2}\left(\mathsf{area}(\mathtt{1},\mathtt{2}) + \mathtt{1} \sqcup\!\sqcup \mathtt{2}\right) \qquad \mathtt{21} = \frac{1}{2}\left(-\mathsf{area}(\mathtt{1},\mathtt{2}) + \mathtt{1} \sqcup\!\sqcup \mathtt{2}\right)$$

In general, however, the expansion is non-unique, as the following example illustrates:

$$\begin{aligned}\mathtt{123} &= \tfrac{1}{3}\mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{2},\mathtt{3})) + \tfrac{1}{6}\mathsf{area}(\mathsf{area}(\mathtt{1},\mathtt{3}),\mathtt{2}) + \tfrac{1}{3}\mathtt{1} \sqcup\!\sqcup \mathsf{area}(\mathtt{2},\mathtt{3}) \\ &\quad - \tfrac{1}{6}\mathtt{2} \sqcup\!\sqcup \mathsf{area}(\mathtt{1},\mathtt{3}) + \tfrac{1}{2}\mathtt{3} \sqcup\!\sqcup \mathsf{area}(\mathtt{1},\mathtt{2}) + \tfrac{1}{6}\mathtt{1} \sqcup\!\sqcup \mathtt{2} \sqcup\!\sqcup \mathtt{3} \\ &= \tfrac{1}{12}\mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{2},\mathtt{3})) - \tfrac{1}{12}\mathsf{area}(\mathsf{area}(\mathtt{1},\mathtt{3}),\mathtt{2}) + \tfrac{1}{4}\mathsf{area}(\mathsf{area}(\mathtt{1},\mathtt{2}),\mathtt{3}) \\ &\quad + \tfrac{1}{12}\mathtt{1} \sqcup\!\sqcup \mathsf{area}(\mathtt{2},\mathtt{3}) + \tfrac{1}{12}\mathtt{2} \sqcup\!\sqcup \mathsf{area}(\mathtt{1},\mathtt{3}) + \tfrac{1}{4}\mathtt{3} \sqcup\!\sqcup \mathsf{area}(\mathtt{1},\mathtt{2}) + \tfrac{1}{6}\mathtt{1} \sqcup\!\sqcup \mathtt{2} \sqcup\!\sqcup \mathtt{3}\end{aligned}$$

To formulate the problem algebraically, let $\mathscr{A} \subset T(\mathbb{R}^d)$ be the smallest linear space containing the letters $\mathtt{1}, \ldots, \mathtt{d}$ that is closed under the (bilinear, non-associative) operation area. The question then becomes:

Is $\mathscr{A}$ a generating set for the shuffle algebra $T(\mathbb{R}^d)$?

The affirmative answer to this question is given in this chapter.

What we really have in mind here is a two-stage numerically-stable procedure for calculating the signature of a physical path. In the first stage one calculates areas, areas of areas and so forth, possibly using an *analog physical apparatus*.[2] The second stage uses these measurements, say on a *digital computer*, and computes polynomial expressions in these.

[2] One physical device that has historically been used to measure area is a planimeter (see e.g. [FS07]). In general, this is related to nonholonomic control, see the textbook [Blo15] and also compare the somewhat similar situation of the rolling ball without slipping at the end of [BD15, Section 3], thanks again to Joscha Diehl for pointing out these references.

The rest of the chapter is structured as follows. In the next subsection we fix notation. In Section 4.2 we revisit results by Rocha from [Roc03a], [Roc03b] and [Roc05] in purely algebraic terms. The outcome of this is a formula for the Dynkin operator applied to the signature. This makes the area operator appear naturally. Together with Section 4.4 this will prove the generating property of areas-of-areas.

For completeness, we show in Section 4.3 how to express coordinates of the first kind using only areas-of-areas. Again, this is basically a purely algebraic revisiting of results by Rocha, in which we also correct some of the expressions he gives.

In Section 4.4 we state a general condition for a set of polynomials to be (free) generators of the shuffle algebra $T(\mathbb{R}^d)$. We then show how a couple of well-known generators fall into this formulation and, how using Section 4.2 (or 4.3), the generating property of areas-of-areas is established in Theorem 4.4.8.

Apart from its geometric interpretation, the area operation possesses some interesting properties. Some of them we present in Section 4.5, where it is shown that it is nicely compatible with discrete integration as well as stochastic integration. In Section 4.6 we collect some results on the linear span generated by the area operator, as it is of interest in its own right.

Acknowledgements.

T.L. is supported by the EPSRC under the program grant EP/S026347/1 and by the Alan Turing Institute under the EPSRC grant EP/N510129/1. R.P. is supported by European Research Council through CoG-683164 and was affiliated to Max Planck Institute for Mathematics in the Sciences, Leipzig, in autumn 2018.

T.L. and J.D. would like to thank Danyu Yang for many insightful discussions and for realizing that it is natural to think of shuffle algebra as a space of operators or sensors that take paths to scalar paths and the consequent insight that it is completely natural to interpret the half-shuffles, and area etc. as binary path operations, and so connect the structures of dendriform algebras with the tensor/shuffle algebra structures in this way.

We would like to thank Cristopher Salvi for extended discussions, valuable presentations and new insights on the area operator, in particular its Jacobian bracketing (which helped shape the interpretation as signed volume) and the interplay with the shuffle product, and on coordinates of the second kind. See [Sal21, Chapter 4], in particular Salvi's "shuffle-pullout" [Sal21, Lemma 4.3.0.2] and "area-Jacobi" [Sal21, Lemma 4.3.0.3] identities.

Tensor algebra and tree computations for this project have been done in `python` and `sage`, where besides standard packages and further custom code by the authors the `python` package `free_lie_algebra.py` [Rei21], which implements a lot of the definitions in [Reu93], has been of central use.

4.1.1 Notation

Denote by $T((\mathbb{R}^d))$ the space of formal *infinite* linear combinations of words in the letters $\mathtt{1},\ldots,\mathtt{d}$. Equip it with the **concatenation product** $\bullet$ (often we write $b \bullet b' = bb'$).

Denote by $T(\mathbb{R}^d)$ its dual, the space of *finite* linear combinations of words. Equip it with the **shuffle product** $\sqcup\!\sqcup$. It decomposes as

$$a \sqcup\!\sqcup a' = a \succ a' + a' \succ a,$$

where $\succ$ is the **half-shuffle**. The half-shuffle is defined on words $a = a_1 \cdots a_m, b = b_1 \cdots b_n$, where b is not the empty word, as

$$a \succ b = (a \sqcup\!\sqcup b_1 \cdots b_{n-1}) \bullet b_n.$$

The dual pairing is written for $a \in T(\mathbb{R}^d), b \in T((\mathbb{R}^d))$ as

$$\Big\langle a, b \Big\rangle.$$

Denote the grouplike elements of $T((\mathbb{R}^d))$ by $\mathcal{G}$.Denote the primitive elements, or Lie elements, of $T((\mathbb{R}^d))$, i.e. the vector space of formal series built from the free Lie algebra, by $\mathfrak{g}((\mathbb{R}^d))$.

Denote by $\mathsf{proj}_n, \mathsf{proj}_{\geq n}$, etc, the projection on $T((\mathbb{R}^d))$ to level n, to levels larger equal to n, ... We write $T_n((\mathbb{R}^d)) = \mathsf{proj}_n T((\mathbb{R}^d)), T_{\geq n}((\mathbb{R}^d)) = \mathsf{proj}_{\geq n} T((\mathbb{R}^d))$, etc. Denote the empty word by e.

Denote by $T\langle \mathcal{R}^d \rangle$ the free tensor algebra over d generators with coefficients in the ring $\mathcal{R}$, where $\mathcal{R} := (T(\mathbb{R}^d), \sqcup\!\sqcup)$, and by $T\langle\langle \mathcal{R}^d \rangle\rangle$ the corresponding space of tensor series. We then canonically have $T\langle \mathcal{R}^d \rangle \subsetneq T\langle\langle \mathcal{R}^d \rangle\rangle$, and identify the $\mathcal{R}$-algebra $T\langle\langle \mathcal{R}^d \rangle\rangle$ with

$$\mathfrak{W} := \prod_{n=1}^{\infty} T(\mathbb{R}^d) \otimes T_n(\mathbb{R}^d), \tag{4.1}$$

where we use the shuffle product on the left and the concatenation product on the right. We denote the product on both $T\langle\langle \mathcal{R}^d \rangle\rangle$ and $\mathfrak{W}$, which are isomorphic as $\mathbb{R}$-algebras, by $\blacksquare$. The $\mathcal{R}$-subalgebra $(T\langle \mathcal{R}^d \rangle, \blacksquare)$ is then $\mathbb{R}$-algebra-isomorphic to $(T(\mathbb{R}^d) \otimes T(\mathbb{R}^d), \blacksquare)$.

We use the usual grading on $T\langle\langle \mathcal{R}^d \rangle\rangle$, that is in the representation (4.1), for a, b words, $|a \otimes b| := |b|$. Then, the projection proj_n makes also sense on $\mathfrak{W}$.

We furthermore introduce an $\mathcal{R}$-linear coproduct on $T\langle\langle \mathcal{R}^d \rangle\rangle$, which maps to the graded completion of the $\mathcal{R}$-module tensor product $\boxtimes$:

$$\Delta_{\underline{\sqcup\!\sqcup}} : T\langle\langle \mathcal{R}^d \rangle\rangle \to T\langle\langle \mathcal{R}^d \rangle\rangle \,\hat{\boxtimes}\, T\langle\langle \mathcal{R}^d \rangle\rangle := \prod_{m,n=1}^{\infty} \mathsf{proj}_m T\langle\langle \mathcal{R}^d \rangle\rangle \boxtimes \mathsf{proj}_n T\langle\langle \mathcal{R}^d \rangle\rangle,$$

where the unshuffle coproduct on $T\langle\langle \mathcal{R}^d \rangle\rangle$ is defined via the usual unshuffle coproduct as

$$\Delta_{\underline{\sqcup\!\sqcup}} \Big(\sum_w a_w \underline{w} \Big) := \sum_w a_w \Delta_{\underline{\sqcup\!\sqcup}} \underline{w} := \sum_w a_w \sum_{(w)}^{\sqcup\!\sqcup} \underline{w_1} \boxtimes \underline{w_2},$$

where the last Sweedler summation is well defined by the unshuffle coproduct on $T(\mathbb{R}^d)$ because there is a unique $\mathbb{R}$-linear map $T(\mathbb{R}^d) \otimes T(\mathbb{R}^d) \to T\langle\langle \mathcal{R}^d \rangle\rangle \hat{\boxtimes} T\langle\langle \mathcal{R}^d \rangle\rangle$ characterized by sending each tensor pair of words $w \otimes v$ to $\underline{w} \boxtimes \underline{v}$ (which is however non-surjective). We have the isomorphism

$$\prod_{m,n=1}^{\infty} T(\mathbb{R}^d) \otimes T_m(\mathbb{R}^d) \otimes T_n(\mathbb{R}^d) \cong T\langle\langle \mathcal{R}^d \rangle\rangle \,\hat{\boxtimes}\, T\langle\langle \mathcal{R}^d \rangle\rangle$$

as $\mathbb{R}$ vector spaces given by the map

$$\sum_{w,v} a_{w,v} \otimes w \otimes v \mapsto \sum_{w,v} a_{w,v} \, (\underline{w} \boxtimes \underline{v}) = \sum_{w,v} (a_{w,v} \underline{w}) \boxtimes \underline{v} = \sum_{w,v} \underline{w} \boxtimes (a_{w,v} \underline{v}).$$

The unshuffle coproduct on $T\langle\langle\mathcal{R}^d\rangle\rangle$ is an $\mathcal{R}$-algebra homomorphism as a consequence of the homomorphism property of the usual unshuffle coproduct, as for words w, v we have

$$\begin{aligned}\Delta_{\underline{\sqcup\!\sqcup}}(p\underline{w} \blacksquare q\underline{v}) &= (p \sqcup\!\sqcup q)\Delta_{\underline{\sqcup\!\sqcup}}\, \underline{w \bullet v} = (p \sqcup\!\sqcup q) \sum_{(w),(v)}^{\sqcup\!\sqcup} \underline{w_1 \bullet v_1} \boxtimes \underline{w_2 \bullet v_2} \\ &= \sum_{(w),(v)}^{\sqcup\!\sqcup} (p\underline{w_1} \blacksquare q\underline{w_2}) \boxtimes (\underline{w_2} \blacksquare \underline{w_1}) = \Big(\sum_{(w)}^{\sqcup\!\sqcup} p\underline{w_1} \boxtimes \underline{w_2}\Big) \tilde{\blacksquare} \Big(\sum_{(v)}^{\sqcup\!\sqcup} q\underline{v_1} \boxtimes \underline{v_2}\Big) \\ &= (\Delta_{\underline{\sqcup\!\sqcup}} p\underline{w}) \,\tilde{\blacksquare}\, (\Delta_{\underline{\sqcup\!\sqcup}} q\underline{v}), \end{aligned} \tag{4.2}$$

where

$$\Big(\sum_{w_1,v_1} a_{w_1,v_1}\underline{w_1} \boxtimes \underline{v_1}\Big) \tilde{\blacksquare} \Big(\sum_{w_2,v_2} b_{w_2,v_2}\underline{w_2} \boxtimes \underline{v_2}\Big) := \sum_{w_1,v_1,w_2,v_2} (a_{w_1,v_1} \sqcup\!\sqcup b_{w_2,v_2})(\underline{w_1} \blacksquare \underline{w_2}) \boxtimes (\underline{v_1} \blacksquare \underline{v_2})$$

is the usual induced product on the tensor product. When restricting to $T\langle\mathcal{R}^d\rangle$, we have $\Delta_{\underline{\sqcup\!\sqcup}} : T\langle\mathcal{R}^d\rangle \to T\langle\mathcal{R}^d\rangle \boxtimes T\langle\mathcal{R}^d\rangle$ and the other compatibility relations of a Hopf algebra are checked along the same lines, so we indeed get an $\mathcal{R}$-Hopf algebra $(T\langle\mathcal{R}^d\rangle, \blacksquare, \Delta_{\underline{\sqcup\!\sqcup}}, \underline{\alpha})$, a Hopf algebra in the category of $\mathcal{R}$-modules, with antipode

$$\underline{\alpha}\Big(\sum_w a_w \underline{w}\Big) = \sum_w (-1)^{|w|} a_w \underline{\overleftarrow{w}},$$

where $\overleftarrow{w}$ is w written backwards.

Now, since we have the homomorphism property of the unshuffle $\Delta_{\underline{\sqcup\!\sqcup}}$ on $T\langle\langle\mathcal{R}^d\rangle\rangle$ according to Equation (4.2), and furthermore

$$\Delta_{\underline{\sqcup\!\sqcup}}(\underline{\mathtt{i}}) = \underline{e} \boxtimes \underline{\mathtt{i}} + \underline{\mathtt{i}} \boxtimes \underline{e}$$

for any letter $\underline{\mathtt{i}}$ in $T\langle\langle\mathcal{R}^d\rangle\rangle$, our $\Delta_{\underline{\sqcup\!\sqcup}}$ is exactly the coproduct from [Reu93, Section 1.3, page 19] for the choice of K as the unital commutative ring $\mathcal{R}$ with characteristic zero. Thus, we may apply all the theory in Reutenauer's book valid for the general setting of an unital commutative ring of characteristic zero to $T\langle\langle\mathcal{R}^d\rangle\rangle$. In particular, we get that the group [Reu93, Corollary 3.3]

$$\underline{G} = \{g \in T\langle\langle\mathcal{R}^d\rangle\rangle | \Delta_{\underline{\sqcup\!\sqcup}} g = g \boxtimes g,\ g \neq 0\}$$

with product $\blacksquare$ and the $\mathcal{R}$-Lie-algebra

$$\mathfrak{g}\langle\langle\mathcal{R}^d\rangle\rangle = \{x \in T\langle\langle\mathcal{R}^d\rangle\rangle | \Delta_{\underline{\sqcup\!\sqcup}} x = \underline{e} \boxtimes x + x \boxtimes \underline{e}\}$$

with Lie bracket $[x, y]_\blacksquare := x \blacksquare y - y \blacksquare x$ are in a one-to-one correspondence [Reu93, Theorem 3.1 and 3.2] via the exponential map [Reu93, Equation (3.1.2)]

$$\exp_\blacksquare : \mathfrak{g}\langle\langle\mathcal{R}^d\rangle\rangle \to \underline{G}, \quad \exp_\blacksquare(x) = \underline{e} + \sum_{n=1}^{\infty} \frac{x^{\blacksquare n}}{n!}$$

with inverse the logarithm [Reu93, Equation (3.1.1)]

$$\log_\blacksquare : \underline{G} \to \mathfrak{g}\langle\langle\mathcal{R}^d\rangle\rangle, \quad \log_\blacksquare(g) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(g - \underline{e})^{\blacksquare n}}{n}.$$

Analogous to G and $\mathfrak{g}((\mathbb{R}^d))$, we call the elements of $\underline{G}$ grouplike and the elements of $\mathfrak{g}\langle\langle\mathcal{R}^d\rangle\rangle$ primitive.

Note however that $(T\langle\mathcal{R}^d\rangle, \blacksquare, \Delta_{\sqcup\!\sqcup})$ does not form a $\mathbb{R}$ Hopf algebra.

Fixing $x \in T((\mathbb{R}^d))$, define for any $F = \sum_w a_w \otimes w \in \mathfrak{W}$, where $a_w \in T(\mathbb{R}^d)$ for all words w,

$$\mathsf{eval}_x(F) := \sum_w \langle x, a_w\rangle \; w \in T((\mathbb{R}^d)).$$

This operation now forms an associative algebra isomorphism from $(\mathfrak{W}, \blacksquare)$ to $(\mathcal{L}(T((\mathbb{R}^d)), T((\mathbb{R}^d))), *)$, where $\mathcal{L}(T((\mathbb{R}^d)), T((\mathbb{R}^d)))$ denotes the linear maps from $T((\mathbb{R}^d))$ to $T((\mathbb{R}^d))$ which are continuous in the product topology and $*$ denotes the convolution product of the Hopf algebra $(T(\mathbb{R}^d), \bullet, \Delta_{\sqcup\!\sqcup})$ extended to $T((\mathbb{R}^d))$. Indeed, for any $F, G \in \mathfrak{W}$, we have $\mathsf{eval}(F), \mathsf{eval}(G) \in \mathcal{L}(T((\mathbb{R}^d)), T((\mathbb{R}^d)))$ by definition with

$$\mathsf{eval}(F \blacksquare G) = \mathsf{eval}(F) * \mathsf{eval}(G),$$

since for $F = \sum_w a_w \otimes w$, $G = \sum_{w'} b_{w'} \otimes w'$, $a_w, b_w \in T(\mathbb{R}^d)$ for all words w, and $x \in T((\mathbb{R}^d))$,

$$\begin{aligned}
\mathsf{eval}_x(F \blacksquare G) &= \sum_{w,w'} \langle x, a_w \sqcup\!\sqcup b_{w'}\rangle \, w \bullet w' = \sum_{w,w'} \langle \Delta_{\sqcup\!\sqcup} x, a_w \otimes b_{w'}\rangle \, w \bullet w' \\
&= \sum_{(x)}^{\sqcup\!\sqcup} \sum_w \langle x_1, a_w\rangle \, w \bullet \sum_{w'} \langle x_2, b_{w'}\rangle w' = \sum_{(x)}^{\sqcup\!\sqcup} \mathsf{eval}_{x_1}(F) \bullet \mathsf{eval}_{x_2}(G) \\
&= \mathsf{conc}(\mathsf{eval}(F) \otimes \mathsf{eval}(G))\Delta_{\sqcup\!\sqcup} x = (\mathsf{eval}(F) * \mathsf{eval}(G))(x),
\end{aligned}$$

where $\sum_{(x)}^{\sqcup\!\sqcup} x_1 \otimes x_2 := \Delta_{\sqcup\!\sqcup} x$ is Sweedler's notation and $\mathsf{conc} : T((\mathbb{R}^d)) \hat{\otimes} T((\mathbb{R}^d)) \to T((\mathbb{R}^d))$ is the continuous linear map corresponding to the bilinear map $\bullet$.

Likewise, for arbitrary $y \in T(\mathbb{R}^d)$ and $F = \sum_w a_w \otimes w \in \mathfrak{W}$, $a_w \in T(\mathbb{R}^d)$, we define

$$\mathsf{coeval}^y(F) := \sum_w \langle w, y\rangle a_w \in T(\mathbb{R}^d).$$

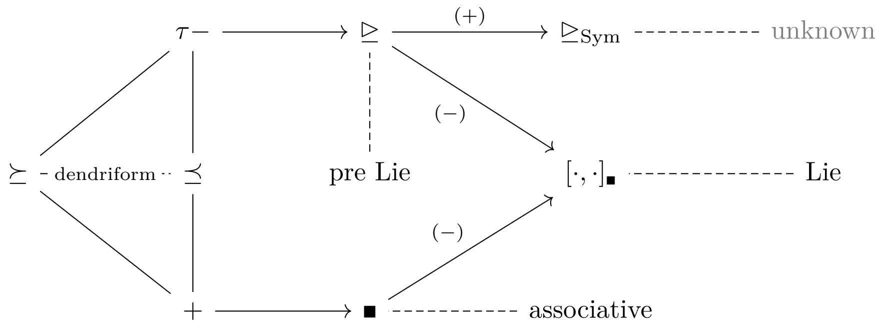
Then, coeval forms an isomorphism from $(\mathfrak{W}, \blacksquare)$ to $(\mathrm{L}(T(\mathbb{R}^d), T(\mathbb{R}^d)), \star)$, where $\mathrm{L}(T(\mathbb{R}^d), T(\mathbb{R}^d))$ denotes all linear maps from $T(\mathbb{R}^d)$ to itself and $\star$ is the convolution product of the Hopf algebra $(T(\mathbb{R}^d), \sqcup\!\sqcup, \Delta_\bullet)$.

We refer to [Reu93] for more details on all of this, except for the half-shuffle, for which a nice entry point to the literature is for example [FP13].

4.1.2 Objectives

4.1.2.1 Revisiting the work of Rocha on coordinates of the first kind

In Sections 4.2 and 4.3, we show how the area operation appears naturally in a purely algebraic formulation of the work of Rocha on coordinates of the first kind. What we may take over from Rocha here is a very interesting network of bilinear operations on $\mathfrak{W}$ refining the basic $\blacksquare$ product. It is based on a dendriform structure, as the following diagram and definition show:



For $A = p \otimes q$, $B = p' \otimes q'$,

$$\begin{aligned}
A \succeq B &:= (p \succ p') \otimes (q \bullet q'),\\
A \preceq B &:= (p' \succ p) \otimes (q \bullet q'),\\
A \blacksquare B &:= A \succeq B + A \preceq B = (p \sqcup\!\sqcup p') \otimes (q \bullet q'),\\
A \trianglerighteq B &:= A \succeq B - B \preceq A = (p \succ p') \otimes [q, q'],\\
A \trianglerighteq_{\mathrm{Sym}} B &:= A \trianglerighteq B + B \trianglerighteq A = A \succeq B + B \succeq A - A \preceq B - B \preceq A = \mathsf{area}(p, p') \otimes [q, q'],\\
[A, B]_{\blacksquare} &:= A \blacksquare B - B \blacksquare A = A \trianglerighteq B - B \trianglerighteq A = A \succeq B + A \preceq B - B \succeq A - B \preceq A\\
&= (p \sqcup\!\sqcup p') \otimes [q, q'].
\end{aligned}$$

In fact, one could describe this network for any dendriform structure, with Rocha's and our work offering a promising use case for talking about all of these operations together, while this system of operations *without* $\trianglerighteq_{\mathrm{Sym}}$ has been explored before e.g. in [EM09]. The symmetrized pre-Lie operation $\trianglerighteq_{\mathrm{Sym}}$ stays the most mysterious also to us, we may only point to the discovery of Bergeron and Loday in [BL11, Theorem 3.1] that the symmetrization of pre-Lie does not in general satisfy any further identities except non-associative commutativity, though since the pre-Lie product $\trianglerighteq$ certainly isn't free, we expect some kind of relations for $\trianglerighteq_{\mathrm{Sym}}$ also, but this is still a question of future work.

With the area operation forming the left part of the symmetrized pre-Lie operation $\trianglerighteq_{\mathrm{Sym}}$, we obtain our main argument to show that the set of all areas of areas forms a shuffle generating set, albeit not a minimal one.

4.1.2.2 Areas of areas and further shuffle generating sets

With our main focus being shuffle-generating sets in terms of areas of areas, in Section 4.4 we first give a general criterion Lemma 4.4.2 for (homogeneous) subsets to form a shuffle generating set (resp. a free shuffle generating set), the condition being that the set contains (resp. forms) a dual basis to some basis of the free Lie algebra $\mathfrak{g}(\mathbb{R}^d) \subsetneq T(\mathbb{R}^d)$. While our actual hands-on proof is based on the characterization of the annihilator of the free Lie algebra which we cite from [Reu93, Theorem 3.1 (iv)], we sketch a more abstract argument in Remark 4.4.3 related to the Cartier-Milnor-Moore theorem.

We continue by illustrating how our statement can be applied to some known shuffle generating sets, as well as to the image of ρ (the dual of the Dynkin map, Section 4.2), which concludes one of our proofs that areas-of-areas generate the shuffle algebra.

4.1.2.3 The area Tortkara algebra

We study the smallest linear subspace $\mathscr{A} \subsetneq T(\mathbb{R}^d)$ closed under the area operation and containing the letters in Section 4.6. Thanks to the work by Dzhumadil'daev, Ismailov and Mashurov, we can use the categorial framework of *Tortkara algebras*, where the objects are characterized as vector spaces with a bilinear antisymmetric operation which furthermore satisfies the *Tortkara identity*, and the morphisms are homomorphisms of the bilinear operations as usual. Also thanks to [DIM19, Section 3], we have a simple linear basis of $\mathscr{A}$ in terms of linear combinations of words, see Lemma 4.6.2. We continue by a very important conjecture that the left bracketings of the area operation yield another basis of $\mathscr{A}$, which was shown for dimension two in both [DIM19, Section 5] and [Rei19, Section 3.2, Theorem 31]. The rest of the section is dedicated to some interesting observations we made while, so far unsuccessfully, trying to prove that conjecture for any dimension.

4.1.2.4 Applications and characterizations

In Section 4.5 we are connecting the purely algebraic considerations of this chapter with the world of (deterministic and probabilistic) path spaces and iterated-integrals signatures on these path spaces, as they have been the motivation for this work to begin with. What we are generally looking at are characterizations of the area Tortkara algebra $\mathscr{A}$ in terms of special properties for given path spaces, like the space of piecewise linear paths (Subsection 4.5.1). The case of piecewise linear paths promises in fact to develop into the main application of the study of areas of areas. They form the most common discretization of general continuous paths that one works with when actually computing iterated-integrals signature numerically, in machine learning for example. It turns out that for piecewise linear paths, the computation of discrete areas is much simpler and better behaved than the computation of discrete integrals.

However, besides the discrete deterministic setting, the study of signatures has, since Lyons' theory of rough paths, been intimately related with stochastic analysis, and we observe how areas of areas preserve the martingale property central in stochastic analysis, while general iterated Stratonovich integrals fail to do so.

4.1.3 Summerized proof of the main result

We give three proofs of the fact that areas-of-areas shuffle generate the tensor algebra, which is Theorem 4.4.8.

4.1.3.1 Via Λ and coordinates of the first kind

Using $Dw = \sum_{uv=w,\,|u|\geq 1} \rho(u) \sqcup\!\sqcup\, v$, $\mathrm{ad}_P = \mathrm{Ad}_P$ for any Lie polynomial P and $\sum_w w \otimes r(w) = \sum_w \rho(w) \otimes w$, we showed $(\underline{D} - \mathrm{id})R = R \trianglerighteq R$ (Lemma 4.2.6).

We then have $R \in \mathfrak{R}$ due to $R_1 = \sum_{\mathtt{i}=1}^{\mathtt{d}} \mathtt{i} \otimes \mathtt{i}$ and $R_n = \frac{1}{2(n-1)} \sum_{l=1}^{n-1} R_l \trianglerighteq_{\mathrm{Sym}} R_{n-l}$ (Corollary 4.2.8).

Using $R = \underline{r} \exp_{\blacksquare} \Lambda$ and Baker's identity for $\underline{r}$ (Lemma 4.3.3), we show that Λ_n is a linear combination of R_n and ■-Lie-bracketings of lower order Λ_i. Thus $\Lambda \in \mathfrak{L}$, and together with the fact $\zeta_h = \mathsf{coeval}^{P_h^*}(\Lambda)$ following from the definition of Λ we conclude that each ζ_h is a linear combination of shuffles of areas of areas.

4.1.3.2 Via R and ρ

From $R \in \mathfrak{R}$ we conclude that the image of ρ lies in $\mathscr{A}$, since $\rho(v) = \mathsf{coeval}^v(R)$ for any word v. Since the image of ρ shuffle generates the shuffle algebra, Point 3.2 in Corollary 4.4.5, so does $\mathscr{A}$.

4.1.3.3 Via [DIM19] and ρ

Via a combinatorial expression for $\rho(\mathtt{12}\dots\mathtt{d})$ for any d (Proposition 4.6.13) we conclude (Corollary 4.6.14) that the image of ρ is a subspace of

$$\operatorname{span}_{\mathbb{R}}\{\mathtt{i} : \mathtt{i} \text{ a letter }\} \oplus \operatorname{span}_{\mathbb{R}}\{w(\mathtt{ij} - \mathtt{ji}) : w \text{ a word}, \mathtt{i}, \mathtt{j} \text{ letters}\},$$

which is nothing but $\mathscr{A}$, according to [DIM19, Theorem 2.1]. Again, since ρ shuffle generates the shuffle algebra, so do areas of areas.

4.2 The Dynkin operator

We recall the linear maps $r, D : T((\mathbb{R}^d)) \to T((\mathbb{R}^d))$ from [Reu93, Section 1, pages 19-20]. The linear right-bracketing map or **Dynkin operator** r is given on a word $w = \mathtt{l}_1 \cdots \mathtt{l}_n$ as

$$r(\mathtt{l}_1 \cdots \mathtt{l}_n) := [\mathtt{l}_1, [\mathtt{l}_2, \dots [\mathtt{l}_{n-1}, \mathtt{l}_n]]], \tag{4.3}$$

with $r(e) = 0$ and $r(\mathtt{i}) = \mathtt{i}$ for any letter $\mathtt{i}$. The map D (for *derivation*) is given on a word w as

$$D(w) := |w|w,$$

where $|w|$ is the length of the word. On $T_{\geq 1}((\mathbb{R}^d))$, D is invertible with inverse $D^{-1}(w) = \frac{1}{|w|}w$.

Remark 4.2.1. 1. The seemingly simple Dynkin operator r has found several applications. It for example characterizes Lie elements of $T((\mathbb{R}^d))$ [Reu93, Theorem 3.1 (vi)]: $x \in T((\mathbb{R}^d))$ is a Lie series if and only if $\langle e, x\rangle = 0$ and $r(x) = D(x)$. See also [PR02b, Section 3], [Gar90, Section 1, 2 and 4] and references therein. In the analysis of numerical schemes it is used for example in [LM13, Section 2.4].

2. Truncated at a fixed level, the grouplike elements / signatures of tree-reduced paths, form a Lie group. The Dynkin operator r is a logarithmic derivative, i.e. the derivative pulled-back to the tangent space at the identity, of an endomorphism of this group in the following sense (see [MP13] for more on this). Let δ_ϵ be the dilation operator, i.e. the operation on tensors which multiplies each level m by ϵ^m, which corresponds to dilating or scaling a path by the factor ϵ. For $g \in \mathcal{G}$, let $g^\epsilon := \delta_\epsilon g$. Then (cf. [MP13, Equation on the bottom of page 906, Section 3.2])

$$\begin{aligned}\Big(\frac{d}{d\epsilon}g^\epsilon\Big) \bullet (g^\epsilon)^{-1} = \Big(\frac{d}{d\epsilon}g^\epsilon\Big) \bullet \alpha[g^\epsilon] = \big(\tfrac{1}{\epsilon}D[g^\epsilon]\big) \bullet \alpha[g^\epsilon] &= \tfrac{1}{\epsilon}\big(\mathsf{conc} \circ (D \otimes \alpha)\big)[g^\epsilon \otimes g^\epsilon] \\ &= \tfrac{1}{\epsilon}\big(\mathsf{conc} \circ (D \otimes \alpha) \circ \Delta_{⧢}\big)[g^\epsilon] \qquad \text{as } g^\epsilon \in G, \text{ [Reu93, Theorem 3.2]} \\ &= \tfrac{1}{\epsilon} r[g^\epsilon], \qquad \text{see [Reu93, Lemma 1.5 or Section 1, page 32]}\end{aligned}$$

where α is the antipode extended from $T(\mathbb{R}^d)$ to $T((\mathbb{R}^d))$ (which is the inverse in the Lie group, and corresponds to reversing a path), $\otimes$ is the external tensor product, conc is the linear map taking $a \otimes b$ to $a \bullet b$, and $\Delta_{⧢}$ is the unshuffle coproduct, which in [Reu93] is denoted with δ.

Let $\underline{r}$, $\underline{D}$, $\underline{D}^{-1}$ act on $\mathfrak{W}$ by letting r, D, D^{-1} act on the right side of the tensor, i.e.

$$\begin{aligned}\underline{r}(a \otimes b) &:= a \otimes r(b)\\ \underline{D}(a \otimes b) &:= a \otimes D(b)\\ \underline{D}^{-1}(a \otimes b) &:= a \otimes D^{-1}(b).\end{aligned}$$

Define[3]

$$\begin{aligned}S &:= \sum_w w \otimes w\\ R &:= \underline{r}(S) = \sum_w w \otimes r(w) = \sum_v \rho(v) \otimes v. \end{aligned} \tag{4.4}$$

Both are elements of $\mathfrak{W}$. The last equality implicitly defines ρ. There also exists a recursive definition given by $\rho(e) = 0$, $\rho(\mathtt{i}) = \mathtt{i}$ for any letter $\mathtt{i}$ and

$$\rho(\mathtt{i}w\mathtt{j}) = \mathtt{i}\rho(w\mathtt{j}) - \mathtt{j}\rho(\mathtt{i}w) \tag{4.5}$$

for any (empty or non-empty) word w and letters $\mathtt{i}, \mathtt{j}$, see [Reu93, Equation (1.5.10)]. Based on this recursion, we derive an expansion of ρ via an action of elements of the symmetric group algebra in Proposition 4.6.13. We repeat that $r(e) = \rho(e) = 0$, so the sum in (4.4) is actually only taken over words of strictly positive length.

We record the following for future use [Reu93, Theorem 1.12]: For any word w[4]

$$Dw = \sum_{uv=w} \rho(u) \sqcup\!\sqcup\, v = \sum_{uv=w, |u|\geq 1} \rho(u) \sqcup\!\sqcup\, v. \tag{4.6}$$

Note that this yields yet another recursive definition of ρ:

$$\rho(e) = 0, \quad \rho(w) = |w|w - \sum_{\substack{uv=w\\ |u|,|v|\geq 1}} \rho(u) \sqcup\!\sqcup\, v,$$

where w is an arbitrary non-empty word.

Proposition 4.2.2. *The map $r : \mathcal{G} \to \mathfrak{g}((\mathbb{R}^d))$ is invertible. To be specific, define for $x \in T_{\geq 1}((\mathbb{R}^d))$ the linear map*

$$\begin{aligned}A_x : T((\mathbb{R}^d)) &\to T((\mathbb{R}^d))\\ z &\mapsto D^{-1}(xz).\end{aligned}$$

Then for $x \in \mathfrak{g}((\mathbb{R}^d))$

$$\begin{aligned}r^{-1}[x] &= \sum_{\ell\geq 0} A_x^\ell e \\ &= e + D^{-1}(x) + D^{-1}(xD^{-1}(x)) + D^{-1}(xD^{-1}(xD^{-1}(x))) + \dots\end{aligned} \tag{4.7}$$

[3] From now on, if we sum over a variable with no given index set, we sum over all words in the alphabet $\mathtt{1}, \dots, \mathtt{d}$, including the empty word e.

[4] If $|w| = 0$ then both sides are equal to zero.

Equivalently, with $\underline{A}_R z := \underline{D}^{-1}[R \blacksquare z]$*,*

$$\begin{aligned} S &= \sum_{\ell \geq 0} (\underline{A}_R)^\ell (e \otimes e) \\ &= e \otimes e + \underline{D}^{-1}[R] + \underline{D}^{-1}[R \blacksquare \underline{D}^{-1}[R]] + \underline{D}^{-1}[R \blacksquare \underline{D}^{-1}[R \blacksquare \underline{D}^{-1}[R]]] + \dots, \end{aligned} \tag{4.8}$$

Remark 4.2.3. Compare [EGP07, Theorem 4.1] for a statement in a more general setting.

Proof of Theorem 4.2.2. The claimed equivalence is shown as follows. For $t \in \mathfrak{W}$, with zero coefficient for $e \otimes e$,

$$\mathsf{eval}_g(\underline{D}^{-1}(t)) = D^{-1}(\mathsf{eval}_g(t)).$$

Hence

$$\begin{aligned} g &= e + D^{-1}(r(g)) + D^{-1}(r(g)D^{-1}(r(g))) + \dots \quad \forall g \in G \\ &\Leftrightarrow \\ \mathsf{eval}_g(S) &= \mathsf{eval}_g(e) + D^{-1}(\mathsf{eval}_g(R)) + D^{-1}(\mathsf{eval}_g(R) D^{-1}(\mathsf{eval}_g(R))) + \dots \quad \forall g \in G \\ &= \mathsf{eval}_g(e) + \mathsf{eval}_g(\underline{D}^{-1}(R)) + \mathsf{eval}_g(\underline{D}^{-1}(R \blacksquare \underline{D}^{-1}(R))) + \dots \\ &\Leftrightarrow \\ S &= e + \underline{D}^{-1}(R) + \underline{D}^{-1}(R \blacksquare \underline{D}^{-1}(R)) + \dots, \end{aligned}$$

where we used the homomorphism property of eval_g and the fact that grouplike elements linearly span $T((\mathbb{R}^d))$ projectively (i.e. truncated, at level n, grouplike elements linearly span $T_{\leq n}((\mathbb{R}^d))$).

We now show (4.7). Write $x := r[g] = D[g] \bullet g^{-1}$ (compare Remark 4.2.1.2). Then

$$g = e + D^{-1}[x \bullet g],$$

i.e.

$$g = e + A_x g. \tag{4.9}$$

Now since x does not contain a component in the empty word, this actually amounts to a recursive formula,

$$\mathsf{proj}_0 g = e, \quad \mathsf{proj}_n g = \sum_{m=1}^{n} A_{\mathsf{proj}_m x}(\mathsf{proj}_{n-m} g), \quad n \geq 1$$

and thus Equation (4.9) has a unique solution. Hence

$$g = \sum_{\ell \geq 0} A_x^\ell e,$$

since the series converges due to being a finite sum for each homogeneous component and obviously provides a solution for Equation (4.9).

This shows that (4.7) gives a left-inverse.

It is also a right inverse. Indeed, first note that for $x \in \mathfrak{g}((\mathbb{R}^d))$ and $n \geq 2$ we have $r[A_x^n e] = 0$. For $n = 2$, using Lemma 4.3.3, this follows from

$$\begin{aligned} r[A_x^2 e] &= r[D^{-1}(xD^{-1}(x))] = D^{-1}(r[xD^{-1}(x)]) = D^{-1}(r[r[D^{-1}x]D^{-1}(x)]) \\ &= D^{-1}([r[D^{-1}x], r[D^{-1}(x)]]) = 0. \end{aligned}$$

Assume it is true for A_x^{n-1}, then

$$\begin{aligned} r[A_x^n e] &= r[D^{-1}(xA_x^{n-1}e)] = D^{-1}(r[xA_x^{n-1}e]) = D^{-1}(r[r[D^{-1}x]A_x^{n-1}e]) \\ &= D^{-1}([r[D^{-1}x], r[A_x^{n-1}e]]) = 0. \end{aligned}$$

Hence

$$r[e + D^{-1}(x) + D^{-1}(xD^{-1}(x)) + \ldots] = x,$$

so that the Lemma indeed provides a right inverse. □

Definition 4.2.4. *Define the following product on* $\mathfrak{W}$*,*

$$(p \otimes q) \trianglerighteq (p' \otimes q') := (p \succ p') \otimes [q, q'],$$

where $[.,.]$ *is the Lie bracket in* $T((\mathbb{R}^d))$ *and* $\succ$ *is the half-shuffle in* $T(\mathbb{R}^d)$*.*

Remark 4.2.5. This product is *pre-Lie*, as the tensor product of a Zinbiel algebra and a Lie algebra is always a pre-Lie algebra (this is shown in Rocha's thesis as [Roc03a, Proposition 4.13 and Corollary 4.14], though there the terminology 'chronological algebra' is used to mean what we call pre-Lie algebra), although we will not use this fact. It comes from the dendriform structure

$$\begin{aligned} (p \otimes q) \succeq (p' \otimes q') &:= (p \succ p') \otimes qq' \\ (p \otimes q) \preceq (p' \otimes q') &:= (p' \succ p) \otimes qq', \end{aligned}$$

i.e. $x \trianglerighteq y = x \succeq y - y \preceq x$. Indeed, the operations $\succeq$ and $\preceq$ together satisfy the three dendriform identities (e.g. [Lod01, Definition 5.1] or [EM09, Equations (8)-(10)]), which is a straightforward consequence of the Zinbiel identity of the halfshuffle and the associativity of the concatenation,

$$\begin{aligned} (A \preceq B) \preceq C &= (p_3 \succ (p_2 \succ p_1)) \otimes q_1q_2q_3 = ((p_3 \succ p_2) \succ p_1) \otimes q_1q_2q_3 + ((p_2 \succ p_3) \succ p_1) \otimes q_1q_2q_3 \\ &= A \preceq (B \preceq C) + A \preceq (C \preceq B), \\ A \succeq (B \succeq C) &= (p_1 \succ (p_2 \succ p_3)) \otimes q_1q_2q_3 = ((p_1 \succ p_2) \succ p_3) \otimes q_1q_2q_3 + ((p_2 \succ p_1) \succ p_3) \otimes q_1q_2q_3 \\ &= (A \succeq B) \succeq C + (B \succeq A) \succeq C, \\ (A \succeq B) \preceq C &= (p_3 \succ (p_1 \succ p_2)) \otimes q_1q_2q_3 = ((p_3 \succ p_1 + p_1 \succ p_3) \succ p_2) \otimes q_1q_2q_3 \\ &= (p_1 \succ (p_3 \succ p_2)) \otimes q_1q_2q_3 = A \succeq (B \preceq C), \end{aligned}$$

for $A = p_1 \otimes q_1$, $B = p_2 \otimes q_2$, $C = p_2 \otimes q_3$.

For more background on pre-Lie products and this relation to dendriform algebras see for example [EM09] and references therein.

The object R satisfies a quadratic fixed-point equation.

Lemma 4.2.6.

$$(\underline{D} - \mathrm{id})R = R \trianglerighteq R. \tag{4.10}$$

Proof. Let $|w| \geq 1$. Starting from (4.6) and concatenating a letter $\mathtt{a}$ from the right on both sides, we get

$$\sum_{uv=w,|u|\geq 1} (\rho(u) \sqcup\!\sqcup v)\mathtt{a} = (Dw)\mathtt{a} = (D - \mathrm{id})(w\mathtt{a}).$$

Hence

$$\sum_{uv=w,|u|\geq 1} \rho(u) \succ v\mathtt{a} = (D - \mathrm{id})(w\mathtt{a}),$$

which means, for $|\bar{w}| \geq 2$,

$$\sum_{uv=\bar{w},|u|\geq 1,|v|\geq 1} \rho(u) \succ v = (D - \mathrm{id})\bar{w}. \tag{4.11}$$

Recall

$$\begin{aligned} \mathrm{ad}_v\, w &= [v, w] \\ \mathrm{Ad}_v\, w &= [\mathtt{k}_1, [\mathtt{k}_2, \ldots, [\mathtt{k}_n, w] \ldots]], \end{aligned}$$

where $v = \mathtt{k}_1 \cdots \mathtt{k}_n$. By [Reu93, Theorem 1.4], for a Lie polynomial P one has

$$\mathrm{ad}_P = \mathrm{Ad}_P\,. \tag{4.12}$$

For a word w define the linear map I_w as

$$I_w x := w \succ x,$$

and extend linearly to the whole tensor algebra. The map

$$I_\bullet \otimes \mathrm{ad}_\bullet : \mathfrak{W} \to \mathrm{Hom}_{\mathbb{R}}\big(\mathfrak{W}, \mathfrak{W}\big),$$

is defined as

$$(I_x \otimes \mathrm{ad}_y) a \otimes b = (I_x a) \otimes (\mathrm{ad}_y\, b).$$

Now

$$\begin{aligned} (I_\bullet \otimes \mathrm{ad}_\bullet) R &= (I_\bullet \otimes \mathrm{ad}_\bullet) \sum w \otimes r(w) = (I_\bullet \otimes \mathrm{Ad}_\bullet) \sum w \otimes r(w) = (I_\bullet \otimes \mathrm{Ad}_\bullet) \sum \rho(v) \otimes v \\ &= \sum_{|v|\geq 1} I_{\rho(v)} \otimes \mathrm{Ad}_v, \end{aligned}$$

where we used (4.12) and then (4.4). Then

$$\begin{aligned} R \trianglerighteq R = ((I_\bullet \otimes \mathrm{ad}_\bullet) R) R &= \left(\sum_{|v|\geq 1} I_{\rho(v)} \otimes \mathrm{Ad}_v \right) \sum_{|w|\geq 1} w \otimes r(w) = \sum_{|v|,|w|\geq 1} (\rho(v) \succ w) \otimes r(vw) \\ &= \sum_{|x|\geq 2} \sum_{vw=x,|v|,|w|\geq 1} (\rho(v) \succ w) \otimes r(vw) = \sum_{|x|\geq 2} (|x| - 1) x \otimes r(x) \\ &= \sum_{|x|\geq 2} x \otimes \big((|x| - 1) r(x)\big) = \sum_{|x|\geq 2} x \otimes r[(D - \mathrm{id})x] = (\underline{D} - \mathrm{id}) R. \end{aligned}$$

□

Remark 4.2.7. We sketch the connection to the ODE approach of [Roc03b]. Let $S^\varepsilon_t := \delta_\varepsilon S(X)_t$ be the signature at time t, dilated by a factor $\varepsilon > 0$. Define

$$Z^\varepsilon_t := \frac{d}{d\varepsilon} S^\varepsilon_t \bullet (S^\varepsilon_t)^{-1},$$

which, as we have seen in Remark 4.2.1.2, is equal to $\varepsilon^{-1}r[S_t^\varepsilon]$. One can show (see [AGS89, Equation (1.8), (3.3) and (3.7)], where again pre-Lie algebras are called 'chronological algebras'), that Z_t^ε satisfies

$$\partial_\varepsilon Z_t^\varepsilon = \int_0^t \left[Z_r^\varepsilon, \dot{Z}_r^\varepsilon\right] dr, \tag{4.13}$$

where $[.,.]$ is the Lie bracket in $\mathfrak{g}((\mathbb{R}^d))$. We may give an alternative proof of (4.13) based on the quadratic fixed-point equation (4.10). For the left-hand side,

$$\begin{aligned}\partial_\varepsilon Z_t^\varepsilon &= \partial_\varepsilon \left(\varepsilon^{-1} r[S_t^\varepsilon]\right) = -\varepsilon^{-2} r[S_t^\varepsilon] + \varepsilon^{-1} r[\partial_\varepsilon S_t^\varepsilon] = -\varepsilon^{-2} r[S_t^\varepsilon] + \varepsilon^{-2} r[D S_t^\varepsilon] = \varepsilon^{-2} r[(D-\mathrm{id}) S_t^\varepsilon] \\ &= \varepsilon^{-2} \sum_w \langle S_t^\varepsilon, w\rangle \, r[(D-\mathrm{id})w] = \varepsilon^{-2} \mathsf{eval}_{S_t^\varepsilon}[(\underline{D}-\mathrm{id})R],\end{aligned}$$

Aiming at the right-hand side, we first note that in general for $p\otimes q, p'\otimes q' \in \mathfrak{W}$, we have

$$\begin{aligned}\int_0^t \left[\left\langle S_s, p\right\rangle q, \left\langle \dot{S}_s, p'\right\rangle q'\right] ds &= \int_0^t \left\langle S_s, p\right\rangle d\left\langle S_s, p'\right\rangle [q,q'] = \left\langle S_t, p \succ p'\right\rangle [q,q'] \\ &= \mathsf{eval}_{S_t}\left[(p\otimes q) \trianglerighteq (p'\otimes q')\right].\end{aligned}$$

This implies

$$\begin{aligned}\int_0^t \left[Z_r^\varepsilon, \dot{Z}_r^\varepsilon\right] dr &= \varepsilon^{-2}\left[r[S_t^\varepsilon], r[\dot{S}_t^\varepsilon]\right] = \varepsilon^{-2}\sum_w\sum_{w'}\int_0^t \left[\left\langle S_s^\varepsilon, w\right\rangle r[w], \left\langle \dot{S}_s^\varepsilon, w'\right\rangle r[w']\right] ds \\ &= \varepsilon^{-2}\sum_w\sum_{w'} \mathsf{eval}_{S_t^\varepsilon}[(w\otimes r[w]) \trianglerighteq (w'\otimes r[w'])] = \varepsilon^{-2}\mathsf{eval}_{S_t^\varepsilon}[R \trianglerighteq R].\end{aligned}$$

Putting things together, we thus have that (4.13) is equivalent to

$$\mathsf{eval}_{S_t^\varepsilon}[(\underline{D}-\mathrm{id})R] = \mathsf{eval}_{S_t^\varepsilon}[R \trianglerighteq R].$$

which is of course an immediate consequence of (4.10).

By symmetrizing the pre-Lie product in the quadratic fixed point equation (4.10), we make the area-operator appear. Define

$$\begin{aligned}(p\otimes q) \trianglerighteq_{\mathrm{Sym}} (p'\otimes q') &:= (p\otimes q) \trianglerighteq (p'\otimes q') + (p'\otimes q') \trianglerighteq (p\otimes q) \\ &= \mathsf{area}(p,p') \otimes [q,q'].\end{aligned}$$

Symmetrized pre-Lie products like this were already introduced by Rocha in [Roc03a, Equation (6.13) and Proposition 6.3].

Corollary 4.2.8 (based on [**?**, Proposition 6.8])**.** *We have*

$$(\underline{D}-\mathrm{id})R = \frac{1}{2} R \trianglerighteq_{\mathrm{Sym}} R.$$

Let $R_n := \mathsf{proj}_n R = \sum_{|w|=n} w\otimes r[w]$ *be the* n*-th level of* R*. Then for* $n\geq 2$ *this spells out as*

$$(n-1)R_n = \frac{1}{2}\sum_{\ell=1}^n R_\ell \trianglerighteq_{\mathrm{Sym}} R_{n-\ell} = \begin{cases} \sum_{\ell=1}^{(n-1)/2} R_\ell \trianglerighteq_{\mathrm{Sym}} R_{n-\ell} & n \text{ odd} \\ \sum_{\ell=1}^{n/2} R_\ell \trianglerighteq_{\mathrm{Sym}} R_{n-\ell} + \frac{1}{2} R_{n/2} \trianglerighteq_{\mathrm{Sym}} R_{n/2} & n \text{ even}\end{cases}$$

with $R_n \in \mathfrak{R} := \langle \mathtt{i}\otimes\mathtt{i},\ \mathtt{i} = \mathtt{1}\ldots\mathtt{d}; \trianglerighteq_{\mathrm{Sym}}\rangle$.

Proof. This follows immediately from Lemma 4.2.6. □

Remark 4.2.9. Note that [Roc03a, Proposition 6.8] has a slightly more complicated recursion. This stems from the fact that the Z^n there relates to our $\frac{R_n}{n!}$ here.

Example 4.2.10. *Let* $R_n := \mathsf{proj}_n R$ *be the* n*-th level of* R*. Then Lemma 4.2.6 and Corollary 4.2.8 give*

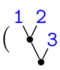
$$\begin{aligned} R_2 &= R_1 \trianglerighteq R_1 = \frac{1}{2} R_1 \trianglerighteq_{\mathrm{Sym}} R_1 \\ 2R_3 &= R_1 \trianglerighteq R_2 + R_2 \trianglerighteq R_1 = \frac{1}{2} R_1 \trianglerighteq_{\mathrm{Sym}} R_2 + \frac{1}{2} R_2 \trianglerighteq R_1 = R_1 \trianglerighteq_{\mathrm{Sym}} R_2 \\ 3R_4 &= R_1 \trianglerighteq R_3 + R_2 \trianglerighteq R_2 + R_3 \trianglerighteq R_1 = \frac{1}{2} R_1 \trianglerighteq_{\mathrm{Sym}} R_3 + \frac{1}{2} R_2 \trianglerighteq_{\mathrm{Sym}} R_2 + \frac{1}{2} R_3 \trianglerighteq_{\mathrm{Sym}} R_1 \\ &= R_1 \trianglerighteq_{\mathrm{Sym}} R_3 + \frac{1}{2} R_2 \trianglerighteq_{\mathrm{Sym}} R_2 \end{aligned}$$

For $d = 2$ *this becomes*

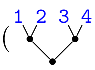
$$\begin{aligned} R_1 &= \mathtt{1} \otimes \mathtt{1} + \mathtt{2} \otimes \mathtt{2} \\ R_2 &= \frac{1}{2} \left(\mathsf{area}(\mathtt{1},\mathtt{2}) \otimes [\mathtt{1},\mathtt{2}] + \mathsf{area}(\mathtt{2},\mathtt{1}) \otimes [\mathtt{2},\mathtt{1}] \right) \\ R_3 &= \frac{1}{4} \Big(\mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{1},\mathtt{2})) \otimes [\mathtt{1},[\mathtt{1},\mathtt{2}]] + \mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{2},\mathtt{1})) \otimes [\mathtt{1},[\mathtt{2},\mathtt{1}]] \\ &\quad + \mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{1},\mathtt{2})) \otimes [\mathtt{2},[\mathtt{1},\mathtt{2}]] + \mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{2},\mathtt{1})) \otimes [\mathtt{2},[\mathtt{2},\mathtt{1}]] \Big) \\ R_4 &= \frac{1}{12} \Big(\mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{1},\mathtt{2}))) \otimes [\mathtt{1},[\mathtt{1},[\mathtt{1},\mathtt{2}]]] + \mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{2},\mathtt{1}))) \otimes [\mathtt{1},[\mathtt{1},[\mathtt{2},\mathtt{1}]]] \\ &\quad + \mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{1},\mathtt{2}))) \otimes [\mathtt{1},[\mathtt{2},[\mathtt{1},\mathtt{2}]]] + \mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{2},\mathtt{1}))) \otimes [\mathtt{1},[\mathtt{2},[\mathtt{2},\mathtt{1}]]] \\ &\quad + \mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{1},\mathtt{2}))) \otimes [\mathtt{2},[\mathtt{1},[\mathtt{1},\mathtt{2}]]] + \mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{2},\mathtt{1}))) \otimes [\mathtt{2},[\mathtt{1},[\mathtt{2},\mathtt{1}]]] \\ &\quad + \mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{1},\mathtt{2}))) \otimes [\mathtt{2},[\mathtt{2},[\mathtt{1},\mathtt{2}]]] + \mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{2},\mathsf{area}(\mathtt{2},\mathtt{1}))) \otimes [\mathtt{2},[\mathtt{2},[\mathtt{2},\mathtt{1}]]] \Big) \end{aligned}$$

In general this looks as follows.

Definition 4.2.11. *Denote by* BPT_n *the set of (complete, rooted) binary planar trees with* n *leaves labelled with the letters* $\mathtt{1},\ldots,\mathtt{d}$*. Given* $\tau \in \mathsf{BPT}_n$ *we define* $\mathsf{area}_\bullet(\tau)$ *(resp.* $\mathsf{lie}_\bullet(\tau)$*) as the bracketing-out using* area *(resp.* $[.,.]$*). For example*

$\mathsf{area}_\bullet($ 

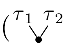
$) = \mathsf{area}(\mathtt{1},\mathsf{area}(\mathtt{2},\mathtt{3}))$

$\mathsf{lie}_\bullet($ 

$) = [[\mathtt{1},\mathtt{2}],[\mathtt{3},\mathtt{4}]].$

Define a function $c : \mathsf{BPT}_n \to \mathbb{R}$*, which does not depend on the specific letter labels, recursively as follows*

$$c(\mathtt{i}) = 1 \qquad \text{for any } \mathtt{i} \text{ in } \mathtt{1},\ldots,\mathtt{d}$$

$c($ 

$) = 2c(\tau_1)c(\tau_2)(|\tau_1|_{\mathsf{leaves}} + |\tau_2|_{\mathsf{leaves}} - 1)$

where $|\tau|_{\mathsf{leaves}}$ *denotes the number of leaves of the tree* τ*. For example*

$$c(2) = 1$$
$$c({}^{1}\!\vee^{2}) = 2 \cdot 1 \cdot 1 \cdot (2-1) = 2$$
$$c({}_{3}\!\vee^{\,{}^{1}\vee^{2}}) = 2 \cdot 1 \cdot 2 \cdot (3-1) = 8$$

Lemma 4.2.12 (based on [Roc03a, Lemma 6.11], [Roc03b, Lemma 1])**.**

$$R_n = \sum_{\tau \in \mathsf{BPT}_n} \frac{1}{c(\tau)} \, \mathsf{area}_\bullet(\tau) \otimes \mathsf{lie}_\bullet(\tau).$$

Remark 4.2.13. We note that [Roc03a, Lemma 6.11] and [Roc03b, Lemma 1] have a slightly more complicated expression for R_n, since there some of the terms are factored out, owing to antisymmetry. We do not pursue this here since the end result, also in Rocha's work, still contains redundant terms, which we do not know how to explicitly get rid of. In fact, due to antisymmetry alone, we already know that $c(\tau)$ is not a unique choice for this equation to hold, however it remains an interesting and more involved question if it is the only choice which is symmetric, i.e. well-defined on *non-planar* trees, and invariant under change of the leaf labels.

A further very interesting question is to find a modified c' which may not be symmetric and may depend on the leaf labels, such that the equations still hold, but such that the number of non-zero summands in the equation is minimized for each n.

Proof of 4.2.12. For the purpose of this proof, let R_n be defined as in Corollary 4.2.8 and

$$R_n' := \sum_{\tau \in \mathsf{BPT}_n} \frac{1}{c(\tau)} \, \mathsf{area}_\bullet(\tau) \otimes \mathsf{lie}_\bullet(\tau).$$

We proceed by induction over n. We have

$$R_1 = \sum_{|w|=1} w \otimes r(w) = \sum_{|w|=1} w \otimes w = \sum_{|w|=1} \frac{1}{c(w)} \mathsf{area}_\bullet(w) \otimes \mathsf{lie}_\bullet(w) = R_1'.$$

Assuming $R_n = R_n'$ holds for some $n \in \mathbb{N}$, we get

$$\begin{aligned}
R_{n+1} &= \frac{1}{2n} \sum_{l=1}^{n+1} R_l \trianglerighteq_{\mathrm{Sym}} R_{n-l} \\
&= \frac{1}{2n} \sum_{l=1}^{n+1} \sum_{\substack{\tau_1 \in \mathsf{BPT}_l, \\ \tau_2 \in \mathsf{BPT}_{n+1-l}}} \frac{1}{c(\tau_1)c(\tau_2)} \mathsf{area}(\mathsf{area}_\bullet(\tau_1), \mathsf{area}_\bullet(\tau_2)) \otimes [\mathsf{lie}_\bullet(\tau_1), \mathsf{lie}_\bullet(\tau_2)] \\
&= \sum_{l=1}^{n+1} \sum_{\substack{\tau_1 \in \mathsf{BPT}_l, \\ \tau_2 \in \mathsf{BPT}_{n+1-l}}} \frac{1}{c({}^{\tau_1}\!\vee^{\tau_2})} \mathsf{area}_\bullet({}^{\tau_1}\!\vee^{\tau_2}) \otimes \mathsf{lie}_\bullet({}^{\tau_1}\!\vee^{\tau_2}) \\
&= \sum_{\tau \in \mathsf{BPT}_{n+1}} \frac{1}{c(\tau)} \mathsf{area}_\bullet(\tau) \otimes \mathsf{lie}_\bullet(\tau) = R_{n+1}'.
\end{aligned}$$

□

Rather than working with the recursion for R_n from Corollary 4.2.8 directly on $T\langle\langle \mathcal{R}^d \rangle\rangle$, in the following theorem we will pursue the alternative approach of first applying the $\mathsf{coeval}^{P_h^*}$ to work on $T(\mathbb{R}^d)$, or, to be more specific, on $\mathscr{A}$ as we will see.

Theorem 4.2.14. *We have* $R = \sum_h \mathfrak{r}_h \otimes P_h$ *where* $\mathfrak{r}_h := \rho(S_h)$ *satisfies the recursion*

$$\mathfrak{r}_h = \frac{1}{|h|-1} \sum_{h_1<h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle \, \mathsf{area}(\mathfrak{r}_{h_1}, \mathfrak{r}_{h_2}). \tag{4.14}$$

More explicitly, we have

$$\mathfrak{r}_h = \sum_\tau \frac{1}{b(\tau)} q_\tau^h \, \mathsf{area}_\bullet(\tau) = \sum_\tau \frac{1}{c(\tau)} p_\tau^h \, \mathsf{area}_\bullet(\tau)$$

with

$$\begin{aligned} q_\tau^h &= \sum_{h_1<h_2} q_{\tau'}^{h_1} q_{\tau''}^{h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle, \\ p_\tau^h &= \langle P_h^*, \mathsf{lie}_\bullet(\tau)\rangle = \sum_{h_1,h_2} q_{\tau'}^{h_1} q_{\tau''}^{h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle \end{aligned}$$

for $|\tau|_{\mathsf{leaves}}, |h| \geq 2$ *and*

$$\begin{aligned} q_{\mathtt{i}}^h &= p_{\mathtt{i}}^h = \delta_{h,\mathtt{i}}, \\ q_\tau^{\mathtt{i}} &= p_\tau^{\mathtt{i}} = \delta_{\tau,\mathtt{i}} \end{aligned}$$

with $b(\tau) = b(\tau')b(\tau'')(|\tau'|_{\mathsf{leaves}} + |\tau''|_{\mathsf{leaves}} - 1)$, $b(\mathtt{i}) = 1$.

Proof. We have $r = \mathsf{eval}(R)$ with $\operatorname{Im} r = \mathfrak{g}(\mathbb{R}^d)$ and thus

$$r = r \circ \mathsf{eval}(\sum_h S_h \otimes P_h) = \mathsf{eval}\Big(\sum_{w,h} \langle w, S_h\rangle \rho(w) \otimes P_h\Big) = \mathsf{eval}\Big(\sum_h \rho(S_h) \otimes P_h\Big),$$

which means $R = \sum_h \rho(S_h) \otimes P_h$ since eval is bijective. Putting $\mathfrak{r}_h := \rho(S_h)$, we get

$$\sum_h (|h|-1)\mathfrak{r}_h \otimes P_h = (\underline{D} - \mathrm{id})R = \tfrac{1}{2}\, R \trianglerighteq_{\mathrm{Sym}} R = \tfrac{1}{2} \sum_{h_1,h_2} \mathsf{area}(\mathfrak{r}_{h_1}, \mathfrak{r}_{h_2}) \otimes [P_{h_1}, P_{h_2}],$$

which by applying $\mathsf{coeval}^{P_h^*}$ on both sides yields

$$(|h|-1)\mathfrak{r}_h = \tfrac{1}{2} \sum_{h_1,h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle \mathsf{area}(\mathfrak{r}_{h_1}, \mathfrak{r}_{h_2}) = \sum_{h_1<h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle \mathsf{area}(\mathfrak{r}_{h_1}, \mathfrak{r}_{h_2}).$$

Since due to $q_\tau^{\mathtt{i}} = \delta_{\tau,\mathtt{i}}$ and $b(\mathtt{i}) = 1$ we have

$$\mathfrak{r}_{\mathtt{i}} = \rho(\mathtt{i}) = \mathtt{i} = \mathsf{area}_\bullet(\mathtt{i}) = \sum_\tau \frac{1}{b(\tau)} q_\tau^{\mathtt{i}} \mathsf{area}_\bullet(\tau),$$

we obtain by induction

$$\begin{aligned}
\mathfrak{r}_h &= \frac{1}{|h|-1} \sum_{h_1<h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle \mathsf{area}(\mathfrak{r}_{h_1}, \mathfrak{r}_{h_2}) \\
&= \frac{1}{|h|-1} \sum_{h_1<h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle \sum_{\tau_1,\tau_2} \frac{1}{b(\tau_1)b(\tau_2)} q_{\tau_1}^{h_1} q_{\tau_2}^{h_2} \mathsf{area}(\mathsf{area}_\bullet(\tau_1), \mathsf{area}_\bullet(\tau_2)) \\
&= \sum_{\tau_1,\tau_2} \frac{1}{(|h|-1)b(\tau_1)b(\tau_2)} \sum_{h_1<h_2} q_{\tau_1}^{h_1} q_{\tau_2}^{h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle \mathsf{area}_\bullet({}^{\tau_1}\!\!\vee\!{}^{\tau_2}) \\
&= \sum_{\tau} \frac{1}{(|\tau'|_{\mathsf{leaves}} + |\tau''|_{\mathsf{leaves}} - 1)b(\tau')b(\tau'')} \sum_{h_1<h_2} q_{\tau'}^{h_1} q_{\tau''}^{h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle \mathsf{area}_\bullet(\tau) \\
&= \sum_{\tau} \frac{1}{b(\tau)} q_\tau^h \mathsf{area}_\bullet(\tau)
\end{aligned}$$

Furthermore, due to Lemma 4.2.12, we have

$$\mathfrak{r}_h = \mathsf{coeval}^{P_h^*}(\mathfrak{r}_{|h|}) = \sum_{\tau \in \mathsf{BPT}_{|h|}} \frac{1}{c(\tau)} \langle P_h^*, \mathsf{lie}_\bullet(\tau)\rangle \mathsf{area}_\bullet(\tau) = \sum_{\tau} \frac{1}{c(\tau)} p_\tau^h \mathsf{area}_\bullet(\tau),$$

where we have $p_{\mathtt{i}}^h = \langle P_h^*, \mathsf{lie}_\bullet(\mathtt{i})\rangle = \langle P_h^*, \mathtt{i}\rangle = \delta_{h,\mathtt{i}}$ as well as $p_\tau^{\mathtt{i}} = \langle P_{\mathtt{i}}^*, \mathsf{lie}_\bullet(\tau)\rangle = \langle \mathtt{i}, \mathsf{lie}_\bullet(\tau)\rangle = \delta_{\tau,\mathtt{i}}$ and by induction over $|\tau|_{\mathsf{leaves}} \geq 2$

$$\begin{aligned}
p_\tau^h &= \langle P_h^*, \mathsf{lie}_\bullet(\tau)\rangle = \langle P_h^*, [\mathsf{lie}_\bullet(\tau'), \mathsf{lie}_\bullet(\tau'')]\rangle = \sum_{h_1,h_2} \langle P_h^*, [\langle P_{h_1}^*, \mathsf{lie}_\bullet(\tau')\rangle P_{h_1}, \langle P_{h_2}^*, \mathsf{lie}_\bullet(\tau')\rangle P_{h_2}]\rangle \\
&= \sum_{h_1,h_2} p_{\tau'}^{h_1} p_{\tau''}^{h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle. \qquad \square
\end{aligned}$$

Remark 4.2.15. Let $h(\tau)$ be the Hall word corresponding to the Hall tree τ. Then,

$$q_\tau^{h_0} = p_\tau^{h_0} = \delta_{h(\tau),h_0}.$$

This is immediate by definition of q for $|h(\tau)| = |\tau|_{\mathsf{leaves}} = 1$, and then by induction over $|\tau|_{\mathsf{leaves}}$ if τ is a Hall tree, then τ', τ'' are Hall trees and thus

$$\begin{aligned}
q_\tau^h &= \sum_{h_1<h_2} q_{\tau'}^{h_1} q_{\tau''}^{h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle = \sum_{h_1<h_2} \delta_{h(\tau'),h_1} \delta_{h(\tau''),h_2} \langle P_h^*, [P_{h_1}, P_{h_2}]\rangle \\
&= \langle P_h^*, [P_{h(\tau')}, P_{h(\tau'')}]\rangle = \delta_{h(\tau),h}
\end{aligned}$$

due to $(P_h)_h$ and $(P_h^*)_h$ being dual bases. For $p_\tau^h = \langle P_h^*, \mathsf{lie}_\bullet(\tau)\rangle$ the claim is immediate.

Note furthermore that we could also have derived

$$\mathfrak{r}_h = \sum_\tau \frac{1}{b(\tau)} q_\tau^h \mathsf{area}_\bullet(\tau)$$

from

$$\mathfrak{r}_h = \sum_\tau \frac{1}{c(\tau)} p_\tau^h \mathsf{area}_\bullet(\tau)$$

by looking at how q_τ^h and p_τ^h relate to each other.

Example 4.2.16. *In the case of $T(\mathbb{R}^2)$ and the Lyndon words H, the values of $\mathfrak{r}_h$ up to level five are*

$$
\begin{aligned}
\mathfrak{r}_{1} &= 1, \quad \mathfrak{r}_{2} = 2,\\
\mathfrak{r}_{12} &= 12 - 21 = \mathsf{area}(1,2),\\
\mathfrak{r}_{112} &= 112 - 121 = \tfrac{1}{2}\,\mathsf{area}(1,\mathsf{area}(1,2)),\\
\mathfrak{r}_{122} &= -212 + 221 = \tfrac{1}{2}\,\mathsf{area}(\mathsf{area}(1,2),2),\\
\mathfrak{r}_{1112} &= 1112 - 1121 = \tfrac{1}{6}\,\mathsf{area}(1,\mathsf{area}(1,\mathsf{area}(1,2)))\\
\mathfrak{r}_{1122} &= -1212 + 1221 - 2112 + 2121 = \tfrac{1}{6}\,\mathsf{area}(1,\mathsf{area}(\mathsf{area}(1,2),2)) + \tfrac{1}{6}\,\mathsf{area}(\mathsf{area}(1,\mathsf{area}(1,2)),2),\\
\mathfrak{r}_{1222} &= 2212 - 2221 = \tfrac{1}{6}\,\mathsf{area}(\mathsf{area}(\mathsf{area}(1,2),2),2),\\
\mathfrak{r}_{11112} &= 11112 - 11121,\\
\mathfrak{r}_{11122} &= -11212 + 11221 - 12112 + 12121 - 21112 + 21121,\\
\mathfrak{r}_{11222} &= 12212 - 12221 + 21212 - 21221 + 22112 - 22121,\\
\mathfrak{r}_{12122} &= 21212 - 21221 + 22112 - 22121,\\
\mathfrak{r}_{11212} &= 21112 - 21121,\\
\mathfrak{r}_{12222} &= -22212 + 22221,
\end{aligned}
$$

where we gave the area bracketings according to the recursion Equation (4.14) *up to level four. The trend of the values being just* $-1, 0, 1$ *combinations of words does not continue to higher levels, e.g.*

$$\mathfrak{r}_{112212} = -211212 + 211221 - 212112 + 212121 - 3\,221112 + 3\,221121.$$

Example 4.2.17. *In the case of $T(\mathbb{R}^3)$ and the Lyndon words H, the values of $\mathfrak{r}_h$ up to level four which are not immediate from the previous example are*

$$
\begin{aligned}
\mathfrak{r}_{123} &= 123 - 132 - 312 + 321,\\
\mathfrak{r}_{132} &= -213 + 231 - 312 + 321,\\
\mathfrak{r}_{1123} &= 1123 - 1132 - 1312 + 1321 - 3112 + 3121,\\
\mathfrak{r}_{1132} &= -1213 + 1231 - 1312 + 1321 - 2113 + 2131 - 3112 + 3121,\\
\mathfrak{r}_{1213} &= -2113 + 2131 + 3112 - 3121,\\
\mathfrak{r}_{1223} &= 1223 - 1232 + 3212 - 3221,\\
\mathfrak{r}_{1232} &= -2123 + 2132 + 2312 - 2321 + 2\,3212 - 2\,3221,\\
\mathfrak{r}_{1233} &= -1323 + 1332 - 3123 + 3132 + 3312 - 3321,\\
\mathfrak{r}_{1322} &= 2213 - 2231 + 2312 - 2321 + 3212 - 3221,\\
\mathfrak{r}_{1323} &= -3123 + 3132 + 3213 - 3231 + 2\,3312 - 2\,3321,\\
\mathfrak{r}_{1332} &= 2313 - 2331 + 3213 - 3231 + 3312 - 3321.
\end{aligned}
$$

Example 4.2.18. *In the case of $T(\mathbb{R}^2)$ and the standard Hall words H, the values of $\mathfrak{r}_h$ up to* *level five are*

$$
\begin{aligned}
\mathfrak{r}_1 &= \mathtt{1}, \quad \mathfrak{r}_2 = \mathtt{2},\\
\mathfrak{r}_{12} &= \mathtt{12} - \mathtt{21},\\
\mathfrak{r}_{121} &= -\mathtt{112} + \mathtt{121},\\
\mathfrak{r}_{122} &= -\mathtt{212} + \mathtt{221},\\
\mathfrak{r}_{1211} &= \mathtt{1112} - \mathtt{1121},\\
\mathfrak{r}_{1221} &= \mathtt{1212} - \mathtt{1221} + \mathtt{2112} - \mathtt{2121},\\
\mathfrak{r}_{1222} &= \mathtt{2212} - \mathtt{2221},\\
\mathfrak{r}_{12111} &= -\mathtt{11112} + \mathtt{11121},\\
\mathfrak{r}_{12211} &= -\mathtt{11212} + \mathtt{11221} - \mathtt{12112} + \mathtt{12121} - \mathtt{21112} + \mathtt{21121},\\
\mathfrak{r}_{12221} &= -\mathtt{12212} + \mathtt{12221} - \mathtt{21212} + \mathtt{21221} - \mathtt{22112} + \mathtt{22121},\\
\mathfrak{r}_{12222} &= -\mathtt{22212} + \mathtt{22221},\\
\mathfrak{r}_{12112} &= -\mathtt{21112} + \mathtt{21121},\\
\mathfrak{r}_{12212} &= -\mathtt{21212} + \mathtt{21221} - \mathtt{22112} + \mathtt{22121},
\end{aligned}
$$

where once again the trend of the values being just $-1, 0, 1$ *combinations of words does not continue to higher levels, e.g.*

$$
\begin{aligned}
\mathfrak{r}_{122112} = \mathtt{121212} &- \mathtt{121221} + \mathtt{122112} - \mathtt{122121} + 2\,\mathtt{211212} - 2\,\mathtt{211221}\\
&+ 2\,\mathtt{212112} - 2\,\mathtt{212121} + 3\,\mathtt{221112} - 3\,\mathtt{221121}.
\end{aligned}
$$

4.3 Coordinates of the first kind

Let $(P_h)_{h\in H}$ be a basis for the free Lie algebra $\mathfrak{g}(\mathbb{R}^d)$. For the index set H we have a Hall set in mind ([Reu93, Section 4.1 and 4.2]), but this is not necessary at this stage. Any grouplike element $g \in \mathcal{G}$ can be written as the exponential of a Lie series,

$$
g = \exp\left(\sum_{h\in H} c_h(g) P_h\right), \tag{4.15}
$$

for some uniquely determined $c_h(g) \in \mathbb{R}$. In fact, there exist unique $\zeta_h \in T(\mathbb{R}^d)$, $h \in H$ such that $c_h(g) = \left\langle \zeta_h, g \right\rangle$. The ζ_h are called the **coordinates of the first kind** (corresponding to $(P_h)_{h\in H}$), see for example [Kaw09, page 1037].

We now formulate this in a way, where we do not have to test against $g \in G$. Recall the product $\blacksquare$ on $\mathfrak{W}$: shuffle product on the left and concatenation product on the right.

For words a, b

$$
\begin{aligned}
\mathsf{eval}_g(a \otimes b) \bullet \mathsf{eval}_g(a' \otimes b') &= \left\langle a, g \right\rangle \left\langle a', g \right\rangle bb' = \left\langle a \sqcup\!\sqcup a', g \right\rangle bb'\\
&= \mathsf{eval}_g\Big((a \sqcup\!\sqcup a') \otimes (b \bullet b')\Big) = \mathsf{eval}_g\Big((a \otimes b) \blacksquare (a' \otimes b')\Big).
\end{aligned}
$$

Both expressions are bilinear, so this is true for general elements in $\mathfrak{W}$. Hence eval_g is an algebra homomorphism from $(\mathfrak{W}, \blacksquare)$ to $(T((\mathbb{R}^d)), \bullet)$. Then on one hand, using first (4.15) and then the homomorphism property

$$g = \exp\left(\sum_{h\in H} c_h P_h\right) = \exp\left(\sum_{h\in H}\langle \zeta_h, g\rangle P_h\right) = \exp\left(\mathsf{eval}_g\left(\sum_{h\in H}\zeta_h\otimes P_h\right)\right)$$
$$= \mathsf{eval}_g\left(\exp_{\blacksquare}\left(\sum_{h\in H}\zeta_h\otimes P_h\right)\right).$$

Here, of course, for $x \in \mathfrak{W}$,

$$\exp_{\blacksquare}(x) := \sum_{n\geq 0}\frac{x^{\blacksquare n}}{n!} := \sum_{n\geq 0}\frac{\overbrace{x\blacksquare\cdots\blacksquare x}^{n \text{ times}}}{n!}.$$

On the other hand, trivially

$$g = \sum_w \langle w, g\rangle\ w = \mathsf{eval}_g\left(\sum_w w\otimes w\right).$$

Since grouplike elements projectively span all of $T((\mathbb{R}^d))$, with that we mean $\text{span}(\mathsf{proj}_{\leq n}\mathcal{G}_d) = T_{\leq n}(\mathbb{R}^d)$ (e.g. [DR19, Lemma 3.4]), we get that for all $x \in T((\mathbb{R}^d))$

$$\mathsf{eval}_x\left(\exp_{\blacksquare}\left(\sum_{h\in H}\zeta_h\otimes P_h\right)\right) = \mathsf{eval}_x\left(\sum_w w\otimes w\right),$$

which is equivalent to

$$\sum_w w\otimes w = \exp_{\blacksquare}\left(\sum_{h\in H}\zeta_h\otimes P_h\right), \tag{4.16}$$

respectively

$$\log_{\blacksquare}\sum_w w\otimes w = \sum_{h\in H}\zeta_h\otimes P_h.$$

We have arrived at a definition of coordinates of the first kind which does not rely on testing against grouplike elements.

Remark 4.3.1. Considering S as an element of $T\langle\langle\mathcal{R}^d\rangle\rangle$, it is grouplike. Indeed, for $a, b \in T(\mathbb{R}^d)$,

$$\begin{aligned}\langle \underline{a}\boxtimes\underline{b}, \Delta_{\underline{\sqcup\!\sqcup}}S\rangle &= \langle \underline{a}\boxtimes\underline{b}, \Delta_{\underline{\sqcup\!\sqcup}}\sum_w w\,\underline{w}\rangle = \sum_w w\,\langle \underline{a}\boxtimes\underline{b}, \Delta_{\underline{\sqcup\!\sqcup}}\underline{w}\rangle = \sum_w w\langle \underline{a}\,\underline{\sqcup\!\sqcup}\,\underline{b}, \underline{w}\rangle\\ &= a\sqcup\!\sqcup b = \langle \underline{a}\boxtimes\underline{b}, \sum_{w,v} w\sqcup\!\sqcup v\ \underline{w}\boxtimes\underline{v}\rangle = \langle \underline{a}\boxtimes\underline{b}, \sum_w w\,\underline{w}\boxtimes\sum_v v\,\underline{v}\rangle\\ &= \langle \underline{a}\boxtimes\underline{b}, S\boxtimes S\rangle.\end{aligned}$$

Here for a word $w \in T(\mathbb{R}^d)$ we write $\underline{w}$ as its realization in $T\langle\langle\mathcal{R}^d\rangle\rangle$. Then

$$\Lambda := \log_{\blacksquare} S,$$

is primitive. The search for coordinates of the first kind then amounts to finding a "simple" expression for this primitive element.

One can construct the coordinates ζ_h as follows. Pick any $S_h \in T(\mathbb{R}^d), h \in H$, such that

$$\langle S_h, P_{h'} \rangle = \delta_{h,h'}.$$

One can actually pick the S_h in such a way that they extend to the dual of the corresponding PBW basis of $T((\mathbb{R}^d))$[5] but this is not necessary here. Then

$$\left\langle S_h, \log\left(\exp\left(\sum_{h' \in H} c_{h'} P_{h'}\right)\right)\right\rangle = \left\langle S_h, \sum_{h' \in H} c_{h'} P_{h'}\right\rangle = c_h.$$

We want "to put the logarithm on the other side". This is indeed possible, since the logarithm on grouplike elements extends to a *linear* map π_1 on all of $T((\mathbb{R}^d))$ (see [Reu93, Section 3.2], and also [MNT13, Section 1 and 10] for a general overview on Lie idempotents), given as [Reu93, Equation (3.2.3)]

$$\pi_1(u) := \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} \sum_{v_1, \ldots, v_n \text{ non-empty}} \langle v_1 \sqcup\!\sqcup \cdots \sqcup\!\sqcup v_n, u \rangle \; v_1 \bullet \cdots \bullet v_n. \tag{4.17}$$

Denote its dual map by $\pi_1^\top$.[6] It is given as [Reu93, proof of Theorem 6.3]

$$\pi_1^\top(v) = \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} \sum_{u_1, \ldots, u_n \text{ non-empty}} \langle v, u_1 \bullet \cdots \bullet u_n \rangle \; u_1 \sqcup\!\sqcup \cdots \sqcup\!\sqcup u_n.$$

Then for all $h \in H$

$$\left\langle \pi_1^\top S_h, \exp\left(\sum_{h' \in H} c_{h'} P_{h'}\right)\right\rangle = c_h,$$

that is, the coordinates of the first kind are given by [GK08, Theorem 1]

$$\zeta_h = \pi_1^\top S_h \qquad h \in H. \tag{4.18}$$

We note that ζ_h must of course be independent of the choice of the S_h and this is indeed the case, since $\ker \pi_1^\top = (\operatorname{im} \pi_1)^\perp = \mathfrak{g}_n(\mathbb{R}^d)^\perp$.

Example 4.3.2. *Let $(P_h)_{h \in H}$ be the Lyndon basis (which is a Hall basis, [Reu93, Theorem 5.1]). In the case $d = 2$, we give in Table 4.1 the first few elements for P_h, S_h and $\pi_1^\top S_h$, where we take $(S_h)_{h \in H}$ as in [Reu93, Theorem 5.3].*

The expressions given by (4.18) can become quite unwieldy. Finding tractable expressions for coordinates of the first kind was a main motivation for Rocha in [Roc03b] and [Roc03a]. We will now reproduce some of his results using purely algebraic arguments.

---

[5] If P_h is a Hall basis, [Reu93, Theorem 5.3] gives recursive formulae for the dual of the corresponding PBW basis

[6] [Reu93, Section 6.2] uses the notation π_1^*.

Lyndon word h	P_h	S_h	$\zeta_h = \pi_1^\top S_h$
1	1	1	1
2	2	2	2
12	$[1, 2]$	12	$+\frac{1}{2}12 - \frac{1}{2}21$
112	$[1, [1, 2]]$	112	$+\frac{1}{6}112 - \frac{1}{3}121 + \frac{1}{6}211$
122	$[[1, 2], 2]$	122	$+\frac{1}{6}122 - \frac{1}{3}212 + \frac{1}{6}221$
1112	$[1, [1, [1, 2]]]$	1112	$-\frac{1}{6}1121 + \frac{1}{6}1211$
1122	$[1, [[1, 2], 2]]$	1122	$+\frac{1}{6}1122 - \frac{1}{6}1212 + \frac{1}{6}2121 - \frac{1}{6}2211$
1222	$[[[1, 2], 2], 2]$	1222	$-\frac{1}{6}2122 + \frac{1}{6}2212$
11112	$[1, [1, [1, [1, 2]]]]$	11112	$\frac{1}{30}\left[-11112 - 11121 + 4\,11211 - 12111 - 21111\right]$
11122	$[1, [1, [[1, 2], 2]]]$	11122	$\frac{1}{30}\left[2\,11122 - 3\,11212 - 3\,11221 + 2\,12112 + 2\,12121 - 3\,12211 + 2\,21112 + 2\,21121 - 3\,21211 + 2\,22111\right]$
11222	$[1, [[[1, 2], 2], 2]]$	11222	$\frac{1}{30}\left[2\,11222 - 3\,12122 + 2\,12212 + 2\,12221 - 3\,21122 + 2\,21212 + 2\,21221 - 3\,22112 - 3\,22121 + 2\,22211\right]$
12122	$[[1, 2], [[1, 2], 2]]$	$12122 + 3\ 11222$	$\frac{1}{30}\left[3\,11222 - 2\,12122 - 2\,12212 + 3\,12221 - 2\,21122 + 3\,21212 - 2\,21221 - 2\,22112 - 2\,22121 + 3\,22211\right]$
11212	$[[1, [1, 2]], [1, 2]]$	$11212 + 2\ 11122$	$\frac{1}{30}\left[11122 + 11212 + 11221 - 4\,12112 + 12121 + 12211 + 21112 - 4\,21121 + 21211 + 22111\right]$
12222	$[[[[1, 2], 2], 2], 2]$	12222	$\frac{1}{30}\left[-12222 - 21222 + 4\,22122 - 22212 - 22221\right]$

Table 4.1: Example values for the Lyndon basis on two elements. The first column shows the Lyndon words, which are the Hall words for this basis. For each Lyndon word h, we show element P_h of the Hall basis which is also the PBW basis element labelled by h. Next we show the corresponding element S_h of the dual PBW basis, which also serves as S_h described above. Finally we show the corresponding coordinate of the second kind.

Lyndon word h	P_h	S_h	$\zeta_h = \pi_1^\top S_h$
1	1	1	1
12	$[1, 2]$	12	$\frac{1}{2}12 - \frac{1}{2}21$
112	$[1, [1, 2]]$	112	$\frac{1}{6}[112 - 2\,121 + 211]$
122	$[[1, 2], 2]$	122	$\frac{1}{6}[122 - 2\,212 + 221]$
123	$[1, [2, 3]]$	123	$\frac{1}{6}[2\,123 - 132 - 213 - 231 - 312 + 2\,321]$
132	$[[1, 3], 2]$	$123 + 132$	$\frac{1}{6}[123 + 132 - 2\,213 + 231 - 2\,312 + 321]$
1123	$[1, [1, [2, 3]]]$	1123	$\frac{1}{6}\left[1123 - 1213 - 1231 + 1321 + 3121 - 3211\right]$
1132	$[1, [[1, 3], 2]]$	$1123 + 1132$	$\frac{1}{6}\left[1123 + 1132 - 1213 - 1312 + 2131 - 2311 + 3121 - 3211\right]$
1213	$[[1, 2], [1, 3]]$	$1123 + 1132 + 1213$	$\frac{1}{6}\left[1213 - 1312 - 2113 + 2131 + 3112 - 3121\right]$

Table 4.2: Example values for the Lyndon basis on three elements. The first column shows the Lyndon words, which are the Hall words for this basis. For each Lyndon word h, we show element P_h of the Hall basis which is also the PBW basis element labelled by h. Next we show the corresponding element S_h of the dual PBW basis, which also serves as S_h described above. Finally we show the corresponding coordinate of the second kind.

Hall word h	P_h	S_h	$\zeta_h = \pi_1^\top S_h$
1	1	1	1
2	2	2	2
12	$[1, 2]$	12	$\frac{1}{2}12 - \frac{1}{2}21$
121	$[[1, 2], 1]$	$112 + 121$	$\frac{1}{6}[2\,121 - 112 - 211]$
122	$[[1, 2], 2]$	122	$\frac{1}{6}[122 + 221 - 2\,212]$
1211	$[[[1, 2], 1], 1]$	$1112 + 1121 + 1211$	$\frac{1}{6}[1211 - 1121]$
1221	$[[[1, 2], 2], 1]$	$1122 + 1212 + 1221$	$\frac{1}{6}[1212 - 1122 - 2121 + 2211]$
1222	$[[[1, 2], 2], 2]$	1222	$\frac{1}{6}[2212 - 2122]$
12111	$[[[[1, 2], 1], 1], 1]$	$11112 + 11121$ $+11211 + 12111$	$\frac{1}{30}[11112 + 11121 - 4\,11211 + 12111 + 21111]$
12211	$[[[[1, 2], 2], 1], 1]$	$11122 + 11212 + 11221$ $+12112 + 12121 + 12211$	$\frac{1}{30}[2\,11122 - 3\,11212 - 3\,11221 + 2\,12112 + 2\,12121$ $-3\,12211 + 2\,21112 + 2\,21121 - 3\,21211 + 2\,22111]$
12221	$[[[[1, 2], 2], 2], 1]$	$11222 + 12122$ $+12212 + 12221$	$\frac{1}{30}[-2\,11222 + 3\,11112 - 2\,12212 - 2\,12221 + 3\,21122$ $-2\,21212 - 2\,21221 + 3\,22112 + 3\,22121 - 2\,22211]$
12222	$[[[[1, 2], 2], 2], 2]$	12222	$\frac{1}{30}[-12222 - 21222 + 4\,22122 - 22212 - 22221]$
12112	$[[[1, 2], 1], [1, 2]]$	$4\,11122 + 3\,11212 + 2\,11221$ $+2\,12112 + 12121$	$\frac{1}{30}[-11122 - 11212 - 11221 + 4\,12112 - 12121$ $-12211 - 21112 + 4\,21121 - 21211 - 22111]$
12112	$[[[1, 2], 2], [1, 2]]$	$3\,11222 + 2\,12122 + 12212$	$\frac{1}{30}[-3\,11222 + 2\,12122 + 2\,12212 - 3\,12221 + 2\,21122$ $-3\,21212 + 2\,21221 + 2\,22112 + 2\,22121 - 3\,22211]$

Table 4.3: Example values for the standard Hall basis on two elements. The first column shows the Hall words. For each Hall word h, we show element P_h of the Hall basis which is also the PBW basis element labelled by h. Next we show the corresponding element S_h of the dual PBW basis, which also serves as S_h described above. Finally we show the corresponding coordinate of the second kind.

4.3.1 Coordinates of the first kind in terms of areas-of-areas

As in Remark 4.3.1 we consider the grouplike element $S \in T\langle\langle \mathcal{R}^d \rangle\rangle$. The goal is to find a "simple expression" for

$$\Lambda := \log_{\blacksquare} S.$$

Following Rocha, we obtain

$$R = \underline{r}(S) = (\underline{r} \circ \exp_{\blacksquare})[\Lambda].$$

The last step consists now in inverting $\underline{r} \circ \exp_{\blacksquare}$ here. We shall need the following version of Baker's identity [Reu93, Equation (1.6.5)].

Lemma 4.3.3. *Let $x, q \in T((\mathbb{R}^d))$ (resp. $L, Q \in T\langle\langle \mathcal{R}^d \rangle\rangle$) with q (resp. Q) having no coefficient in the empty word e (resp. $\underline{e}$) and x (resp. L) primitive. Then*

$$r(x \bullet q) = [x, r(q)], \qquad \underline{r}(L \blacksquare Q) = [L, \underline{r}(Q)]_{\blacksquare}.$$

Proof. For x, L Lie, by [Reu93, Theorem 1.4 (ii)], $\mathrm{ad}_x = \mathrm{Ad}_x$ on $T((\mathbb{R}^d))$ and $\mathrm{ad}_L = \mathrm{Ad}_L$ on $T\langle\langle \mathcal{R}^d \rangle\rangle$. Hence for $q \in T((\mathbb{R}^d))$ and $Q \in T\langle\langle \mathcal{R}^d \rangle\rangle$ any polynomial having no coefficient in the empty word,

$$r[x \bullet q] = \mathrm{ad}_x\, r[q] = \mathrm{Ad}_x\, r[q] = [x, r[q]],$$
$$\underline{r}[L \blacksquare Q] = \mathrm{ad}_L\, \underline{r}[Q] = \mathrm{Ad}_L\, \underline{r}[Q] = [L, \underline{r}[Q]]_{\blacksquare}. \qquad \square$$

We denote by $[.,.]_{\blacksquare}$ the Lie bracket on $\mathfrak{W}$ coming from the product $\blacksquare$. Note that

$$[p \otimes p', q \otimes q']_{\blacksquare} = (p \sqcup\!\sqcup q) \otimes [p', q'].$$

Remark 4.3.4. This is the Lie structure for the pre-Lie structure $\trianglerighteq$ (Theorem 4.2.5), i.e.

$$[x, y]_{\blacksquare} = x \blacksquare y - y \blacksquare x = x \trianglerighteq y - y \trianglerighteq x = x \succeq y - y \succeq x - y \preceq x + x \preceq y.$$

For $x \in \mathfrak{W}$ denote by $\mathrm{ad}_{\blacksquare;x}$ the corresponding adjunction operator, i.e. $\mathrm{ad}_{\blacksquare;x}\, y := [x, y]_{\blacksquare}$.
Let $\Lambda \in T\langle\langle \mathcal{R}^d \rangle\rangle$ be primitive. Then, using Lemma 4.3.3,

$$\underline{r}\left(\Lambda^{\blacksquare n}\right) = [\Lambda, \underline{r}\left(\Lambda^{n-1}\right)]_{\blacksquare}.$$

Iterating this, we get

$$\underline{r}(\Lambda^{\blacksquare n}) = (\mathrm{ad}_{\blacksquare;\Lambda})^{n-1} \underline{D}\Lambda.$$

Hence

$$R = \underline{r}[\exp_{\blacksquare}(\Lambda)] = \underline{r}\left[\sum_{n \geq 0} \frac{\Lambda^{\blacksquare n}}{n!}\right] = \sum_{n \geq 1} \frac{(\mathrm{ad}_{\blacksquare;\Lambda})^{n-1}}{n!} \underline{D}\Lambda. \tag{4.19}$$

This can now be used to recursively construct Λ from R. Put

$$[x_1, \ldots, x_n]_{\blacksquare} := [x_1, [\ldots, [x_{n-1}, x_n] \ldots]_{\blacksquare}, \quad [x_1, x_2]_{\blacksquare} := [x_1, x_2]_{\blacksquare}, \quad [x]_{\blacksquare} := x.$$

Proposition 4.3.5. *We have* $\Lambda_1 = R_1$ *and*

$$\Lambda_n = \frac{1}{n} R_n - \frac{1}{n} \sum_{i=2}^{n} \frac{1}{i!} \sum_{\substack{n_1,\dots,n_i \\ n_1+\dots+n_i=n}} n_i \left[\Lambda_{n_1}, \dots, \Lambda_{n_i}\right]_\blacksquare.$$

Proof. Rewriting Equation (4.19) for the homogeneous part R_n yields

$$\begin{aligned} R_n &= \underline{D}\Lambda_n + \frac{1}{2} \sum_{m=1}^{n-1} [\Lambda_m, \underline{D}\Lambda_{n-m}]_\blacksquare + \sum_{i=3}^{n} \frac{1}{i!} \sum_{\substack{n_1,\dots,n_1 \\ n_1+\dots+n_1=n}} [\Lambda_{n_1}, \dots, \Lambda_{n_{i-1}}, \underline{D}\Lambda_{n_i}]_\blacksquare \\ &= n\Lambda_n + \sum_{m=1}^{n-1} (n-m)[\Lambda_m, \Lambda_{n-m}]_\blacksquare + \sum_{i=3}^{n} \frac{1}{i!} \sum_{\substack{n_1,\dots,n_1 \\ n_1+\dots+n_1=n}} n_i[\Lambda_{n_1}, \dots, \Lambda_{n_{i-1}}, \Lambda_{n_i}]_\blacksquare \\ &= n\Lambda_n + \sum_{i=2}^{n} \frac{1}{i!} \sum_{\substack{n_1,\dots,n_1 \\ n_1+\dots+n_1=n}} n_i[\Lambda_{n_1}, \dots, \Lambda_{n_i}]_\blacksquare, \end{aligned}$$

which shows the claim. □

Example 4.3.6. *Let us spell out the first few summands of* (4.19)*,*

$$R = \underline{D}\Lambda + \frac{1}{2!}\,[\Lambda, \underline{D}\Lambda]_\blacksquare + \frac{1}{3!}\,[\Lambda, [\Lambda, \underline{D}\Lambda]_\blacksquare]_\blacksquare + \frac{1}{4!}\,\big[\Lambda, [\Lambda, [\Lambda, \underline{D}\Lambda]_\blacksquare]_\blacksquare\big]_\blacksquare + \dots$$

Level by level (remember that $\Lambda_0 = 0$*), we see*

$$\begin{aligned} R_1 &= \Lambda_1 \\ R_2 &= 2\Lambda_2 + \tfrac{1}{2!}[\Lambda_1, \Lambda_1]_\blacksquare \\ R_3 &= 3\Lambda_3 + \tfrac{1}{2!}\big([\Lambda_1, 2\Lambda_2]_\blacksquare + [\Lambda_2, \Lambda_1]_\blacksquare\big) + \tfrac{1}{3!}[\Lambda_1, [\Lambda_1, \Lambda_1]_\blacksquare]_\blacksquare \\ R_4 &= 4\Lambda_4 + \tfrac{1}{2!}\big([\Lambda_1, 3\Lambda_3]_\blacksquare + [\Lambda_2, 2\Lambda_2]_\blacksquare + [\Lambda_3, \Lambda_1]_\blacksquare\big) \\ &\quad + \tfrac{1}{3!}\big([\Lambda_1, [\Lambda_1, 2\Lambda_2]_\blacksquare]_\blacksquare + [\Lambda_1, [\Lambda_2, \Lambda_1]_\blacksquare]_\blacksquare + [\Lambda_2, [\Lambda_1, \Lambda_1]_\blacksquare]_\blacksquare\big) \\ &\quad + \tfrac{1}{4!}[\Lambda_1, [\Lambda_1, [\Lambda_1, \Lambda_1]_\blacksquare]_\blacksquare]_\blacksquare \end{aligned}$$

Plugging in the expressions from Example 4.2.10 in for R, we get for $d = 2$

$$
\begin{aligned}
\Lambda_1 &= R_1 = 1 \otimes 1 + 2 \otimes 2 \\
\Lambda_2 &= \tfrac{1}{2} R_2 = \tfrac{1}{4}\big(\mathsf{area}(1,2) \otimes [1,2] + \mathsf{area}(2,1) \otimes [2,1]\big) = \tfrac{1}{2}\mathsf{area}(1,2) \otimes [1,2] \\
\Lambda_3 &= \tfrac{1}{3}\left(R_3 - \tfrac{1}{2}[\Lambda_1, 2\Lambda_2]_{\blacksquare} - \tfrac{1}{2}[\Lambda_2, \Lambda_1]_{\blacksquare} - \tfrac{1}{3}[\Lambda_1,[\Lambda_1,\Lambda_1]_{\blacksquare}]_{\blacksquare}\right) = \tfrac{1}{3}\left(R_3 - \tfrac{1}{2}[\Lambda_1, \Lambda_2]_{\blacksquare}\right) \\
&= \tfrac{1}{6}\mathsf{area}(1, \mathsf{area}(1,2)) \otimes [1,[1,2]] + \tfrac{1}{6}\mathsf{area}(2, \mathsf{area}(1,2)) \otimes [2,[1,2]] \\
&\qquad - \tfrac{1}{12}\left(1 \sqcup\!\sqcup\, \mathsf{area}(1,2)\right) \otimes [1,[1,2]] - \tfrac{1}{12}\left(2 \sqcup\!\sqcup\, \mathsf{area}(1,2)\right) \otimes [2,[1,2]] \\
\Lambda_4 &= \tfrac{1}{4}\Big\{R_4 - \tfrac{1}{2!}\big([\Lambda_1, 3\Lambda_3]_{\blacksquare} + [\Lambda_2, 2\Lambda_2]_{\blacksquare} + [\Lambda_3, \Lambda_1]_{\blacksquare}\big) \\
&\qquad - \tfrac{1}{3!}\big([\Lambda_1,[\Lambda_1,2\Lambda_2]_{\blacksquare}]_{\blacksquare} + [\Lambda_1,[\Lambda_2,\Lambda_1]_{\blacksquare}]_{\blacksquare} + [\Lambda_2,[\Lambda_1,\Lambda_1]_{\blacksquare}]_{\blacksquare}\big) \\
&\qquad - \tfrac{1}{4!}[\Lambda_1,[\Lambda_1,[\Lambda_1,\Lambda_1]_{\blacksquare}]_{\blacksquare}]_{\blacksquare}\Big\} \\
&= \tfrac{1}{4}\left(R_4 - [\Lambda_1,\Lambda_3]_{\blacksquare} - \tfrac{1}{3!}[\Lambda_1,[\Lambda_1,\Lambda_2]_{\blacksquare}]_{\blacksquare}\right) \\
&= \tfrac{1}{24}\Big(\mathsf{area}(1,\mathsf{area}(1,\mathsf{area}(1,2))) \otimes [1,[1,[1,2]]] + \mathsf{area}(1,\mathsf{area}(2,\mathsf{area}(1,2))) \otimes [1,[2,[1,2]]] \\
&\qquad + \mathsf{area}(2,\mathsf{area}(1,\mathsf{area}(1,2))) \otimes [2,[1,[1,2]]] + \mathsf{area}(2,\mathsf{area}(2,\mathsf{area}(1,2))) \otimes [2,[2,[1,2]]] \\
&\qquad - (1 \sqcup\!\sqcup\, \mathsf{area}(1,\mathsf{area}(1,2))) \otimes [1,[1,[1,2]]] - (2 \sqcup\!\sqcup\, \mathsf{area}(1,\mathsf{area}(1,2))) \otimes [2,[1,[1,2]]] \\
&\qquad - (1 \sqcup\!\sqcup\, \mathsf{area}(2,\mathsf{area}(1,2))) \otimes [1,[2,[1,2]]] - (2 \sqcup\!\sqcup\, \mathsf{area}(2,\mathsf{area}(1,2))) \otimes [2,[2,[1,2]]]\Big)
\end{aligned}
$$

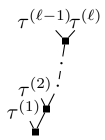
Remark 4.3.7. Comparing with [Roc03b, Section 3, page 322] we note that we correct some of the coefficients appearing in Λ_3 and Λ_4 there.

Definition 4.3.8. *Let $\widetilde{\mathsf{BPT}}_n$ be binary planar trees, with two types of inner nodes, $\bullet$ and $\blacksquare$, and such that the subset of all $\blacksquare$ nodes is either empty or forms a subtree with the same root as the tree itself. In other words the square nodes are all connected to the root.*

Define $e : \widetilde{\mathsf{BPT}}_n \to \mathbb{R}$ as follows. If the root of τ is $\bullet$, then

$$e(\tau) := \frac{1}{nc(\tau)},$$

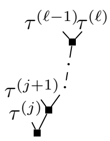
where c was defined in Lemma 4.2.12. Otherwise, we can write τ uniquely as

$\tau =$  $= (\tau^{(1)} \to_{\blacksquare} (\tau^{(2)} \to_{\blacksquare} (\cdots \to_{\blacksquare} (\tau^{(\ell-1)} \to_{\blacksquare} \tau^{(\ell)})))),$

for some $\ell = \ell(\tau) \geq 2$, $\tau^{(1)}, \ldots, \tau^{(\ell-1)} \in \widetilde{\mathsf{BPT}}$ and $\tau^{(\ell)} \in \mathsf{BPT}$. Here $\sigma \to_{\blacksquare} \rho$ is the grafting, to a new root of type $\blacksquare$, with σ on the left and ρ on the right. Then

$$e(\tau) := -\sum_{j=2}^{\ell(\tau)} \frac{|\tau^{(\geq j)}|_{\mathsf{leaves}}\, e(\tau^{(\geq j)})}{j!\,|\tau|_{\mathsf{leaves}}} \left(\prod_{i=1}^{j-1} e(\tau^{(i)})\right),$$

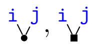
where

$\tau^{(\geq j)} :=$  $= (\tau^{(j)} \to_{\blacksquare} (\tau^{(j+1)} \to_{\blacksquare} (\cdots \to_{\blacksquare} (\tau^{(\ell-1)} \to_{\blacksquare} \tau^{(\ell)})))), \qquad j = 1, \ldots, \ell - 1$

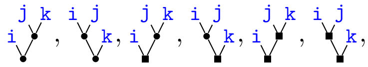
$\tau^{(\geq \ell)} := \tau^{(\ell)}.$

Finally, for a tree $\tau \in \widetilde{\mathsf{BPT}}_n$ *and a word* w *of length* n*, define* $\widetilde{\mathsf{area}}_\bullet(\tau)$ *as bracketing out using* area *if a node of type* $\bullet$ *is encountered and multiplying using* $\sqcup\!\sqcup$ *when a node* $\blacksquare$ *is encountered.*

Example 4.3.9. *The trees in* $\widetilde{\mathsf{BPT}}_2$ *are*



for letters $\mathtt{i}$ *and* $\mathtt{j}$*, and the trees in* $\widetilde{\mathsf{BPT}}_3$ *are*



for letters $\mathtt{i}$, $\mathtt{j}$ *and* $\mathtt{k}$.
We have

$$
\begin{aligned}
e(2) &= c(2) = 1\\
e({}^{2\,3}_{\ \blacksquare}) &= -\frac{1}{2\cdot 2} e(2)|3|_{\mathsf{leaves}} e(3) = -\frac{1}{4}\\
e({}_{1}{}^{2\,3}_{\ \blacksquare\,\blacksquare}) &= -\frac{1}{3}\left(\frac{1}{2!}e(1)|{}^{2\,3}_{\ \blacksquare}|_{\mathsf{leaves}} e({}^{2\,3}_{\ \blacksquare}) + \frac{1}{3!}e(2)e(2)|3|_{\mathsf{leaves}} e(3)\right) = -\frac{1}{3}\left(-\frac{1}{4}+\frac{1}{6}\right) = \frac{1}{36}.
\end{aligned}
$$

And

$$\widetilde{\mathsf{area}}_\bullet({}_{1}{}^{2\,3}_{\ \bullet\,\blacksquare}) = 1 \sqcup\!\sqcup \mathsf{area}(2,3).$$

Theorem 4.3.10. *Then*

$$\Lambda_n = \sum_{\tau \in \widetilde{\mathsf{BPT}}_n} e(\tau)\ \widetilde{\mathsf{area}}_\bullet(\tau) \otimes \mathsf{lie}_\bullet(\tau)\ \in \mathfrak{L} := \langle \mathfrak{R}; [\cdot,\cdot]_\blacksquare \rangle\,.$$

Proof. Define

$$
\begin{aligned}
\tau \in \widetilde{\mathsf{BPT}}_{n;\geq i} &:= \{\tau \in \widetilde{\mathsf{BPT}}_n \mid \ell(\tau) \geq i\},\\
[x_1,\dots,x_n] &:= [x_1,[\dots,[x_{n-1},x_n]\dots], \quad [x_1,x_2] := [x_1,x_2], \quad [x] := x.
\end{aligned}
$$

Due to $e(\mathtt{i}) = c(\mathtt{i}) = 1$ for all letters $\mathtt{i}$, we have

$$\Lambda_1 = R_1 = \sum_{\mathtt{i}=1}^{\mathtt{d}} \mathtt{i} \otimes \mathtt{i} = \sum_{\tau \in \mathsf{BPT}_1} e(\tau)\widetilde{\mathsf{area}}_\bullet(\mathtt{i}) \otimes \mathsf{lie}_\bullet(\mathtt{i}),$$

and then via induction over n

$$\Lambda_n = \frac{1}{n}R_n - \frac{1}{n}\sum_{i=2}^{n}\frac{1}{i!}\sum_{\substack{n_1,\dots,n_i\\ n_1+\dots+n_i=n}} n_i\,[\Lambda_{n_1},\dots,\Lambda_{n_i}]_{\blacksquare}$$
$$= \sum_{\tau\in\mathsf{BPT}_n}\frac{1}{nc(\tau)}\mathsf{area}_{\bullet}(\tau)\otimes\mathsf{lie}_{\bullet}(\tau)$$
$$-\frac{1}{n}\sum_{i=2}^{n}\frac{1}{i!}\sum_{\substack{n_1,\dots,n_i\\ n_1+\dots+n_i=n}}\sum_{\substack{\tau_1,\dots\tau_i\\ \tau_j\in\widetilde{\mathsf{BPT}}_{n_j}}}|\tau_i|_{\mathsf{leaves}}\prod_{j=1}^{i}e(\tau_j)\widetilde{\mathsf{area}}_{\bullet}(\tau_1)\sqcup\!\sqcup\cdots\sqcup\!\sqcup\,\widetilde{\mathsf{area}}_{\bullet}(\tau_i)\otimes[\mathsf{lie}_{\bullet}(\tau_1),\dots,\mathsf{lie}_{\bullet}(\tau_i)]$$
$$= \sum_{\tau\in\mathsf{BPT}_n} e(\tau)\,\mathsf{area}_{\bullet}(\tau)\otimes\mathsf{lie}_{\bullet}(\tau)$$
$$-\frac{1}{n}\sum_{i=2}^{n}\frac{1}{i!}\sum_{\tau\in\widetilde{\mathsf{BPT}}_{n;\geq i}}|\tau^{(\geq i)}|_{\mathsf{leaves}}e(\tau^{(\geq i)})\prod_{j=1}^{i-1}e(\tau^{(j)})\,\widetilde{\mathsf{area}}_{\bullet}(\tau)\otimes\mathsf{lie}_{\bullet}(\tau)$$
$$= \sum_{\tau\in\mathsf{BPT}_n} e(\tau)\,\mathsf{area}_{\bullet}(\tau)\otimes\mathsf{lie}_{\bullet}(\tau) - \sum_{\tau\in\widetilde{\mathsf{BPT}}_n^{\geq 2}}\sum_{i=2}^{\ell(\tau)}\frac{|\tau^{(\geq i)}|_{\mathsf{leaves}}e(\tau^{(\geq i)})}{i!|\tau|_{\mathsf{leaves}}}\prod_{j=1}^{i-1}e(\tau^{(k)})\,\widetilde{\mathsf{area}}_{\bullet}(\tau)\otimes\mathsf{lie}_{\bullet}(\tau)$$
$$= \sum_{\tau\in\widetilde{\mathsf{BPT}}_n} e(\tau)\,\widetilde{\mathsf{area}}_{\bullet}(\tau)\otimes\mathsf{lie}_{\bullet}(\tau). \qquad \square$$

Remark 4.3.11. Recall, from Theorem 4.2.8,

$$\mathfrak{R} = \langle \mathtt{i}\otimes\mathtt{i},\ \mathtt{i}=\mathtt{1}\dots\mathtt{d}; \trianglerighteq_{\mathrm{Sym}}\rangle\,.$$

Define

$$\mathfrak{P} := \langle \mathtt{i}\otimes\mathtt{i},\ \mathtt{i}=\mathtt{1}\dots\mathtt{d}; \trianglerighteq\rangle$$
$$\mathfrak{D} := \langle \mathtt{i}\otimes\mathtt{i},\ \mathtt{i}=\mathtt{1}\dots\mathtt{d}; \succeq,\preceq\rangle\,.$$

Then, we have $S_n, R_n, \Lambda_n \in \mathfrak{D}$, and the chain of inclusions

$$\mathfrak{R} \subsetneq \mathfrak{L} \subseteq \mathfrak{P} \subsetneq \mathfrak{D}.$$

Indeed, the mere inclusions are clear since $\trianglerighteq_{\mathrm{Sym}}$ and $[\cdot,\cdot]_{\blacksquare}$ are symmetrization and antisymmetrization of $\trianglerighteq$, and $\trianglerighteq$ itself is defined as a combination of $\succeq$ and $\preceq$. Regarding the strictness of two of the inclusions, on the one hand for any $d\geq 2$, the only anagram axis of $\mathtt{12}\otimes\mathtt{12}$ contained in $\mathfrak{R}$ is spanned by

$$(\mathtt{1}\otimes\mathtt{1})\trianglerighteq_{\mathrm{Sym}}(\mathtt{2}\otimes\mathtt{2}) = (\mathtt{2}\otimes\mathtt{2})\trianglerighteq_{\mathrm{Sym}}(\mathtt{1}\otimes\mathtt{1}) = \mathsf{area}(\mathtt{1},\mathtt{2})\otimes[\mathtt{1},\mathtt{2}] = (\mathtt{12}-\mathtt{21})\otimes(\mathtt{12}-\mathtt{21}),$$

and thus the $\mathfrak{L}$ element

$$[\mathtt{1}\otimes\mathtt{1},\mathtt{2}\otimes\mathtt{2}]_{\blacksquare} = (\mathtt{1}\sqcup\!\sqcup\,\mathtt{2})\otimes[\mathtt{1},\mathtt{2}] = (\mathtt{12}+\mathtt{21})\otimes(\mathtt{12}-\mathtt{21})$$

is not contained in $\mathfrak{R}$. On the other hand, the anagram space of $\mathtt{12}\otimes\mathtt{12}$ in $\mathfrak{P}$ is spanned by the two vectors

$$(\mathtt{1}\otimes\mathtt{1})\trianglerighteq(\mathtt{2}\otimes\mathtt{2}) = (\mathtt{1}\succ\mathtt{2})\otimes[\mathtt{1},\mathtt{2}] = \mathtt{12}\otimes(\mathtt{12}-\mathtt{21}),$$
$$(\mathtt{2}\otimes\mathtt{2})\trianglerighteq(\mathtt{1}\otimes\mathtt{1}) = (\mathtt{2}\succ\mathtt{1})\otimes[\mathtt{2},\mathtt{1}] = -\mathtt{21}\otimes(\mathtt{12}-\mathtt{21}),$$

and is thus easily seen to not contain the $\mathfrak{D}$ element

$$(1 \otimes 1) \succeq (2 \otimes 2) = (1 \succ 2) \otimes (1 \bullet 2) = 12 \otimes 12.$$

However, it remains an open problem whether $\mathfrak{L}$ and $\mathfrak{P}$ coincide.

Finally, we note the inclusion $\mathfrak{D} \subseteq \mathfrak{A}$, where $(\mathfrak{A}, \succeq, \preceq)$ is the dendriform algebra with linear basis given by all $w \otimes v$ such that w is a word and v is an anagram of w, and leave as a further question for future work whether $\mathfrak{D}$ and $\mathfrak{A}$ actually coincide.

Since the expansion in this theorem is not in terms of a *basis* of the Lie algebra, these are not yet coordinates of the first kind. But, by a straightforward projection procedure we get

Corollary 4.3.12 (based on [Roc03b, Theorem 1 and Corollary 1])**.**

$$S = \exp_{\blacksquare}(\Lambda),$$

with

$$\Lambda = \sum_h \zeta_h \otimes P_h,$$

where h runs over Hall words, P_h are the corresponding Lie Hall basis elements, and the ζ_h are expressed as linear combinations of shuffles of areas-of-areas,

$$\zeta_h = \sum_{\substack{\tau \in \widetilde{\mathsf{BPT}}_{|h|}, \\ \textit{foliage of } \tau \in \mathsf{Anagrams}(h)}} e(\tau) \Big\langle P_h^*, \mathsf{lie}_\bullet(\tau) \Big\rangle \, \widetilde{\mathsf{area}_\bullet}(\tau).$$

Remark 4.3.13. Again, this result is not satisfying because the ζ_h are expensive to calculate due to the large number of summands, which are not even linearly independent. We mention it only for completeness.

Proof of Corollary 4.3.12. Let P_h be Lie basis and P_h^* its dual basis. Then

$$\begin{aligned} \Lambda &= \sum_h \sum_{\tau \in \widetilde{\mathsf{BPT}}_n} e(\tau) \; \widetilde{\mathsf{area}_\bullet}(\tau) \langle P_h^*, \mathsf{lie}_\bullet(\tau) \rangle \otimes P_h \\ &=: \sum_h \zeta_h \otimes P_h, \end{aligned}$$

where P_h^* can be expressed as an element of $T(\mathbb{R}^d)$ which is a linear combination of anagrams of h, thus $\langle P_h^*, \mathsf{lie}_\bullet(\tau) \rangle = 0$ if the foliage of τ is not an anagram of h. □

4.4 Shuffle generators

For a countable index set I consider the free commutative algebra $\mathbb{R}[x_i : i \in I]$ over the indeterminates $x_i, i \in I$ ([Row88, Definition 1.2.12]). If V is a vector space with a countable basis, we also write $\mathbb{R}[V]$ for $\mathbb{R}[x_i : i \in I]$ where I is some basis of V. A commutative algebra $\mathcal{A}$ is **generated** by some elements $z_i \in \mathcal{A}$, $i \in I$, if the commutative algebra morphism

$$\mathbb{R}[x_i : i \in I] \to \mathcal{A},$$

extended from $x_i \mapsto z_i$, is surjective. If it is also injective, the algebra is **freely generated** by the elements z_i. The goal of this section is to find a simple condition on a countable family $z_i \in T(\mathbb{R}^d), i \in I$, to be (freely) generating.

Before stating the general results, let us begin with the example of the image of ρ.

Proposition 4.4.1. *Any basis for the image of* $\operatorname{Im}\rho$ *is generating. More explicitly, for any non-empty word* w*, we have*

$$w = \sum_{\substack{w_1,\dots,w_n \\ w_1\cdots w_n = w}} \frac{1}{k_{|w_1|,\dots,|w_n|}} \rho(w_1) \sqcup\!\sqcup \cdots \sqcup\!\sqcup \rho(w_n), \tag{4.20}$$

where $k_{m_1,\dots,m_n} = (m_1 + \cdots + m_n) k_{m_2,\dots,m_n}$*, with* $k_m = m$*.*

Proof. For any letter $\mathtt{i}$, we have $\mathtt{i} = \rho(\mathtt{i})$ in accordance with Equation (4.20). Assume the equation holds for all non-empty words v with $|v| \le \ell$ for some $\ell \ge 1$, and let w be a word with $|w| = \ell + 1$. Then, by Equation (4.6) we have

$$\begin{aligned} |w|w = Dw = \sum_{uv=w} \rho(u) \sqcup\!\sqcup v = \sum_{uv=w} \rho(u) \sqcup\!\sqcup \sum_{\substack{v_1,\dots,v_n \\ v_1\cdots v_n = v}} \frac{1}{k_{|v_1|,\dots,|v_n|}} \rho(v_1) \sqcup\!\sqcup \dots \sqcup\!\sqcup \rho(v_n) \\ = \sum_{\substack{w_1,\dots,w_n \\ w_1\cdots w_n = w}} \frac{1}{k_{|w_2|,\dots,|w_n|}} \rho(w_1) \sqcup\!\sqcup \dots \sqcup\!\sqcup \rho(w_n), \end{aligned}$$

again in accordance with Equation (4.20).

Note that in order for the induction to work, we made use again of the fact that $\rho(e) = 0$, so we only sum over non-empty words. ☐

Lemma 4.4.2. *For each* $n \ge 1$*, let* $X_n \subset T_n(\mathbb{R}^d)$ *be a subset of the shuffle algebra at level* n*. Let* $X := \bigcup_{n\ge1} X_n$*. Then:*

For all $n \ge 1$*, for all nonzero* $L \in \mathfrak{g}_n(\mathbb{R}^d)$ *there is an* $x \in X_n$ *such that* $\langle x, L\rangle \ne 0$

if and only if

X *generates the shuffle algebra* $T(\mathbb{R}^d)$*.*

If moreover $|X_n| = \dim \mathfrak{g}_n(\mathbb{R}^d)$*,* $n \ge 1$*, then* X *is* freely *generating.*

Remark 4.4.3. This lemma can also be seen as a consequence of (the proof of) the Cartier-Milnor-Moore theorem, see for example [Car07, Section 3.8]. Let us sketch this. Let $T(\mathbb{R}^d)^{*\,\mathrm{gr}}$ be the graded dual of $T(\mathbb{R}^d)$, the subspace of $T((\mathbb{R}^d))$ consisting of only finite linear combinations of words. Endowed with the unshuffle coproduct, the dual of the shuffle product, this is a cocommutative, conilpotent coalgebra. Then, by the proof of [Car07, Theorem 3.8.1], there exists an isomorphism of cocommutative coalgebras

$$e_{T(\mathbb{R}^d)^{*\,\mathrm{gr}}} : \Gamma[\mathfrak{g}] \to T(\mathbb{R}^d)^{*\,\mathrm{gr}}.$$

Here $\Gamma[\mathfrak{g}] \subset T(\mathfrak{g})$ are the symmetric tensors over $\mathfrak{g}$, generated, as a vector space, by the elements $\underbrace{v \otimes \cdots \otimes v}_{n \text{ times}}$, $v \in \mathfrak{g}$, $n \ge 0$, and endowed with the deconcatenation coproduct. The map $e_{T(\mathbb{R}^d)^{*\,\mathrm{gr}}}$

acts on these elements as

$$e_{T(\mathbb{R}^d)^{*\,\mathrm{gr}}}\left(\underbrace{v\otimes\cdots\otimes v}_{n\text{ times}}\right)=\frac{v^n}{n!},$$

where the n-th power on the right-hand side is taken with respect to the concatenation product (under which $T(\mathbb{R}^d)^{*\,\mathrm{gr}}$ is closed). The grading on $T(\mathbb{R}^d)^{*\,\mathrm{gr}}$ induces a grading on $\Gamma[\mathfrak{g}]$ via the isomorphism $e_{T(\mathbb{R}^d)^{*\,\mathrm{gr}}}$.

Let us now follow the end of the proof of [Car07, Theorem 3.8.3]. The graded dual (with respect to this induced grading) of $\Gamma[\mathfrak{g}]$ is given by $\mathbb{R}[\mathfrak{g}^{*\,\mathrm{gr}}]$, i.e. the symmetric algebra over $\mathfrak{g}^{*\,\mathrm{gr}}$, where $\mathfrak{g}^{*\,\mathrm{gr}}$ is the graded dual of $\mathfrak{g}$. Since $e_{T(\mathbb{R}^d)^{*\,\mathrm{gr}}}$ is an isomorphism of cocommutative coalgebras, the dual map

$$e^{*\,\mathrm{gr}}_{T(\mathbb{R}^d)^{*\,\mathrm{gr}}}:T(\mathbb{R}^d)\to\mathbb{R}[\mathfrak{g}^{*\,\mathrm{gr}}],$$

is an isomorphism of commutative algebras.

X (freely) generating $T(\mathbb{R}^d)$ is then equivalent to $e^{*\,\mathrm{gr}}_{T(\mathbb{R}^d)^{*\,\mathrm{gr}}}(X)$ (freely) generating $\mathbb{R}[\mathfrak{g}^{*\,\mathrm{gr}}]$, which is equivalent to our condition, using Lemma 4.8.1.

Proof. We show for every level N:

$$\forall n\le N\ \forall\ 0\ne L\in\mathfrak{g}_n(\mathbb{R}^d)\text{ there is }x\in X_n\text{ with }\langle x,L\rangle\ne 0$$
$$\textit{if and only if}$$
$$\bigcup_{1\le n\le N}X_n\text{ shuffle generates }T_{\le N}(\mathbb{R}^d).$$

It is clearly true for $N=1$. Let it be true for some N. We show it for $N+1$.

Let $\mathsf{shuff}_{N+1}\subset T_{N+1}(\mathbb{R}^d)$ denote the linear space of shuffles of everything "from below", i.e.

$$\mathsf{shuff}_{N+1}:=\bigcup_{n=1}^{N}\left\{T_n(\mathbb{R}^d)\sqcup\!\sqcup\, T_{N-n}(\mathbb{R}^d)\right\}.$$

By [Reu93, Theorem 3.1 (iv)]

$$\left\langle\mathsf{shuff}_{N+1},L\right\rangle=0,$$

for all $L\in\mathfrak{g}_{N+1}(\mathbb{R}^d)$. In other words, shuff_{N+1} is contained in the annihilator of $\mathfrak{g}_{N+1}(\mathbb{R}^d)$. By [Reu93, Theorem 6.1], the shuffle algebra is freely generated by the Lyndon words in $\mathtt{1},\dots,\mathtt{d}$, which have dimension $\dim\mathfrak{g}_n(\mathbb{R}^d)$ on level n. Hence

$$\dim\mathsf{shuff}_{N+1}=\dim T_{N+1}(\mathbb{R}^d)-\dim\mathfrak{g}_{N+1}(\mathbb{R}^d).$$

By dimension counting, we hence have that shuff_{N+1} must actually be *equal* to the annihilator of $\mathfrak{g}_{N+1}(\mathbb{R}^d)$. Then, a fortiori, $\mathfrak{g}_{N+1}(\mathbb{R}^d)$ is the annihilator of shuff_{N+1}.

By Lemma 4.4.4,

$$T_{N+1}(\mathbb{R}^d)=\mathsf{shuff}_{N+1}+\operatorname{span}_{\mathbb{R}}X_{N+1}$$
$$\textit{if and only if}$$
$$\forall\ 0\ne L\in\mathfrak{g}_{N+1}(\mathbb{R}^d)\text{ there is }x\in\operatorname{span}_{\mathbb{R}}X_{N+1}\text{ with }\langle x,L\rangle\ne 0.$$

But this is the case if and only if $\forall\, 0 \neq L \in \mathfrak{g}_{N+1}(\mathbb{R}^d)$ there is $x \in X_{N+1}$ with $\langle x, L\rangle \neq 0$. This finishes the proof regarding the generating property.

Regarding freeness: denote $\iota : \mathbb{R}[x_v : v \in X] \to T(\mathbb{R}^d)$ the extension, as a commutative algebra morphism, of the map $x_v \mapsto v$. Denote $\iota_{Lyndon} : \mathbb{R}[y_w : w \in L] \to T(\mathbb{R}^d)$ the extension, as a commutative algebra morphism, of the map $y_w \mapsto w$, where L are the Lyndon words. By [Reu93, Theorem 6.1], ι_{Lyndon} is an isomorphism. By what we have shown so far, ι is surjective. Since X consists of homogeneous elements, we can grade $\mathbb{R}[x_v : v \in X]$ induced from the grading of $T(\mathbb{R}^d)$ and analogously for $\mathbb{R}[y_w : w \in L]$. By assumption, the graded dimensions match. Hence, there is an isomorphism of graded, commutative algebras

$$\Phi : \mathbb{R}[x_v : v \in X] \to \mathbb{R}[y_w : w \in L].$$

Since ι_{Lyndon} is an isomorphism of graded, commutative algebras and ι is an epimorphism of graded, commutative algebras (where each homogeneous subspace is finite dimensional!) we must have that ι is in fact an isomorphism. □

We used the following simple lemma.

Lemma 4.4.4. *Let V be a finite dimensional vector space with dual $W := V^*$. We denote the pairing by $\langle w, v\rangle$, for $w \in W, v \in V$. Let W_1, W_2 be subspaces of W and let*

$$W_1^\perp := \{v \in V : \langle w_1, v\rangle = 0 \ \forall w_1 \in W_1\},$$

be the annihilator of W_1. Then:

$$\forall\, 0 \neq v_1 \in W_1^\perp \text{ there is } w_2 \in W_2 \text{ with } \langle w_2, v_1\rangle \neq 0$$
$$\text{if and only if}$$
$$W_1 + W_2 = W.$$

Proof. Recall the well-known identity (e.g. [Hal17, Exercise 8.(c) of Section 17])

$$(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp.$$

Then

$$W_1 + W_2 = W \Leftrightarrow W_1^\perp \cap W_2^\perp = \{0\},$$

which is the claim. □

Corollary 4.4.5. *Let X be a set of homogeneous elements of $T(\mathbb{R}^d)$. Then, the following are equivalent:*

(i) X freely shuffle generates $T(\mathbb{R}^d)$,

(ii) X is a homogeneous realization of a dual basis to a homogeneous basis of $\mathfrak{g}(\mathbb{R}^d)$,

(iii) X is a homogeneous basis for the image of a projection $\pi^\top$, where π is a graded projection $\pi : T(\mathbb{R}^d) \to \mathfrak{g}(\mathbb{R}^d) \subset T(\mathbb{R}^d)$ [7].

Examples include:

[7] Identifying $\mathfrak{g}(\mathbb{R}^d)$ as a subset of $T(\mathbb{R}^d)$.

1. *$\pi := \pi_1$, the Eulerian idempotent* (4.17)

   ⇝ *Coordinates of the first kind.*

2. *A rescaling of the Dynkin map r* (4.3) *(to make it a projection)*

   ⇝ *A basis for the image of ρ, for example $\mathfrak{r}_h$ from Theorem 4.2.14, which by Corollary 4.2.8 can be expressed as areas-of-areas.*

3. *π the orthogonal projection (with respect to the inner product in the ambient space $T(\mathbb{R}^d)$) onto $\mathfrak{g}(\mathbb{R}^d)$ (the* Garsia idempotent*, see [Duc91] where formulas for the idempotent of multilinear elements are given in Proposition 5.1 and Theorem 6.3, and [PRS05])*

   ⇝ *Any (homogeneous) basis for the Lie algebra $\mathfrak{g}(\mathbb{R}^d) \subset T(\mathbb{R}^d)$, identified as elements of $T(\mathbb{R}^d)$.*

Remark 4.4.6. 1. Point 3. is shown in [Reu93, Section 6.5.1], We include it here, as it falls nicely into the setting of Lemma 4.4.2.

2. Coordinates of the first kind must - by definition - contain all the information of the signature, so it is reasonable that they shuffle generate $T(\mathbb{R}^d)$. For the other sets this is not immediately evident. The basis for the Lie algebra is one such example and it does not even live in the correct space (formally, it is an element of the concatenation algebra $T((\mathbb{R}^d))$ not of the shuffle algebra $T(\mathbb{R}^d)$).

3. All of the above examples of Lie idempotents can be expressed in terms of an action of a symmetric group algebra element (for the Garsia idempotent, this fact is [Duc91, Theorem 4.1 (v)]), and are thus part of the algebra introduced in [PR02a].

Proof. (iii)⇒(i): Assume first that we have given a graded projection π with image $\mathfrak{g}(\mathbb{R}^d)$, and a homogeneous basis $(x_i)_i$ of $\operatorname{Im}\pi^\top$. Because of the grading it makes sense to speak of the component $\pi_n : T_n((\mathbb{R}^d)) \to T_n((\mathbb{R}^d))$. Then $\pi_n^\top : T_n(\mathbb{R}^d) \to T_n(\mathbb{R}^d)$ and

$$\operatorname{im}\left(\pi_n^\top\right)^\perp = \ker\left(\pi_n\right).$$

Since π_n itself is also a projection, we have that

$$T_n(\mathbb{R}^d) = \ker(\pi_n) \oplus \operatorname{im}(\pi_n).$$

Hence, for every $L \in \operatorname{im}(\pi)$ there is $x \in \operatorname{im}\left(\pi^\top\right)$ with $\langle x, L\rangle \neq 0$. Then Lemma 4.4.2 applies.

(i)⇒(ii): Let now X be a homogeneous free shuffle generating set. Then $\{\langle x, \cdot\rangle | x \in X_n\}$ spans the whole dual space of $\mathfrak{g}_n(\mathbb{R}^d)$. Indeed, assume this is not the case, then a comparison with some $\mathbb{R}^n$ shows that there is a nonempty annihilator of X_n inside $\mathfrak{g}_n(\mathbb{R}^d)$, but this contradicts the criterion from Lemma 4.4.2. Hence, X does span the dual space of $\mathfrak{g}(\mathbb{R}^d)$, thus contains a dual basis to some basis of $\mathfrak{g}(\mathbb{R}^d)$, and since X is freely generating, X is actually that dual basis (otherwise, a subset of X would already generate, which contradicts the assumption that X freely generates).

(ii)⇒(iii): Let $P_h \in T(\mathbb{R}^d), h \in H$, be some homogeneous basis for the Lie algebra. Let $D_h \in T(\mathbb{R}^d), h \in H$ be a realization of a dual basis. That is

$$\left\langle D_h, P_{h'} \right\rangle = \delta_{h,h'}.$$

Then choose π such that $\ker\pi = \left(\operatorname{span}_{\mathbb{R}}\{D_h : h \in H\}\right)^\top$. □

Proposition 4.4.7. *If X is a homogeneous set and $\pi : T(\mathbb{R}^d) \to \mathfrak{g}(\mathbb{R}^d) \subset T(\mathbb{R}^d)$ is a graded projection, then $\pi^\top X$ is a shuffle generating set if and only if $\pi^\top X$ spans $\operatorname{Im}\pi^\top$ if and only if X is a shuffle generating set. If X is freely generating, then so is $\pi^\top X$.*

Proof. Since for any $x \in X$ and $p \in \mathfrak{g}(\mathbb{R}^d)$ we have

$$\langle \pi^\top x, p\rangle = \langle x, \pi p\rangle = \langle x, p\rangle,$$

the condition for being a shuffle generating set in Lemma 4.4.2 is fulfilled for X if and only if it is fulfilled for $\pi^\top X$. Since any basis of $\operatorname{Im}\pi^\top$ is a free and thus also a minimal shuffle generating set by Corollary 4.4.5, $\pi^\top X \subseteq \operatorname{Im}\pi^\top$ shuffle generates if and only if it linearly spans $\operatorname{Im}\pi^\top$. If X freely shuffle generates, then the shuffle generating set $\pi^\top X$ must also have minimal dimension for each homogeneity, and thus freely generate due to the freeness of the shuffle algebra. □

Point 3.2 in Corollary 4.4.5 proves, using Corollary 4.2.8, what we set out to prove: areas-of-areas do shuffle generate $T(\mathbb{R}^d)$.

Corollary 4.4.8. *The set $\mathscr{A}$ of the Introduction is a generating set for $T(\mathbb{R}^d)$. A free generating set is given e.g. by any basis for the image of ρ.*

Proof. See Corollary 4.4.5, Point 2. □

Remark 4.4.9. Corollary 4.4.8 is an a priori stronger statement than the following easy-to-prove statement, with which it is occasionally confused.

> *(A) Any word is a linear combination of shuffles of letters and* area*s of arbitrary words.*

An illustration of (A) is as follows.

$$\begin{aligned} 123 &= (1 \succ 2) \succ 3 = \tfrac{1}{2}\{1 \sqcup\!\sqcup 2 + \mathsf{area}(1,2)\} \succ 3 \\ &= \tfrac{1}{4}\big[\{1 \sqcup\!\sqcup 2 + \mathsf{area}(1,2)\} \sqcup\!\sqcup 3 + \mathsf{area}(\{1 \sqcup\!\sqcup 2 + \mathsf{area}(1,2)\}, 3)\big] \\ &= \tfrac{1}{4}\big[1 \sqcup\!\sqcup 2 \sqcup\!\sqcup 3 + \mathsf{area}(1,2) \sqcup\!\sqcup 3 + \mathsf{area}(1 \sqcup\!\sqcup 2, 3) + \mathsf{area}(\mathsf{area}(1,2),3)\big] \end{aligned}$$

Corollary 4.4.8 implies that this can be done with all the shuffles *outside* all the areas, namely

> *(B) Any word is a linear combination of shuffles of letters and iterated* area*s of letters.*

For example

$$\begin{aligned} 123 = \tfrac{1}{3}\mathsf{area}(1, \mathsf{area}(2,3)) + \tfrac{1}{6}\mathsf{area}(\mathsf{area}(1,3),2) + \tfrac{1}{3} 1 \sqcup\!\sqcup \mathsf{area}(2,3) \\ - \tfrac{1}{6} 2 \sqcup\!\sqcup \mathsf{area}(1,3) + \tfrac{1}{2} 3 \sqcup\!\sqcup \mathsf{area}(1,2) + \tfrac{1}{6} 1 \sqcup\!\sqcup 2 \sqcup\!\sqcup 3. \end{aligned}$$

4.5 Applications

The antisymmetrizing feature of the area operation leads to pleasant properties for piecewise linear paths and semimartingales.

4.5.1 Piecewise linear paths: computational aspects

For two time series $a_0, \dots, a_n, b_0, \dots, b_n \in \mathbb{R}$ define the new time series

$$\begin{aligned}\mathsf{DiscreteArea}(a,b)_\ell &:= \mathrm{Corr}_1(a,b)_\ell - \mathrm{Corr}_1(b,a)_\ell \\ &:= \sum_{i=0}^{\ell-1} a_{i+1}b_i - \sum_{i=0}^{\ell-1} b_{i+1}a_i, \quad \ell = 0,\dots,n,\end{aligned}$$

set to be 0 for $\ell = 0$. It is known ([DR19, Section 3.2]), that for a piecewise linear curve X through the points $0, x_1, \dots, x_n \in \mathbb{R}^2$, one has

$$\Big\langle \mathsf{area}(1,2), S(X)_{0,n} \Big\rangle = \mathsf{DiscreteArea}\left(x^1, x^2\right)_n. \tag{4.21}$$

We will show that this iterates nicely.

Lemma 4.5.1. *If X, Y are piecewise linear then $\mathsf{Area}(X,Y)$ is piecewise linear.*

Proof.

$$\begin{aligned}\frac{d}{dt}\mathsf{Area}(X,Y)_t &= \int_0^t dX_r\, \dot{Y}_t - \int_0^t dY_r\, \dot{X}_t \\ \frac{d^2}{dt^2}\mathsf{Area}(X,Y)_t &= \int_0^t dX_r\, \ddot{Y}_t + \dot{X}_t\dot{Y}_t - \int_0^t dY_r\, \ddot{X}_t - \dot{Y}_t\dot{X}_t \\ &= \int_0^t dX_r\, \ddot{Y}_t - \int_0^t dY_r\, \ddot{X}_t = 0,\end{aligned}$$

since X, Y are piecewise linear. Hence, $\mathsf{Area}(X,Y)$ is indeed piecewise linear. □

In particular, for $\phi \in \mathscr{A}$ (defined in the Introduction) and X piecewise linear

$$t \mapsto \Big\langle \phi, S(X)_{0,t} \Big\rangle,$$

is piecewise linear. Note that by Lemma 4.6.2, ϕ can be written as linear combination of elements of the form $w(ij - ji)$. One can also see directly that such elements yield something piecewise linear:

$$\begin{aligned}&\frac{d^2}{dt^2}\Big\langle w(ij-ji), S(X)_{0,t} \Big\rangle \\ &= \frac{d^2}{dt^2}\left\{ \int_0^t \int_0^s \Big\langle w, S(X)_{0,r} \Big\rangle dX^{(i)}_r dX^{(j)}_s - \int_0^t \int_0^s \Big\langle w, S(X)_{0,r} \Big\rangle dX^{(j)}_r dX^{(i)}_s \right\} \\ &= \int_0^t \Big\langle w, S(X)_{0,r} \Big\rangle dX^{(i)}_r\, \ddot{X}^{(j)}_t + \Big\langle w, S(X)_{0,t} \Big\rangle \dot{X}^{(i)}_t \dot{X}^{(j)}_t \\ &\quad - \int_0^t \Big\langle w, S(X)_{0,r} \Big\rangle dX^{(j)}_r\, \ddot{X}^{(i)}_t - \Big\langle w, S(X)_{0,t} \Big\rangle \dot{X}^{(i)}_t \dot{X}^{(j)}_t \\ &= 0.\end{aligned}$$

Lemma 4.5.2. *For all nonzero $z \in T(\mathbb{R}^d)$, there is a piecewise linear path X such that $\langle z, S(X) \rangle \neq 0$.*

Proof. Let $z \in T(\mathbb{R}^d) \setminus \{0\}$ be arbitrary and let n be its degree (the length of the longest word in the word expansion of z). Since $G_{\leq n} := \mathsf{proj}_{\leq n} G$ spans $T_{\leq n}(\mathbb{R}^d)$ (see e.g. [DR19, Lemma 3.4]), there are $g_1, \dots, g_k \in G$ and $r_1, \dots, r_k \in \mathbb{R}$ such that

$$\langle z, r_1 g_1 + \dots + r_k g_k \rangle = \langle z, z \rangle \neq 0,$$

and hence there is $g_i \in G$ such that $\langle z, g_i \rangle \neq 0$. Now, due to Chow's theorem according to [FV10, Theorem 7.28], there is a piecewise linear X such that $\mathsf{proj}_{\leq n} g_i = \mathsf{proj}_{\leq n} S(X)$, which implies $\langle z, S(X) \rangle = \langle z, g_i \rangle \neq 0$. ☐

Theorem 4.5.3. *$\langle S(X)_{0,t}, \phi \rangle$ is piecewise linear for all piecewise linear paths X if and only if $\phi \in \mathbb{R} \oplus \mathscr{A}$.*

Proof. We already showed in Lemma 4.5.1 that for piecewise linear X, $\langle S(X)_{0,t}, \phi \rangle$ is again piecewise linear for all $\phi \in \mathscr{A}$. Since the whole tensor space $T(\mathbb{R}^d) = \mathbb{R} \oplus \mathscr{A} \oplus B$, where B is $\operatorname{span}_{\mathbb{R}}\{w\mathtt{ij},\ w \text{ a word},\ \mathtt{i} \leq \mathtt{j} \text{ letters}\}$, and since the sum of a function which is not piecewise linear with a piecewise linear function is again not piecewise linear, it only remains to show that for any $b \in B \setminus \{0\}$, there is a piecewise linear X such that $t \mapsto \langle b, S(X)_{0,t} \rangle$ is not piecewise linear.

To this end, let $b = \sum_{i \leq j} d_{ij} \mathtt{ij} \in B \setminus \{0\}$ be arbitrary. If there is a letter $\mathtt{l}$ such that $d_{ll} \neq 0$ (case 1), choose a piecewise linear path $X : [0,2] \to \mathbb{R}^d$ such that $\langle d_{ll}, S(X)_{0,1} \rangle \neq 0$ and such that $X\restriction_{[1,2]}$ is linear with $x_l = 1$ and $x_i = 0$ for $i \neq l$, where $x_i := \dot{X}^i_{3/2}$. Otherwise, since b is nonzero, there are letters $\mathtt{k} < \mathtt{l}$ such that $d_{kl} \neq 0$ (case 2), and in this case, choose a piecewise linear path $X : [0,2] \to \mathbb{R}^d$ such that $\langle d_{kl}, S(X)_{0,1} \rangle \neq 0$ and such that $X\restriction_{[1,2]}$ is linear with $x_k = 1, x_l = 1$ and $x_i = 0$ for $i \notin \{k, l\}$, where $x_i := \dot{X}^i_{3/2}$. In both cases, such a piecewise linear X exists due to Theorem 4.5.2.

Since $X_{[1,2]}$ is linear, we have for arbitrary $z \in T(\mathbb{R}^d$ that $t \mapsto \langle z, S(X)_{0,t} \rangle$ is polynomial on $[1,2]$, and thus arbitrarily often continuously differentiable on $(1,2)$. Thus, since $\ddot{X} = 0$ and $\dot{X}$ constant on $(1,2)$, we have

$$\lim_{t \searrow 1} \frac{d^2}{dt^2} \Big\langle \sum_{i \leq j} d_{ij} \mathtt{ij}, S(X)_{0,t} \Big\rangle = \sum_{i \leq j} \Big\langle d_{ij}, S(X)_{0,1} \Big\rangle x_i x_j = \begin{cases} \langle d_{ll}, S(X)_{0,1} \rangle \neq 0, & \text{case 1} \\ \langle d_{kl}, S(X)_{0,1} \rangle \neq 0, & \text{case 2} \end{cases}$$

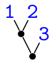
In both cases, we conclude that $t \mapsto \langle b, S(X)_{0,t} \rangle$ is not piecewise linear on any interval $[1, s]$, $1 < s \leq 2$, which finishes the proof. ☐

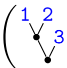
The fact that "being linear" is preserved under the Area-operation immediately leads to the following theorem.

Theorem 4.5.4. *Let X in $\mathbb{R}^d$ be a piecewise linear curve through the points $0, x_1, \dots, x_n \in \mathbb{R}^d$. Then: for every tree τ,*

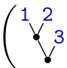
$$\Big\langle \mathsf{area}_{\bullet}(\tau), S(X)_{0,n} \Big\rangle = \mathsf{DiscreteArea}_{\bullet}(\tau, x)_n \,.$$

Here, area$_\bullet$ *is defined in Lemma 4.2.12 and* DiscreteArea$_\bullet$ *is defined similarly, as iterated bracketing using the* DiscreteArea*-operator.*

Example 4.5.5. *For* $\tau =$  *the statement reads as*

$$\left\langle \mathsf{area}_\bullet\left(\right.\right.$$  $$\left.\left.\right), S(X)_{0,n}\right\rangle := \left\langle \mathsf{area}\left(\mathsf{area}\left(1,2\right),3\right), S(X)_{0,n}\right\rangle$$

$$=$$

$$\mathsf{DiscreteArea}_\bullet\left(\right.$$  $$\left., x\right)_n := \mathsf{DiscreteArea}\left(\mathsf{DiscreteArea}\left(x^{(1)}, x^{(2)}\right), x^{(3)}\right)_n,$$

which one can verify by a direct, but tedious, calculation.

Remark 4.5.6. This is not obvious at all. Indeed, if we just look at the discrete integration operator (still assuming $x_0 = 0$)

$$\left\langle 12, S(X)_{0,n}\right\rangle = \sum_{i=0}^{n-1}\frac{1}{2}\left(x_i^1 + x_{i+1}^1\right)\left(x_{i+1}^2 - x_i^2\right) =: \mathsf{DiscreteIntegral}\left(x^1, x^2\right)_n,$$

this does *not* iterate. Indeed,

$$\left\langle 123, S(X)_{0,n}\right\rangle = \sum_{i=0}^{n-1}\left(\sum_{j=0}^{i-1}\frac{1}{2}\left(x_j^1 + x_{j+1}^1\right)x_{j,j+1}^2 + \left(\frac{1}{2}x_i^1 + \frac{1}{3!}x_{i,i+1}^1\right)x_{i,i+1}^2\right)x_{i,i+1}^3$$
$$\neq \mathsf{DiscreteIntegral}\left(\mathsf{DiscreteIntegral}\left(x^1, x^2\right), x^3\right)_n.$$

Proof. If Y, Z are piecewise linear between the points $0, y_1, \dots$ and $0, z_1, \dots$, then $\mathsf{Area}(Y, Z)$ is piecewise linear between the points

$$0, \mathsf{DiscreteArea}(y, z)_1, \dots, \mathsf{DiscreteArea}(y, z)_n.$$

We can hence iterate (4.21). □

4.5.2 Martingales and Martingaloids

Another pleasant property of the area operation presents itself when working with a continuous semimartingale M. One has (see e.g. [IW89, Chapter III, Equation (1.10) and Theorem 1.4])

$$\int_0^T M_{0,r}^i d_{\text{Strat}} M_r^j = \int_0^T M_{0,r}^i d_{\text{Itō}} M_r^j + \frac{1}{2}[M^i, M^j]_T. \tag{4.22}$$

where d_{Strat} denotes Fisk-Stratonovich integration, $d_{\text{Itō}}$ denotes Itō integration and $[.,.]$ denotes the quadratic covariation.

Proposition 4.5.7. *Let* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ *be a filtered probability space. Let* M *be a continuous,* $\mathcal{F}_t$*-martingale such that all iterated Itō integrals are martingales. Then* $t \mapsto \langle \phi, S_{\text{Strat}}(M)_{0,t}\rangle$ *is an* $\mathcal{F}_t$*-martingale for all* $\phi \in \mathscr{A}$.

Proof. Let X, Y be as in the statement. Assume for simplicity $X_0 = Y_0 = 0$ almost surely. Then, using Equation (4.22),

$$\mathsf{Area}_{\text{Strat}}(X, Y)_t := \int_0^t X_r d_{\text{Strat}} Y_r - \int_0^T Y_r d_{\text{Strat}} X_r = \int_0^t X_r d_{\text{Itō}} Y_r - \int_0^t Y_r d_{\text{Itō}} X_r,$$

is again an $(\mathcal{F}_t)_t$-martingale. Hence for every $\phi \in \mathscr{A}$, $\langle \phi, S_{\text{Strat}}(M)_{0,t}\rangle$ is a $\mathcal{F}_t$-martingale. □

Proposition 4.5.8. *Let M be the piecewise linear interpolation of a time discrete (local) martingale whose moments all exist. Then $t \mapsto \langle \phi, S(M)_{0,t} \rangle$ is again the piecewise linear interpolation of a time discrete (local) martingale.*

Proof. Let $(a_n)_n, (b_n)_n$ be two time discrete martingales for the filtration $\mathcal{F}_k$ whose moments all exist. Then,

$$\begin{aligned}
&\mathbb{E}[\mathsf{DiscreteArea}(a,b)_{k+1}|\mathcal{F}_k] - \mathsf{DiscreteArea}(a,b)_k \\
&\quad = \mathbb{E}[\mathsf{DiscreteArea}(a,b)_{k+1} - \mathsf{DiscreteArea}(a,b)_k|\mathcal{F}_k] = \mathbb{E}[a_{k+1}b_k - b_{k+1}a_k|\mathcal{F}_k] \\
&\quad = b_k\mathbb{E}[a_{k+1}|\mathcal{F}_k] - a_k\mathbb{E}[b_{k+1}|\mathcal{F}_k] = b_k a_k - a_k b_k = 0,
\end{aligned}$$

thus $(\mathsf{DiscreteArea}(a,b)_n)_n$ is again an $(\mathcal{F}_n)_n$ martingale whose moments all exist. If $(c_n)_n, (d_n)_n$ are τ_n local martingales whose moments all exits, then

$$\mathsf{DiscreteArea}(a,b)^{\tau_k} = \mathsf{DiscreteArea}(a^{\tau_k}, b^{\tau_k})^{\tau_k} = \mathsf{DiscreteArea}(a^{\tau_k}, b^{\tau_k})$$

is a martingale for any k and thus $\mathsf{DiscreteArea}(a,b)$ is a τ_k local martingale whose moments all exist. □

The previous results imply that for M a martingale with all iterated integrals being in $L^1(\Omega)$, or for M a linear interpolation of a time discrete martingale, all expectancies of areas of areas vanish for M. This naturally leads to the very interesting question of what is the class of all semimartingale paths such that all area expectancies vanish? For the moment, we feel justified to simply take it as a definition for further investigation.

Definition 4.5.9.

1. *We call a probability measure μ on $\mathcal{G}_d$ a* martingaloid measure *if $\int_{\mathcal{G}_d} \langle g, x \rangle \, \mathrm{d}\mu(g)$ exists and is equal to zero for all $x \in \mathscr{A}$.*

2. *We call a random map $\mathbf{X} : [0,T]^2 \times \Omega \to \mathcal{G}_d$ a* terminal p-variation martingaloid *(resp.* terminal γ-Hölder martingaloid*) if it is almost surely a continuous finite p-variation (resp. γ-Hölder) weakly geometric rough path and the law of $\mathbf{X}_{0T}$ is a martingaloid measure.*

3. *We call a random map $\mathbf{X} : [0,T]^2 \times \Omega \to \mathcal{G}_d$ a* p-variation martingaloid *(resp.* γ-Hölder martingaloid*) if it is almost surely a continuous finite p-variation (resp. γ-Hölder) weakly geometric rough path and the law of $\mathbf{X}_{st}$ is a martingaloid measure for all $0 \le s < t \le T$.*

In particular, for all these paths, we have, by Theorem 4.6.14, that the expected Stratonovich signature lies in the kernel of r,

$$\begin{aligned}
r(\mathbb{E}[S_{\mathrm{Strat}}(M)_{0,T}]) &= \mathbb{E}\left[r\left(S_{\mathrm{Strat}}(M)_{0,T}\right)\right] = \sum_w \mathbb{E}[\langle w, S_{\mathrm{Strat}}(M)_{0,T}\rangle]\, r(w) \\
&= \sum_w \mathbb{E}[\langle \rho(w), S_{\mathrm{Strat}}(M)_{0,T}\rangle]\, w = 0.
\end{aligned}$$

In fact this last property obviously holds for any Semimartingale with $\langle \rho(w), S_{\mathrm{Strat}} \rangle = 0$ for any w, which is a priori a larger class than just those paths where the expectancy of all areas of areas vanishes since $\operatorname{Im} \rho$ does not linearly span $\mathscr{A}$.

Proposition 4.5.10. *Let μ and ν be measures on $\mathcal{G}_d$ where μ has finite expected signature. Then if two out of the three measures $\mu, \nu, \mu \circledast \nu$ are martingaloid measures, so is the third. In particular, martingaloid measures of a fixed dimension with finite expected signature form a monoid under stochastic convolution. Its unit $\delta_{\mathbf{e}}$ is the only Dirac measure on $\mathcal{G}_d$ that is martingaloid.*

This result can be directly applied to rough paths: If $\mathbf{X}$ and $\mathbf{Y}$ are independent random rough paths where $\mathbf{X}$ has finite expected signature, then again if two out of the three $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{X} \sqcup \mathbf{Y}$ are martingaloids, so is the third. This Proposition also means that Lévy processes on $\mathcal{G}_d$ with finite expected signature are martingaloids if and only if they are terminal martingaloids.

Proof. Let g_1, g_2 be independent $\mathcal{G}_d$-valued random variables with g_1 having finite expected signature. We have for all homogeneous $x \in \mathscr{A}$, with Sweedler's notation $\sum_{(x)}^{\bullet} x' \otimes x'' = \Delta_{\bullet} x - x \otimes \mathbf{e} - \mathbf{e} \otimes x$ for the reduced coproduct,

$$\mathbb{E}[\langle g_1 \bullet g_2, x\rangle] = \mathbb{E}[\sum_{(x)}^{\bullet} \langle g_1, x_1\rangle\langle g_2, x_2\rangle] = \mathbb{E}[\langle g_1, x\rangle] + \mathbb{E}[\langle g_2, x\rangle] + a\sum_{(x)}^{\bullet} \mathbb{E}[\langle g_1, x'\rangle]\mathbb{E}[\langle g_2, x''\rangle],$$

provided that both all expectancies $\mathbb{E}[\langle g_2, x''\rangle]$ and $\mathbb{E}[\langle g_1 \bullet g_2, x\rangle]$ or $\mathbb{E}[\langle g_2, x\rangle]$ exist. Since all x'' in the sum can be expressed as sums of homogeneous terms of lower homogeneity than x and are by Corollary 4.6.4 elements of $\mathscr{A}$, it can be easily seen by induction over the homogeneity of x that all the area expectancies of g_1, g_2 and $g_1 \bullet g_2$ exist and vanish whenever g_1 and g_2, or g_1 and $g_1 \bullet g_2$, or g_2 and $g_1 \bullet g_2$ have vanishing area expectancies.

If we now assume that g_1 and g_2 are independent, have finite expected signature and vanishing area expectancies, then by the previous induction argument $g_1 \bullet g_2$ has vanishing area expectancies, and furthermore for all words w

$$\mathbb{E}[\langle g_1 \bullet g_2, w\rangle = \mathbb{E}[\sum_{(w)} \langle g_1, w_1\rangle\langle g_2, w_2\rangle] = \sum_{(w)} \mathbb{E}[\langle g_1, w_1\rangle]\mathbb{E}[\langle g_2, w_2\rangle]$$

exists, thus $g_1 \bullet g_2$ has finite expected signature. Together with $\delta_{\mathbf{e}}$ obviously being a unit, this shows the monoid property of the set of all martingaloid measures on $\mathcal{G}_d$ with finite expected signature.

Finally, for a Dirac measure δ_g on $\mathcal{G}_d$ the expected signature is just the single grouplike element g the measure is supported on, and if δ_g is martingaloid we have $\langle g, \mathscr{A}\rangle = \{0\}$ which by the main theorem of this chapter Corollary 4.4.8 means $g = \mathbf{e}$. ☐

4.6 Linear span of area expressions

In [Dzh07] the antisymmetric, non-associative operation area was studied in detail. It was shown in [Dzh07, Lemma 6.1 and 6.2] that

- area does not satisfy any new identity of arity 3; in particular it does *not* satisfy the Jacobi identity.
- In arity 4 there is exactly one new identity, the *Tortkara identity*.

Over a field of characteristic different from two, we have the following equivalent formulations of the Tortkara identity:

$$\begin{aligned}
&\mathsf{area}(\mathsf{area}(a,b),\mathsf{area}(c,b)) = \mathsf{area}(\mathsf{vol}(a,b,c),b),\\
&\mathsf{area}(\mathsf{area}(a,b),\mathsf{area}(c,d)) + \mathsf{area}(\mathsf{area}(a,d),\mathsf{area}(c,b)) = \mathsf{area}(\mathsf{vol}(a,b,c),d) + \mathsf{area}(\mathsf{vol}(a,d,c),b),\\
&2\cdot \mathsf{area}(\mathsf{area}(a,b),\mathsf{area}(c,d))\\
&\ = \mathsf{area}(\mathsf{vol}(a,b,c),d) + \mathsf{area}(\mathsf{vol}(a,d,c),b) + \mathsf{area}(\mathsf{vol}(b,a,d),c) + \mathsf{area}(\mathsf{vol}(b,c,d),a)
\end{aligned}$$

where $\mathsf{vol}(x,y,z) := \mathsf{area}(\mathsf{area}(x,y),z)+\mathsf{area}(\mathsf{area}(y,z),x)+\mathsf{area}(\mathsf{area}(z,x),y)$, and where the first identity is [DIM19, Equation (2)] while the second identity was the first version of the Tortkara identity presented in [Dzh07, Section 2 and 6] and can also be found in [DIM19, Equation (3)].

We chose the notation vol because $\langle \mathsf{vol}(u,v,w), S(X)_{0,T}\rangle$ is six times the signed volume ([DR19, Equation (4), Definition 3.27 and Theorem 3.28]) of the curve (U,V,W), where

$$U_t = \Big\langle u, S(X)_{0,t}\Big\rangle, \quad V_t = \Big\langle v, S(X)_{0,t}\Big\rangle, \quad W_t = \Big\langle w, S(X)_{0,t}\Big\rangle.$$

The Tortkara identity is readily verified on all of $T^{\geq 1}(\mathbb{R}^d)$ by computing

$$\begin{aligned}
\mathsf{area}(\mathsf{area}(1,2),\mathsf{area}(3,2)) &= -2\,1223 + 2\,1232 + 2\,2213 - 2\,2231 - 2\,3212 + 2\,3221\\
&= \mathsf{area}(\mathsf{vol}(1,2,3),2),
\end{aligned}$$

where ([DR19, Section 3.1])

$$\mathsf{vol}(1,2,3) = 123 - 132 - 213 + 231 + 312 - 321.$$

Indeed, this computation suffices to show the Tortkara identity on $T^{\geq 1}(\mathbb{R}^d)$ due to the universal property of the free Zinbiel algebra $(T^{\geq 1}(\mathbb{R}^3), \succ)$, i.e. for any $a,b,c \in T^{\geq 1}(\mathbb{R}^d)$, there is a unique Zinbiel homomorphism $(T^{\geq 1}(\mathbb{R}^3), \succ) \to (T^{\geq 1}(\mathbb{R}^d), \succ)$ with $1 \mapsto a$, $2 \mapsto b$, $3 \mapsto c$, and then the Tortkara identity follows for a,b,c from the above computation by the homomorphism property and the fact that area is nothing but the antisymmetrization of $\succ$.

Remark 4.6.1. In [Bre18], Bremner studies the *Tortkara triple product* which in our notation is given by

$$\mathsf{tri}(a,b,c) := \mathsf{area}(\mathsf{area}(a,b),c),$$

and thus vol relates to tri by being the sum of the cyclic permutations of its arguments,

$$\mathsf{vol}(a,b,c) = \mathsf{tri}(a,b,c) + \mathsf{tri}(b,c,a) + \mathsf{tri}(c,b,a).$$

Besides antisymmetry in the first two arguments $\mathsf{tri}(a,a,b) = 0$, Bremner found a new relation in arity 5 [Bre18, Theorem 5] and another new relation in arity 7 [Bre18, Theorem 7] for the Tortkara triple product of the free Zinbiel algebra.

Proof of equivalence of the Tortkara identities. Let $\mathscr{T}$ be a vector space over an arbitrary field with a bilinear antisymmetric operation t and

$$\mathsf{J}(x,y,z) := \mathsf{t}(\mathsf{t}(x,y),z) + \mathsf{t}(\mathsf{t}(y,z),x) + \mathsf{t}(\mathsf{t}(z,x),y).$$

1. Assume first that for all $x,y,z \in \mathscr{T}$ we have

$$\mathsf{t}(\mathsf{t}(x,y),\mathsf{t}(z,y)) = \mathsf{t}(\mathsf{J}(x,y,z),y).$$

Then, for all $a, b, c, d \in \mathscr{T}$, due to bilinearity, we have

$$\begin{aligned}&\mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,b)) + \mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,d)) + \mathsf{t}(\mathsf{t}(a,d),\mathsf{t}(c,b)) + \mathsf{t}(\mathsf{t}(a,d),\mathsf{t}(c,d))\\ &= \mathsf{t}(\mathsf{t}(a,b+d),\mathsf{t}(c,b+d)) = \mathsf{t}(\mathsf{J}(a,b+d,c),b+d)\\ &= \mathsf{t}(\mathsf{J}(a,b,c),b) + \mathsf{t}(\mathsf{J}(a,b,c),d) + \mathsf{t}(\mathsf{J}(a,d,c),b) + \mathsf{t}(\mathsf{J}(a,d,c),d).\end{aligned}$$

Since $\mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,b)) = \mathsf{t}(\mathsf{J}(a,b,c),b)$ and $\mathsf{t}(\mathsf{t}(a,d),\mathsf{t}(c,d)) = \mathsf{t}(\mathsf{J}(a,d,c),d)$, we obtain the identity

$$\mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,d)) + \mathsf{t}(\mathsf{t}(a,d),\mathsf{t}(c,b)) = \mathsf{t}(\mathsf{J}(a,b,c),d) + \mathsf{t}(\mathsf{J}(a,d,c),b)$$

for all $a, b, c, d \in \mathscr{T}$. Using antisymmetry, we furthermore get

$$\begin{aligned}&\mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,d)) + \mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,d))\\ &= \mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,d)) + \mathsf{t}(\mathsf{t}(a,d),\mathsf{t}(c,b)) + \mathsf{t}(\mathsf{t}(b,a),\mathsf{t}(d,c)) + \mathsf{t}(\mathsf{t}(b,c),\mathsf{t}(d,a))\\ &= \mathsf{t}(\mathsf{J}(a,b,c),d) + \mathsf{t}(\mathsf{J}(a,d,c),b) + \mathsf{t}(\mathsf{J}(b,a,d),c) + \mathsf{t}(\mathsf{J}(b,c,d),a)\end{aligned}$$

If the field is of characteristic different from two, this reads as

$$2\mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,d)) = \mathsf{t}(\mathsf{J}(a,b,c),d) + \mathsf{t}(\mathsf{J}(a,d,c),b) + \mathsf{t}(\mathsf{J}(b,a,d),c) + \mathsf{t}(\mathsf{J}(b,c,d),a)$$

for all $a, b, c, d \in \mathscr{T}$.

2. Assume now that for all $a, b, c, d \in \mathscr{T}$ we have

$$\mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,d)) + \mathsf{t}(\mathsf{t}(a,d),\mathsf{t}(c,b)) = \mathsf{t}(\mathsf{J}(a,b,c),d) + \mathsf{t}(\mathsf{J}(a,d,c),b).$$

This immediately implies

$$\mathsf{t}(\mathsf{t}(x,y),\mathsf{t}(z,y)) + \mathsf{t}(\mathsf{t}(x,y),\mathsf{t}(z,y)) = \mathsf{t}(\mathsf{J}(x,y,z),y) + \mathsf{t}(\mathsf{J}(x,y,z),y)$$

for all $x, y, z \in \mathscr{T}$, which is an empty statement in characteristic two, but in characteristic different from two reduces to

$$\mathsf{t}(\mathsf{t}(x,y),\mathsf{t}(z,y)) = \mathsf{t}(\mathsf{J}(x,y,z),y). \tag{4.23}$$

3. For the last implication we want to show, assume that the characteristic of the underlying field is different from two and for all $a, b, c, d \in \mathscr{T}$ we have

$$2\mathsf{t}(\mathsf{t}(a,b),\mathsf{t}(c,d)) = \mathsf{t}(\mathsf{J}(a,b,c),d) + \mathsf{t}(\mathsf{J}(a,d,c),b) + \mathsf{t}(\mathsf{J}(b,a,d),c) + \mathsf{t}(\mathsf{J}(b,c,d),a).$$

This implies

$$2\mathsf{t}(\mathsf{t}(x,y),\mathsf{t}(y,z)) = 2\mathsf{t}(\mathsf{J}(x,y,z),y) + \mathsf{t}(\mathsf{J}(y,x,y),z) + \mathsf{t}(\mathsf{J}(y,z,y),x) = 2\mathsf{t}(\mathsf{J}(x,y,z),y),$$

since due to antisymmetry

$$\mathsf{J}(y,x,y) = \mathsf{J}(y,z,y) = 0.$$

Since the characteristic is different from two, we can divide by two, and thus again arrive at (4.23).

□

In [DIM19, Section 5] it is shown that in $d = 2$, $(\mathscr{A}, \mathsf{area})$ is the free Tortkara algebra.[8]

This linear space has a surprisingly simple description. The following is [DIM19, Theorem 2.1] (for the proof see [DIM19, Section 3]; see also [Rei19, Section 3.2] for another proof in the special case $d = 2$).

Lemma 4.6.2.

$$\mathscr{A} = \operatorname{span}_{\mathbb{R}}\{\mathtt{i} : \mathtt{i} \textit{ a letter }\} \oplus \operatorname{span}_{\mathbb{R}}\{w(\mathtt{ij} - \mathtt{ji}) : w \textit{ a word}, \mathtt{i}, \mathtt{j} \textit{ letters}\}. \tag{4.24}$$

Example 4.6.3. *We have that [DR19, Equation (4)]*

$$\mathrm{Inv}_n := \sum_{\sigma \in S_n} \mathsf{sign}(\sigma)\, \sigma(\mathtt{1}) \cdots \sigma(\mathtt{i}),$$

where we interpret σ as a permutation of the letters, is in $\mathscr{A}$ for $d \geq n \geq 2$ by Lemma 4.6.2, an element which plays an important role as the lowest order SL *invariant component of the signature in dimension $d = n$, see [DR19, Section 3.3], and can be interpreted as the $d = n$ dimensional signed volume of the path underlying the signature [DR19, Definition 3.27 and Theorem 3.28]. In particular, we recover*

$$\mathrm{Inv}_2 = \mathsf{area}(\mathtt{1}, \mathtt{2}),$$
$$\mathrm{Inv}_3 = \mathsf{vol}(\mathtt{1}, \mathtt{2}, \mathtt{3}),$$

and in fact this can be generalized by defining the multilinear map

$$\mathsf{vol}^n : T(\mathbb{R}^d)^n \to T(\mathbb{R}^d)$$

such that $\mathsf{vol}^n(a_1, \ldots, a_n)$ is the image of Inv_n under the unique Zinbiel homomorphism (unique due to freeness of the halfshuffle algebra as a Zinbiel algebra) that maps $\mathtt{i} \mapsto a_i$ for $\mathtt{i} = \mathtt{1}, \ldots, \mathtt{n}$. Written out, this means

$$\mathsf{vol}^n(a_1, \ldots, a_n) = \sum_{\sigma \in S_n} \mathsf{sign}(\sigma)\, (((a_{\sigma(1)} \succ a_{\sigma(2)}) \succ a_{\sigma(3)}) \succ \cdots) \succ a_{\sigma(n)}.$$

Through the fact that $\mathrm{Inv}_n \in \mathscr{A}$ for $d = n$, it is immediate that we obtain the restriction

$$\mathsf{vol}^n : \mathscr{A}^n \to \mathscr{A}$$

for any $d \geq 2$.

As a direct consequence of the theorem by Dzhumadil'daev, Ismailov and Mashurov 4.6.2, we get the following property with respect to the deconcatenation coproduct, which is very important for expected signature computations in the context of martingaloids as we saw in Proposition 4.5.10.

Corollary 4.6.4. *$\mathscr{A} \oplus \mathbb{R}$ is a right coideal with respect to $\Delta_\bullet$, i.e. $\Delta_\bullet \mathscr{A} \in T(\mathbb{R}^d) \otimes (\mathscr{A} \oplus \mathbb{R})$*

Proof. For letters $\mathtt{i}$, we have $\Delta_\bullet \mathtt{i} = \mathbf{e} \otimes \mathtt{i} + \mathtt{i} \otimes \mathbf{e}$, and thus the statement is clear. For and (empty or non-empty) word w and letters $\mathtt{i}, \mathtt{j}$ we have

$$\Delta_\bullet\, w(\mathtt{ij} - \mathtt{ji}) = \sum_{u,v:\, uv = w} u \otimes v(\mathtt{ij} - \mathtt{ji}) + w\mathtt{i} \otimes \mathtt{j} - w\mathtt{j} \otimes \mathtt{i} + w(\mathtt{ij} - \mathtt{ji}) \otimes \mathbf{e} \in T(\mathbb{R}^d) \otimes (\mathscr{A} \oplus \mathbb{R}),$$

where we sum over all (empty or non-empty) words u, v such that $uv = w$. □

[8] Recall the definition of $\mathscr{A}$ from the introduction: the smallest linear space containing the letters $\mathtt{1}, \ldots \mathtt{d}$ and being closed under the area operation.

We note the following conjecture, which was shown to hold true in the case $d = 2$ in [DIM19, Section 5] as well as in [Rei19, Section 3.2, Theorem 31]. The case $d \geq 3$ is still open.

Conjecture 4.6.5. *$\mathscr{A}$ is linearly generated by strict left-bracketings of the* area *operation. In particular, a linear basis for $\mathscr{A}$ (without the single letters) is given by*

$$\mathsf{area}(\mathsf{area}(\mathsf{area}(\mathtt{i}_1, \mathtt{i}_2), \mathtt{i}_3), \ldots, \mathtt{i}_n), \qquad n \geq 2, \mathtt{i}_1, \ldots, \mathtt{i}_n \in \{\mathtt{1}, \ldots, \mathtt{d}\}, \mathtt{i}_1 < \mathtt{i}_2.$$

Example 4.6.6. *For example, with $d = 2$, the tensor $\mathtt{12}(\mathtt{12} - \mathtt{21})$, which is in $\mathscr{A}$, can be written as*

$$\mathtt{12}(\mathtt{12} - \mathtt{21}) = \tfrac{1}{6}[2\,\mathsf{area}(\mathsf{area}(\mathsf{area}(\mathtt{1}, \mathtt{2}), \mathtt{1}), \mathtt{2}) - \mathsf{area}(\mathsf{area}(\mathsf{area}(\mathtt{1}, \mathtt{2}), \mathtt{2}), \mathtt{1})].$$

It turns out that for a bilinear, antisymmetric operation, showing that all bracketings can be rewritten as linear combinations of left-bracketings reduces to showing that this is possible for a small subset of bracketings. We have not been able to show that this subset of bracketings can be rewritten, but want to record this general fact nonetheless. We formulate the statement imprecisely here, and leave the exact statement Proposition 4.8.2 and its proof to the appendix.

Proposition 4.6.7. *To be able to rewrite any bracketing to a linear combination of left-bracketings, it is enough to verify this for bracketings of the form*

$$\mathsf{area}(\mathsf{area}(\ldots, \mathsf{area}(\mathsf{area}(a_1, a_2), a_3), \ldots, a_{n-2}), \mathsf{area}(a_{n-1}, a_n)), \quad a_i \in \mathscr{A}.$$

While trying to find a proof for Conjecture 4.6.5 for $d \geq 3$, we investigated in detail the operator $\overleftarrow{\mathsf{area}}$ given by the following definition.

Definition 4.6.8. *If $w = \mathtt{l}_1 \cdots \mathtt{l}_n$ is a word, we define $\overleftarrow{\mathsf{area}}(w)$ to be the left-bracketing expression*

$$\mathsf{area}(\ldots\mathsf{area}(\mathsf{area}(\mathsf{area}(\mathtt{l}_1, \mathtt{l}_2), \mathtt{l}_3), \mathtt{l}_4), \ldots, \mathtt{l}_n).$$

This is expanded linearly to an operation on the tensor algebra with $\overleftarrow{\mathsf{area}}(e) = 0$ and $\overleftarrow{\mathsf{area}}(\mathtt{l}) = \mathtt{l}$ for any letter $\mathtt{l}$.

We came across some interesting properties.

First, we show that there is an expansion formula for $\overleftarrow{\mathsf{area}}(w)$ in terms of permutations of the letters in the word w. To this end, define the right action of a permutation $\sigma \in S_n$ on words of length n as

$$\sigma := \mathtt{l}_{\sigma(1)} \cdots \mathtt{l}_{\sigma(n)},$$

where $w = \mathtt{l}_1 \cdots \mathtt{l}_n$.

Proposition 4.6.9. *We have*

$$\overleftarrow{\mathsf{area}}(\mathtt{l}_1 \cdots \mathtt{l}_n) = \mathtt{l}_1 \cdots \mathtt{l}_n\ \theta_n,$$

where

$$\theta_n := \sum_{\sigma \in S_n} f_n(\sigma)\sigma$$

and $f_n : S_n \to \{-1, 1\}$ *is given as*

$$f_n(\sigma) = \prod_{i=1}^{n} g_i(\sigma)$$

with

$$g_i(\sigma) = \begin{cases} +1, & \text{if } \sigma^{-1}(j) < \sigma^{-1}(i) \text{ for all } j \in \mathbb{N} \text{ with } j < i, \\ -1, & \text{else.} \end{cases}$$

Proof. For $n = 1$, there is only the identity permutation and $f_1(\mathrm{id}) = g_1(\mathrm{id}) = 1$, thus the statement is obviously true. For $n = 2$, we have $S_2 = \{\mathrm{id}, (12)\}$, $f_2(\mathrm{id}) = -f_2((12)) = 1$ and

$$\mathsf{area}(\mathtt{l}_1\mathtt{l}_2) = \mathtt{l}_1\mathtt{l}_2 - \mathtt{l}_2\mathtt{l}_1 = f_2(\mathrm{id})\mathtt{l}_1\mathtt{l}_2 + f_2((12))\mathtt{l}_2\mathtt{l}_1.$$

Assume the statement holds for some $n \in \mathbb{N} \setminus \{1\}$. Then,

$$\begin{aligned}
&\overleftarrow{\mathsf{area}}(\mathtt{l}_1 \cdots \mathtt{l}_{n+1}) = \mathsf{area}(\overleftarrow{\mathsf{area}}(\mathtt{l}_1 \cdots \mathtt{l}_n), \mathtt{l}_{n+1}) \\
&= \sum_{\sigma \in S_n} f_n(\sigma)\, \mathtt{l}_{\sigma(1)} \cdots \mathtt{l}_{\sigma(n)}\mathtt{l}_{n+1} - \sum_{\sigma \in S_n} f_n(\sigma)\, \mathtt{l}_{n+1} \succ (\mathtt{l}_{\sigma(1)} \cdots \mathtt{l}_{\sigma(n)}) \\
&= \sum_{\substack{\tilde\sigma \in S_{n+1}: \\ g_{n+1}(\tilde\sigma)=1}} f_{n+1}(\tilde\sigma)\, \mathtt{l}_{\tilde\sigma(1)} \cdots \mathtt{l}_{\tilde\sigma(n+1)} - \sum_{\sigma \in S_n} f_n(\sigma) \left(\mathtt{l}_{n+1} \sqcup\!\sqcup (\mathtt{l}_{\sigma(1)} \cdots \mathtt{l}_{\sigma(n-1)})\right)\mathtt{l}_{\sigma(n)} \\
&= \sum_{\substack{\tilde\sigma \in S_{n+1}: \\ g_{n+1}(\tilde\sigma)=1}} f_{n+1}(\tilde\sigma)\, \mathtt{l}_{\tilde\sigma(1)} \cdots \mathtt{l}_{\tilde\sigma(n+1)} + \sum_{\substack{\tilde\sigma \in S_{n+1}: \\ g_{n+1}(\tilde\sigma)=-1}} f_{n+1}(\tilde\sigma)\, \mathtt{l}_{\tilde\sigma(1)} \cdots \mathtt{l}_{\tilde\sigma(n+1)} \\
&= \sum_{\tilde\sigma \in S_{n+1}} f_{n+1}(\tilde\sigma)\, \mathtt{l}_{\tilde\sigma(1)} \cdots \mathtt{l}_{\tilde\sigma(n)}.
\end{aligned}$$

□

We furthermore have the following surprising identity.

Proposition 4.6.10. *For all integers* $n > 2$, *we have*

$$\overleftarrow{\mathsf{area}}(\mathtt{1}\, l(\mathtt{2} \cdots \mathtt{n})) = \mathtt{1}\, \overleftarrow{\mathsf{area}}(l(\mathtt{2} \cdots \mathtt{n})), \tag{4.25}$$

where l *is the left Lie bracketing. This implies (by freeness of the half-shuffle algebra)*

$$\overleftarrow{\mathsf{area}}(v\, l(w)) = \overleftarrow{\mathsf{area}}(v)\, \overleftarrow{\mathsf{area}}(l(w))$$

for words v, w *such that* $|w| \geq 2$.

Remark 4.6.11. Note that due to the well-known fact that strict left Lie bracketings linearly generate the free Lie algebra, more generally formulated, it holds that

$$\overleftarrow{\mathsf{area}}(vx) = \overleftarrow{\mathsf{area}}(v)\, \overleftarrow{\mathsf{area}}(x)$$

for any $v \in T(\mathbb{R}^d)$ and any Lie polynomial x with $\langle x, \mathtt{i} \rangle = 0$ for all letters $\mathtt{i}$.

Remark 4.6.12. In particular, we have

$$\overleftarrow{\mathsf{area}}(\mathtt{l}_1 \cdots \mathtt{l}_n\mathtt{l}_{n+1}) - \overleftarrow{\mathsf{area}}(\mathtt{l}_1 \cdots \mathtt{l}_{n+1}\mathtt{l}_n) = 2\, \overleftarrow{\mathsf{area}}(\mathtt{l}_1 \cdots \mathtt{l}_{n-1})(\mathtt{l}_n\mathtt{l}_{n+1} - \mathtt{l}_{n+1}\mathtt{l}_n)$$

for any letters $\mathtt{l}_1, \ldots, \mathtt{l}_{n+1}$.

Proof. For the base case, we compute

$$\overleftarrow{\mathsf{area}}(\mathtt{123} - \mathtt{132}) = 2(\mathtt{123} - \mathtt{132}) = \mathtt{1}\overleftarrow{\mathsf{area}}(\mathtt{23} - \mathtt{32}).$$

Assume (4.25) holds for some integer $n > 2$ and let w be a word of length $n-1$. Then, for any letter $\mathtt{i}$,

$$\begin{aligned}
\overleftarrow{\mathsf{area}}(\mathtt{1}\, l(w\mathtt{i})) =& \overleftarrow{\mathsf{area}}(\mathtt{1}l(w)\mathtt{i} - \mathtt{1i}l(w)) = \overleftarrow{\mathsf{area}}(\mathtt{1}l(w)\mathtt{i}) - \overleftarrow{\mathsf{area}}(\mathtt{1i})\, \overleftarrow{\mathsf{area}}(l(w)) \\
=& \overleftarrow{\mathsf{area}}(\mathtt{1}l(w))\mathtt{i} - \mathtt{i} \succ \overleftarrow{\mathsf{area}}(\mathtt{1}l(w)) - \overleftarrow{\mathsf{area}}(\mathtt{1i})\, \overleftarrow{\mathsf{area}}(l(w)) \\
=& \mathtt{1}\, \overleftarrow{\mathsf{area}}(l(w))\mathtt{i} - \mathtt{i} \succ (\mathtt{1}\, \overleftarrow{\mathsf{area}}(l(w))) - \mathtt{1i}\, \overleftarrow{\mathsf{area}}(l(w)) + \mathtt{i1}\, \overleftarrow{\mathsf{area}}(l(w)) \\
=& \mathtt{1}\, \overleftarrow{\mathsf{area}}(l(w))\mathtt{i} - \mathtt{1}(\mathtt{i} \succ \overleftarrow{\mathsf{area}}(l(w))) - \mathtt{1i}\, \overleftarrow{\mathsf{area}}(l(w)) \\
=& \mathtt{1}\, \overleftarrow{\mathsf{area}}(l(w)\mathtt{i}) - \mathtt{1i}\, \overleftarrow{\mathsf{area}}(l(w)) \\
=& \mathtt{1}\, \overleftarrow{\mathsf{area}}(l(w)\mathtt{i}) - \mathtt{1}\, \overleftarrow{\mathsf{area}}(\mathtt{i}l(w)) \\
=& \mathtt{1}\, \overleftarrow{\mathsf{area}}(l(w\mathtt{i})),
\end{aligned}$$

where we used that $\mathtt{i} \succ (\mathtt{1}\, \overleftarrow{\mathsf{area}}(l(w))) = \mathtt{i1}\, \overleftarrow{\mathsf{area}}(l(w)) + \mathtt{1}(\mathtt{i} \succ \overleftarrow{\mathsf{area}}(l(w)))$ due to the combinatorial expansion formula for the half-shuffle. □

Interestingly, ρ admits a permutation expansion which is quite similar to that of $\overleftarrow{\mathsf{area}}$, in fact again via f_n, just that now only a subset T_n of all permutations S_n is involved.

Proposition 4.6.13. *We have*

$$\rho(\mathtt{l}_1 \cdots \mathtt{l}_n) = \mathtt{l}_1 \cdots \mathtt{l}_n\ \vartheta_n,$$

where

$$\vartheta_n := \sum_{\sigma \in T_n} f_n(\sigma)\sigma,$$

$f_n : S_n \to \{-1, 1\}$ is as in Lemma 4.6.9 and T_n is the set of all $\sigma \in S_n$ such that $\{\sigma(i), \dots, \sigma(n)\}$ is an interval of integers (a set of the form $[a, b] \cap \mathbb{N}$) for all $i \in \{1, \dots, n-1\}$.

Proof. For $n = 1$, there is only the identity permutation and $f_1(\mathrm{id}) = g_1(\mathrm{id}) = 1$, thus the statement is obviously true. For $n = 2$, we have $T_2 = \{\mathrm{id}, (12)\}$, $f_2(\mathrm{id}) = -f_2((12)) = 1$ and

$$\rho(\mathtt{l}_1\mathtt{l}_2) = \mathtt{l}_1\mathtt{l}_2 - \mathtt{l}_2\mathtt{l}_1 = f_2(\mathrm{id})\mathtt{l}_1\mathtt{l}_2 + f_2((12))\mathtt{l}_2\mathtt{l}_1.$$

Assume the statement holds for some $n \in \mathbb{N} \setminus \{1\}$. Then, using the recursive definition of ρ from Equation (4.5),

$$\begin{aligned}
\rho(\mathtt{l}_1 \cdots \mathtt{l}_{n+1}) &= \mathtt{l}_1 \rho(\mathtt{l}_2 \cdots \mathtt{l}_{n+1}) - \mathtt{l}_{n+1}\rho(\mathtt{l}_1 \cdots \mathtt{l}_n) \\
&= \sum_{\sigma \in T_n} f_n(\sigma)\, \mathtt{l}_1 l'_{\sigma(1)} \cdots l'_{\sigma(n)} - \sum_{\sigma \in T_n} f_n(\sigma)\, \mathtt{l}_{n+1}\mathtt{l}_{\sigma(1)} \cdots \mathtt{l}_{\sigma(n)} \\
&= \sum_{\substack{\tilde\sigma \in T_{n+1}: \\ \tilde\sigma(1) = 1}} f_{n+1}(\tilde\sigma)\, \mathtt{l}_{\tilde\sigma(1)} \cdots \mathtt{l}_{\tilde\sigma(n+1)} + \sum_{\substack{\tilde\sigma \in T_n: \\ \tilde\sigma(1) = n+1}} f_{n+1}(\tilde\sigma)\, \mathtt{l}_{\tilde\sigma(1)} \cdots \mathtt{l}_{\tilde\sigma(n+1)} \\
&= \sum_{\tilde\sigma \in T_{n+1}} f_{n+1}(\tilde\sigma)\, \mathtt{l}_{\tilde\sigma(1)} \cdots \mathtt{l}_{\tilde\sigma(n)},
\end{aligned}$$

where $l'_i = \mathtt{l}_{i+1}$. □

Via the recursive formula for ρ, we also get an alternative proof of the following.

Corollary 4.6.14. $\operatorname{Im}\rho \subset \mathscr{A}$

Proof. It suffices to show $\rho(w) \in \mathscr{A}$ for any word w. We have $\rho(e) = 0$, $\rho(\mathtt{i}) = \mathtt{i} \in \mathscr{A}$ and $\rho(\mathtt{ij}) = \mathtt{ij} - \mathtt{ji} \in \mathscr{A}$ for any letters $\mathtt{i}, \mathtt{j}$. Let $n \geq 2$ and assume that $\rho(w) \in \mathscr{A}$ for any word w with $|w| = n$. Then, for any word v with $|v| = n-1$ and any letters $\mathtt{i}, \mathtt{j}$, we have

$$\rho(\mathtt{i}v\mathtt{j}) = \mathtt{i}\rho(v\mathtt{j}) - \mathtt{j}\rho(\mathtt{i}v) \in \mathscr{A}$$

since $\rho(v\mathtt{j}), \rho(\mathtt{i}v) \in \mathscr{A}$ with $|\rho(v\mathtt{j})| = |\rho(\mathtt{i}v)| = n \geq 2$ due to the induction hypothesis, and by Lemma 4.6.2, the non-letter part of $\mathscr{A}$ is stable under concatenation of any element of the tensor algebra from the left. Thus, the induction hypothesis also holds for all words of length $n+1$. $\square$

4.7 Conclusion

We have linked the area operation in the tensor algebra to work in control theory and more abstract work on Tortkara algebras. We have shown that starting from letters and applying the area operation, one obtains enough elements to shuffle-generate the tensor algebra.

There are many open directions for research. We have not identified a minimal set of areas-of-areas which is just enough to shuffle-generate the tensor algebra – i.e. to shuffle-generate it exactly. The linear span of the areas-of-areas has been identified, but a basis for it in terms of areas-of-areas has not.

4.7.1 Open combinatorial problems

1. What is $\operatorname{span}\{\mathsf{area}(\mathtt{i}_1 ⧢ \cdots ⧢ \mathtt{i}_n, \mathtt{j}_1 ⧢ \cdots ⧢ \mathtt{j}_m),\ n, m \in \mathbb{N},\ \mathtt{i}_1, \ldots \mathtt{i}_n, \mathtt{j}_1, \ldots \mathtt{j}_m \text{ letters}\}$? Does it shuffle generate together with the letters, and if not what is the smallest subalgebra of the associative shuffle algebra containing it and the letters? This is the algebraic formulation of the question "what do we know about a path if we are only allowed to collect its increment and the values of the first area of any two dimensional polynomial image of the path"?

2. Give linear bases for $\mathscr{A} ⧢ \mathscr{A}$, $\mathscr{A} ⧢ \mathscr{A} ⧢ \mathscr{A}$, ... Does $\sum_{m=1}^{n} \mathscr{A}^{⧢ m}$ already arrive at $T(\mathbb{R}^d)$ for a finite n?

3. Give a minimal generating set for $T(\mathbb{R}^d)$ as a Tortkara algebra. Is it free?

4. In light of Proposition 4.6.10, look at $\langle x \in \mathfrak{g}(\mathbb{R}^d) : \langle x, \mathtt{i} \rangle = 0 \,\forall\, \mathtt{i}; \cdot \rangle$ and its image under $\overleftarrow{\mathsf{area}}$

5. What are the eigenspaces of $\overleftarrow{\mathsf{area}}$?

4.8 Appendix

Lemma 4.8.1. *Let $V = \bigoplus_{n\geq 1} V_n$ be a graded vector space, each V_n finite dimensional, and denote the grading $|.|_V$.*

Consider $\mathbb{R}[V]$, the symmetric algebra over V (see Section 4.4), with two different gradings, defined on monomials as follows

- $|x^m|_{\text{deg}} := m$ *(denote the corresponding projection onto degree* 1 *by* $\mathsf{proj}_1^{\text{deg}}$*).*
- $|x^m|_{\text{weight}} := m \cdot |x|_V$.

Let $Y \subset \mathbb{R}[V]$*, countable, be such that every* $y \in Y$ *is homogeneous with respect to* $|.|_{\text{weight}}$*. Then:*

$$Y \text{ generates } \mathbb{R}[V] \text{ (as a commutative algebra)}$$
$$\text{if and only if}$$
$$\text{span}_{\mathbb{R}}\, \mathsf{proj}_1^{\text{deg}} Y = V$$

If moreover $\mathsf{proj}_1^{\text{deg}} y$*,* $y \in Y$*, are linearly independent, then* Y *freely generates* $\mathbb{R}[V]$ *(as commutative algebra).*

Proof. We show the first statement.

$\Rightarrow$: Assume $v \in V \subset \mathbb{R}[V]$ is not in the span of $\mathsf{proj}_1^{\text{deg}} Y$. Then it is clearly not in the algebra generated by Y. Hence Y does not generate $\mathbb{R}[V]$. This proves the contrapositive.

$\Leftarrow$: Denote, local to this proof, by $\langle M \rangle$ the subalgebra generated by $M \subset \mathbb{R}[V]$.

Claim: $\langle V_1 \rangle \subset \langle Y \rangle$. Indeed, $v \in V_1$ can, by assumption be written as a linear combination of some

$$\mathsf{proj}_1^{\text{deg}} y_i,$$

where $y_i \in Y$. Since the y_i are homogeneous they must be of weight 1. Hence $\mathsf{proj}_1^{\text{deg}} y_i = y_i$, hence $V_1 \subset \langle Y \rangle$, hence $\langle V_1 \rangle \subset \langle Y \rangle$, which proves the claim.

Now let $\langle V_1 \oplus \dots \oplus V_n \rangle \subset \langle Y \rangle$. Claim: $V_{n+1} \subset \langle Y \rangle$. Indeed, $v \in V_{n+1}$ can be written as a linear combination of some

$$\mathsf{proj}_1^{\text{deg}} y_i,$$

where $y_i \in Y$, of weight $n+1$. Then

$$y_i = \mathsf{proj}_1^{\text{deg}} y_i + r_i,$$

with r_i monomials (of order 2 an higher) in terms from $V_1 \oplus \dots \oplus V_n$, i.e. $r_i \in \langle V_1 \oplus \dots \oplus V_n \rangle \subset \langle Y \rangle$. Hence $v \in \langle Y \rangle$.

Hence $\langle V_1 \oplus \dots \oplus V_n \oplus V_{n+1} \rangle \subset \langle Y \rangle$. Iterating, we see that $\mathbb{R}[V] = \langle V \rangle \subset \langle Y \rangle$, which proves the first claim. □

We finally give a precise statement and proof of Theorem 4.6.7

Let V be an $\mathbb{R}$-vector space and let

$$\mathfrak{B} : V \times V \to V,$$

be a bilinear map.[9] We encode bracketings as planar trees. Define the complete left-bracketed tree with n leaves as

$$\mathsf{LeftBracketTree}_1 := \bullet$$
$$\mathsf{LeftBracketTree}_n := \mathsf{LeftBracketTree}_{n-1} \to_\bullet \bullet, \qquad n \geq 2,$$

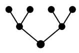
[9] This section would be most comfortably be formulated in the language of operads. But this would require more mathematical setup, which we want to avoid.

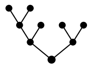
her $\to_\bullet$ denotes grafting to a new root.

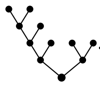
Define

$$\mathsf{SpecialTree}_n := \mathsf{LeftBracketTree}_{n-2} \to_\bullet \mathsf{LeftBracketTree}_2.$$

For example

$\mathsf{SpecialTree}_4 =$ 

$\mathsf{SpecialTree}_5 =$ 

$\mathsf{SpecialTree}_6 =$ .

For any tree τ with n leaves, and $a_1, \ldots, a_n \in V$ write

$$\tau(a_1, \ldots, a_n),$$

as the corresponding bracketing. We extend this definition to the case where (some of) the a_i are planar trees (with labeled leaves) themselves, by just replacing the respective leaf of τ with a_i. (This is consistent, when considering $a \in V$ as the tree with exactly on vertex, labeled a.)

On every new level $n+1$, it is enough to check that $\mathsf{SpecialTree}_{n+1}$ can be expressed in terms of left brackets:

Proposition 4.8.2. *Assume that $\mathfrak{B}$ is symmetric or anti-symmetric.*

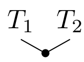
Assume, for some n, that all trees τ with $|\tau|_{\mathsf{leaves}} \le n$ can be expressed in terms of left brackets, i.e. for some $c(\tau, \sigma) \in \mathbb{R}$,

$$\tau(a_1, \ldots, a_n) = \sum_{\sigma \in S_n} c(\tau, \sigma) \mathsf{LeftBracketTree}_n(a_{\sigma(1)}, \ldots, a_{\sigma(n)}). \quad \forall a_1, \ldots, a_n \in V.$$

Assume that (every labeling of) $\mathsf{SpecialTree}_{n+1}$ can be expressed in terms of left brackets. Then:

(every labeling of) every tree σ with $|\sigma|_{\mathsf{leaves}} = n+1$ can be expressed in terms of left brackets.

Proof. Consider

$\tau =$ 

,

with

$$T_1 = \tau_1(a_1, \ldots, a_m)$$
$$T_2 = \tau_2(a_{m+1}, \ldots, a_{m+\ell})$$

with $|\tau|_{\mathsf{leaves}} = n + 1 = m + \ell$. By using symmetry/antisymmetry, we can assume $|\tau_1|_{\mathsf{leaves}} = m \ge |\tau_2|_{\mathsf{leaves}} = \ell$.

By assumption, we can write both τ_1 and τ_2 in terms of left-bracketings. It is hence enough to consider

$$\tau_1 = \mathsf{LeftBracketTree}_m$$
$$\tau_2 = \mathsf{LeftBracketTree}_\ell,$$

with $m+\ell=n+1$ and $m \geq \ell$.

Claim: we can reduce to $\ell=1$ and $\ell=2$.

Indeed, write

$$\begin{aligned}\tau_1 &= ((t_1,t_2),\ldots,t_{m-1})\\ \tau_2 &= (t_m,t_{m+1}),\end{aligned}$$

with

$$\begin{aligned}t_1 &= (a_1,a_2)\\ t_2 &= a_3\\ &\ldots\\ t_{m-1} &= a_m\\ t_m &= ((a_{m+1},a_{m+2}),\ldots,a_{m+\ell-1})\\ t_{m+1} &= a_{m+\ell}.\end{aligned}$$

If $\ell \geq 2$ we have $m+1 \leq n$. Hence by assumption

$$\tau = \sum \text{ left bracketings } (t_1,\ldots,t_{m+1}).$$

Now, consider the rightmost spot in each left-bracketing.

- If it is taken by a letter: $\rightsquigarrow \ell = 1$.
- If it is taken by t_1: $\rightsquigarrow \ell = 2$.
- If it is taken by t_m: $|t_m|_{\text{leaves}} = |\tau_2|_{\text{leaves}} - 1 = \ell - 1$. So we go from ℓ to $\ell-1$.

We can finish by induction.

□

Chapter 5

Conclusion

5.1 Scientific contribution and impact of our results

The result on renormalizing RDEs, Theorem 2.5.10, has been an important inspiration for the article [BCCH21], where a similar result for renormalizing SPDEs in the framework of regularity structures was presented as [BCCH21, Theorem 3.25]. In [BCE20], the duality Proposition 2.3.14 in the form of Remark 2.3.16 (there [BCE20, Lemma 4]) is used for $\mathcal{H}(A)$, where A is the algebra corresponding to a commutative semigroup, to introduce the *arborified Hoffman exponential* M_v^* [BCE20, Definition 1 and Theorem 2] with $v = (v_i)_i$ (in our formalism) given by

$$v_i = \sum_{\substack{\tau:\{\tau\}=i \\ |\tau|\geq 2}} \frac{1}{\tau!}\tau$$

where $\{\tau\}$ denotes the product of the labels of τ in A and $\tau!$ is the tree factorial. Our translation-renormalization map M_v can furthermore be embedded into the free transitive action on the space of branched rough paths developed in [TZ20], with the embedding specified in [TZ20, Equation (6.5)]. In [Rah21, Section 6], translation is introduced for planarly branched rough paths, and similar to our Theorem 2.5.10, compatibility with RDEs on homogenous spaces is shown in [Rah21, Proposition 6.2]. The special case of translating in time direction only with a non-time containing primitive element from Theorem 2.5.1 (ii) is described as an addition of a log-linear path in [BFPP22, Corollary 4.18]. Furthermore, [BFPP22, Section 3.4] discusses translation in the case of (smooth) quasi-geometric rough paths. Overall, the investigation of higher order translation of rough paths as renormalization and the direct, functional connection with regularity structures provides an introduction to the formalism of [BHZ19] for researchers fluent with rough paths theory.

Even though the theory behind Chapter 3 of the signature being a Zinbiel homomorphism has been laid out before, our presentation and application provides an explicit formulisation that is new in its accessibility for readers from both algebraic geometry and signature/rough path theory. The author of this thesis isn't aware of any prior work that lays out the compatibilty of the signature with the action of polynomial maps on paths space so clearly. Analogs of our results Theorem 3.1.2 and Corollary 3.3.5 for generalized iterated-sums signatures defined on the free commutative tridendriform algebra are given in [DET20a, Theorem 3.14 and Corollary 3.16].

The results in Chapter 4 establish areas of areas as an important, meaningful subset of signature components, whose importance to applications, including its characterization as those components which are piecewise linear for every piecewise linear path, is in particular explored in Section 4.5. Furthermore, in Sections 4.2 and 4.3 we reformulate the study of coordinates of the first kind by Eugénio M Rocha in a purely algebraic manner, unfortunately again with the conclusion that the resulting formula doesn't simplify their computation (Remark 4.3.13). Our characterization of (free) homogeneous shuffle generating sets is applicable far beyond the areas of areas case, and up to the best knowledge of the author of this thesis, the results haven't been written down that explicitly before, although certain experts might have already been aware of these facts (especially those having a deep understanding of the Milnor Moore theorem, as Joscha Diehl pointed out to the author of the thesis). Our main result is reproven in a more direct fashion in [Sal21, Theorem 4.4.2]. In [DET20b, Definition 5.7], a "discrete area map" **area** is introduced on the free commutative tridendriform algebra and it is shown in [DET20b, Theorem 5.9] that the Hoffman exponential constitutes a Tortkara isomorphism from $\mathscr{A}(S(V))$ to the smallest subspace of $T^{\geq 1}(S(V))$ containing $S(V)$ and being closed under **area**.

5.2 Outlook

Together, the individual projects collected in this thesis lay out in detail how Hopf, (pre-)Lie, Zinbiel, dendriform and Tortkara algebras have a direct application in the theory of rough paths and iterated integral signatures of paths. However, as the many open problems especially in the case of Tortkara algebras and their direct linking with questions about the signature, as well as the very modern classes of quasi-geometric and planarly branched rough paths show, this story is far from being told completely. With more and more meaningful examples of rough paths for commutative connected graded Hopf algebras emerging, the interest in (more or less) universal statements for these objects will grow further, and the author of this thesis is convinced that one of the most central questions in this general picture is how to algebraically characterize those commutative connected graded Hopf algebras which give rise to a rough path theory that includes a property of *uniqueness of the signature up to tree-like equivalence*. Furthermore, in prominent cases where such a property does not hold, one would ask what is the kernel of the signature map, and how one could minimally extend the given Hopf algebra structure to restore this property. Ultimately, the goal is a full picture of how the analysis works together with the graded Hopf algebra structure, as well as the 'finer structure' of certain Hopf algebras as we observe it with the Zinbiel structure in the case of weakly geometric rough paths, the commutative tridendriform structure in the case of quasi-geometric rough paths (both concerning the commutative product) and the pre-Lie structure in the case of branched rough paths as well as the post-Lie structure in the case of planarly branched rough paths (both concerning the Lie algebra underlying the Lie group of characters the rough paths lives in).

Given how our translation operators were defined in Sections 2.2.2 and 2.3.2.3, the following definition seems justified to obtain a generalization. Independently from this thesis, a very similar construction has been described in [Rah21, Definiton 8 and Theorem 5.1].

Definition 5.2.1. *We call a graded Lie admissible algebra $L = \bigoplus_{n=1}^{\infty} L_i$ with graded Lie admissible product $\triangleright$ and a (empty or non-empty) set of further graded multilinear maps (s_n^j), $s_n^j : L^n \to L$, where n may be any nonnegative integer, a* fully renormalizable graded Lie-admissible system *over W,*
if $L_1 = W$ generates $(L, \triangleright, (s_n^j))$, i.e. the smallest subspace of L containing $L_1 = W$ and being closed under $\triangleright$ and all (s_n^j) is L itself, and $(L, \triangleright, (s_n^j))$ is 'internally free' over W, i.e. for any linear map $v : W \to L$, there exists a homomorphism T_v of $\triangleright$ and all s_n^j, i.e. $(T_v x) \triangleright (T_v y) = T_v(x \triangleright y)$

for all $x, y \in L$ *and* $s_n^j \circ T_v^{\times n} = T_v \circ s_n^j$ *for all* s_n^j*, such that*

$$T_v w = w + v(w) \quad \forall\, w \in W.$$

Lie admissibility of $\triangleright$ means nothing more than that $(L, [\cdot,\cdot]_\triangleright)$ with $[x, y]_\triangleright = x \triangleright y - y \triangleright x$ is a Lie algebra. Thus, we can form the universal enveloping algebra $\mathcal{U}(L)$ to obtain a connected graded cocommutative Hopf algebra whose space of homogeneity one is W, together with a monoid of Hopf algebra endomorphisms $(T_v)_v$ indexed over all linear maps $v : W \to L$, L being the space of primitive elements of $\mathcal{U}(L)$, such that $T_v w = w + v(w)$. Finally, we can look at rough paths over the graded dual of $\mathcal{U}(L)$, which is then a connected graded commutative Hopf algebra, and study the effect of the translations-renormalizations $(T_v)_v$ on those rough paths.

In the well-established case of weakly geometric rough paths and the iterated-integral signature with Ree's shuffle identity, half a century after the initiation through Chen's series of papers, we as a community of researchers investigating signature-rough-paths theory are closing in on a more and more complete understanding of individual results as part of a big picture of an algebraic geometry of paths and time series.

Bibliography

[AFS19] Carlos Améndola, Peter Friz, and Bernd Sturmfels. Varieties of Signature Tensors. *Forum of Mathematics, Sigma*, 7:e10, 2019. doi:10.1017/fms.2019.3. 5, 65, 80

[AGS89] A. A. Agrachev, R. V. Gamkrelidze, and A. V. Sarychev. Local invariants of smooth control systems. *Acta Applicandae Mathematica*, 14(3):191–237, March 1989. doi:10.1007/BF01307214. 99

[Apo74] Tom M. Apostol. *Mathematical Analysis*. Addison-Wesley Publishing Company, Reading, Massachusetts, second edition, 1974. 3

[Bal13] Jennifer S. Balakrishnan. Iterated Coleman integration for hyperelliptic curves. In *ANTS X—Proceedings of the Tenth Algorithmic Number Theory Symposium*, volume 1 of *The Open Book Series*, pages 41–61. Mathematical Science Publishers, Berkeley, CA, 2013. doi:10.2140/obs.2013.1.41. 65

[BBM17] Fatemeh Bagherzadeh, Murray Bremner, and Sara Madariaga. Jordan trialgebras and post-Jordan algebras. *Journal of Algebra*, 486:360–395, September 2017. doi:10.1016/j.jalgebra.2017.04.022. 10

[BC19] Horatio Boedihardjo and Ilya Chevyrev. An isomorphism between branched and geometric rough paths. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*, 55(2):1131–1148, May 2019. doi:10.1214/18-AIHP912. 6, 30

[BCCH21] Yvain Bruned, Ajay Chandra, Ilya Chevyrev, and Martin Hairer. Renormalising SPDEs in regularity structures. *Journal of the European Mathematical Society*, 23(3):869–947, November 2021. doi:10.4171/JEMS/1025. 19, 137

[BCE20] Yvain Bruned, Charles Curry, and Kurusch Ebrahimi-Fard. Quasi-shuffle algebras and renormalisation of rough differential equations. *Bulletin of the London Mathematical Society*, 52(1):43–63, February 2020. doi:10.1112/blms.12305. 5, 45, 137

[BCF18] Yvain Bruned, Ilya Chevyrev, and Peter K. Friz. Examples of Renormalized SDEs. In Andreas Eberle, Martin Grothaus, Walter Hoh, Moritz Kassmann, Wilhelm Stannat, and Gerald Trutnau, editors, *Stochastic Partial Differential Equations and Related Fields*, volume 229 of *Springer Proceedings in Mathematics & Statistics*, pages 303–317, Cham, 2018. Springer International Publishing. doi:10.1007/978-3-319-74929-7_19. 38, 39, 40

[BCFP19] Y. Bruned, I. Chevyrev, P. K. Friz, and R. Preiß. A rough path perspective on renormalisation. *Journal of Functional Analysis*, 277(11):108283, December 2019. `doi:10.1016/j.jfa.2019.108283`. xiii, 15, 18, 45

[BD15] I. Bailleul and J. Diehl. The Inverse Problem for Rough Controlled Differential Equations. *SIAM Journal on Control and Optimization*, 53(5):2762–2780, 2015. `doi:10.1137/140995982`. 87

[Bel20] Carlo Bellingeri. Quasi-geometric rough paths and rough change of variable formula, September 2020. `arXiv:2009.00903v1 [math.PR]`. 6

[Ben89] Gérard Ben Arous. Flots et séries de Taylor stochastiques. *Probability Theory and Related Fields*, 81(1):29–77, February 1989. `doi:10.1007/BF00343737`. 22, 36

[BFG$^+$20] Christian Bayer, Peter K. Friz, Paul Gassiat, Jorg Martin, and Benjamin Stemper. A regularity structure for rough volatility. *Mathematical Finance*, 30(3):782–832, 2020. `doi:10.1111/mafi.12233`. 41

[BFPP22] Carlo Bellingeri, Peter K. Friz, Sylvie Paycha, and Rosa Preiß. Smooth rough paths, their geometry and algebraic renormalization. *Vietnam Journal of Mathematics*, 50(3):719–761, July 2022. `10.1007/s10013-022-00570-7`. 137

[BGLY16] Horatio Boedihardjo, Xi Geng, Terry Lyons, and Danyu Yang. The signature of a rough path: Uniqueness. *Advances in Mathematics*, 293:720–737, April 2016. `doi:10.1016/j.aim.2016.02.011`. 4, 5, 6, 82

[BHZ19] Y. Bruned, M. Hairer, and L. Zambotti. Algebraic renormalisation of regularity structures. *Inventiones mathematicae*, 215(3):1039–1156, March 2019. `doi:10.1007/s00222-018-0841-x`. xxi, 8, 12, 17, 18, 19, 22, 34, 49, 50, 51, 52, 53, 54, 56, 61, 137

[BL11] Nantel Bergeron and Jean-Louis Loday. The symmetric operation in a free pre-Lie algebra is magmatic. *Proceedings of the American Mathematical Society*, 139(5):1585–1597, May 2011. `http://www.jstor.org/stable/41102954`. 92

[Blo15] A. M. Bloch. *Nonholonomic Mechanics and Control*, volume 24 of *Interdisciplinary Applied Mathematics*. Springer-Verlag New York, second edition, 2015. `doi:10.1007/978-1-4939-3017-3`. 87

[Boe18] Horatio Boedihardjo. Decay rate of iterated integrals of branched rough paths. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 35(4):945–969, July 2018. `doi:10.1016/j.anihpc.2017.09.002`. 42

[Bre18] Murray Bremner. On tortkara triple systems. *Communications in Algebra*, 46(6):2396–2404, 2018. `doi:10.1080/00927872.2017.1384003`. 126

[Car07] Pierre Cartier. A primer of Hopf algebras. In Pierre Cartier, Pierre Moussa, Bernard Julia, and Pierre Vanhove, editors, *Frontiers in Number Theory, Physics, and Geometry II: On Conformal Field Theories, Discrete Groups and Renormalization*, pages 537–615. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. `doi:10.1007/978-3-540-30308-4_12`. 116, 117

[Cay57] A. Cayley Esq. XXVIII. On the theory of the analytical forms called trees. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 13(85):172–176, 1857. `doi:10.1080/14786445708642275`. 12

[CEM11] Damien Calaque, Kurusch Ebrahimi-Fard, and Dominique Manchon. Two interacting Hopf algebras of trees: a Hopf-algebraic approach to composition and substitution of B-series. *Advances in Applied Mathematics*, 47(2):282–308, August 2011. doi:10.1016/j.aam.2009.08.003. 30, 34, 45

[CEMM20] C. Curry, K. Ebrahimi-Fard, D. Manchon, and H. Z. Munthe-Kaas. Planarly branched rough paths and rough differential equations on homogeneous spaces. *Journal of Differential Equations*, 269(11):9740–9782, November 2020. doi:10.1016/j.jde.2020.06.058. 6

[CFO11] Michael Caruana, Peter K. Friz, and Harald Oberhauser. A (rough) pathwise approach to a class of non-linear stochastic partial differential equations. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 28(1):27–46, January-February 2011. doi:10.1016/j.anihpc.2010.11.002. 17

[CGM20] Laura Colmenarejo, Francesco Galuppi, and Mateusz Michałek. Toric geometry of path signature varieties. *Advances in Applied Mathematics*, 121:102102, August 2020. doi:10.1016/j.aam.2020.102102. 5

[CH18] Ajay Chandra and Martin Hairer. An analytic BPHZ theorem for regularity structures, January 2018. Version 5, arXiv:1612.08138v5 [math.PR]. 19

[Cha10] Frédéric Chapoton. Free pre-Lie algebras are free as Lie algebras. *Canadian Mathematical Bulletin*, 53(3):425–437, September 2010. doi:10.4153/CMB-2010-063-2. 30

[Che54] Kuo-Tsai Chen. Iterated Integrals and Exponential Homomorphisms. *Proceedings of the London Mathematical Society*, s3-4(1):502–512, 1954. doi:10.1112/plms/s3-4.1.502. 1, 2, 3, 69, 85

[Che57] Kuo-Tsai Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Annals of Mathematics Second Series*, 65(1):163–178, January 1957. doi:10.2307/1969671. 3, 69

[Che58] Kuo-Tsai Chen. Integration of Paths – A Faithful Representation of Paths by Noncommutative Formal Power Series. *Transactions of the American Mathematical Society*, 89(2):395–407, 1958. doi:10.2307/1993193. 4

[Che77] Kuo-Tsai Chen. Iterated path integrals. *Bulletin of the American Mathematical Society*, 83(5):831–879, September 1977. doi:10.1090/S0002-9904-1977-14320-6. 5

[Che18] Ilya Chevyrev. Random walks and Lévy processes as rough paths. *Probability Theory and Related Fields*, 170(3):891–932, April 2018. doi:10.1007/s00440-017-0781-1. 19, 36, 37

[CK98] Alain Connes and Dirk Kreimer. Hopf Algebras, Renormalization and Noncommutative Geometry. *Communications in Mathematical Physics*, 199(1):203–242, 1998. doi:10.1007/s002200050499. 11, 20

[CK16] Ilya Chevyrev and Andrey Kormilitzin. A Primer on the Signature Method in Machine Learning, March 2016. arXiv:1603.03788v1 [stat.ML]. 5, 86

[CL01] Frédéric Chapoton and Muriel Livernet. Pre-Lie algebras and the rooted trees operad. *International Mathematics Research Notices*, 2001(8):395–408, 2001. doi:10.1155/S1073792801000198. 10, 11, 31, 45

[CO22] Ilya Chevyrev and Harald Oberhauser. Signature moments to characterize laws of stochastic processes. *Journal of Machine Learning Research*, 23(176):1–42, June 2022. http://jmlr.org/papers/v23/20-1466.html. 5

[CP20] Laura Colmenarejo and Rosa Preiß. Signatures of paths transformed by polynomial maps. *Beiträge zur Algebra und Geometrie/Contributions to Algebra and Geometry*, 61(4):695–717, December 2020. doi:10.1007/s13366-020-00493-9. xiii, xxi, 65

[CQ02] Laure Coutin and Zhongmin Qian. Stochastic analysis, rough path analysis and fractional Brownian motions. *Probability Theory and Related Fields*, 122(1):108–140, January 2002. doi:10.1007/s004400100158. 19

[CW17] Thomas Cass and Martin P. Weidner. Tree algebras over topological vector spaces in rough path theory, October 2017. Version 2, arXiv:1604.07352v2 [math.PR]. 45

[DET20a] Joscha Diehl, Kurusch Ebrahimi-Fard, and Nikolas Tapia. Generalized iterated-sums signatures, December 2020. arXiv:2012.04597v1 [math.RA]. 10, 137

[DET20b] Joscha Diehl, Kurusch Ebrahimi-Fard, and Nikolas Tapia. Time-Warping Invariants of Multidimensional Time Series. *Acta Applicandae Mathematicae*, 170(1):265–290, December 2020. doi:10.1007/s10440-020-00333-x. 138

[Die13] Joscha Diehl. Rotation invariants of two dimensional curves based on iterated integrals, May 2013. arXiv:1305.6883v1 [cs.CV]. 3, 5

[DIM19] A. S. Dzhumadil'daev, N. A. Ismailov, and F. A. Mashurov. On the speciality of Tortkara algebras. *Journal of Algebra*, 540:1–19, December 2019. doi:10.1016/j.jalgebra.2019.08.010. xvi, 9, 14, 93, 94, 126, 128, 129

[DL02] Askar Dzhumadil'daev and Clas Löfwall. Trees, free right-symmetric algebras, free Novikov algebras and identities. *Homology, Homotopy and Applications*, 4(2):165–190, 2002. doi:10.4310/HHA.2002.v4.n2.a8. 31

[DLPR21] Joscha Diehl, Terry Lyons, Rosa Preiß, and Jeremy Reizenstein. Areas of areas generate the shuffle algebra, July 2021. Version 2, arXiv:2002.02338v2 [math.RA]. xiii, 80, 85

[DPRT22] Joscha Diehl, Rosa Preiß, Michael Ruddy, and Nikolas Tapia. The Moving-Frame Method for the Iterated-Integrals Signature: Orthogonal Invariants. *Foundations of Computational Mathematics*, June 2022. doi:10.1007/s10208-022-09569-5. 3, 5

[DR19] Joscha Diehl and Jeremy Reizenstein. Invariants of Multidimensional Time Series Based on Their Iterated-Integral Signature. *Acta Applicandae Mathematicae*, 164(1):83–122, December 2019. doi:10.1007/s10440-018-00227-z. 3, 5, 13, 65, 80, 86, 106, 121, 122, 126, 128

[Duc91] Gérard Duchamp. Orthogonal projection onto the free Lie algebra. *Theoretical Computer Science*, 79(1):227–239, February 1991. doi:10.1016/0304-3975(91)90153-S. 119

[Dzh07] A. S. Dzhumadil'daev. Zinbiel algebras under q-commutators. *Journal of Mathematical Sciences*, 144(2):3909–3925, July 2007. doi:10.1007/s10958-007-0244-9. 9, 10, 13, 125, 126

[EGP07] Kurusch Ebrahimi-Fard, José M. Gracia-Bondía, and Frédéric Patras. A Lie theoretic approach to renormalization. *Communications in mathematical physics*, 276(2):519–549, December 2007. doi:10.1007/s00220-007-0346-8. 96

[EM53] Samual Eilenberg and Saunders Mac Lane. On the Groups $H(\Pi, n)$, I. *Annals of Mathematics Second Series*, 58(1):55–106, July 1953. doi:10.2307/1969820. 75, 76

[EM09] Kurusch Ebrahimi-Fard and Dominique Manchon. Dendriform equations. *Journal of Algebra*, 322(11):4053–4079, December 2009. doi:10.1016/j.jalgebra.2009.06.002. 92, 97

[EP15] Kurusch Ebrahimi-Fard and Frédéric Patras. Cumulants, free cumulants and half-shuffles. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2176):20140843, April 2015. doi:10.1098/rspa.2014.0843. 10

[EP21] Kurusch Ebrahimi-Fard and Frédéric Patras. Quasi-shuffle Algebras in Non-commutative Stochastic Calculus. In Stefania Ugolini, Marco Fuhrman, Elisa Mastrogiacomo, Paola Morando, and Barbara Rüdiger, editors, *Geometry and Invariance in Stochastic Dynamics*, volume 378 of *Springer Proceedings in Mathematics & Statistics*, pages 89–112, Cham, 2021. Springer International Publishing. doi:10.1007/978-3-030-87432-2_6. 10

[FGL15] Peter Friz, Paul Gassiat, and Terry Lyons. Physical Brownian motion in a magnetic field as a rough path. *Transactions of the American Mathematical Society*, 367(11):7939–7955, November 2015. doi:10.1090/S0002-9947-2015-06272-2. 18, 38, 39

[FH14] Peter K. Friz and Martin Hairer. *A Course on Rough Paths: With an Introduction to Regularity Structures*. Universitext. Springer, Cham, 2014. doi:10.1007/978-3-319-08332-2. 6, 16, 17, 18, 65

[FHL16] Guy Flint, Ben Hambly, and Terry Lyons. Discretely sampled signals and the rough Hoff process. *Stochastic Processes and their Applications*, 126(9):2593–2614, September 2016. doi:10.1016/j.spa.2016.02.011. 40

[FO09] Peter Friz and Harald Oberhauser. Rough path limits of the Wong-Zakai type with a modified drift term. *Journal of Functional Analysis*, 256(10):3236–3256, May 2009. doi:10.1016/j.jfa.2009.02.010. 17, 18, 46

[Foi02] L. Foissy. Finite-dimensional comodules over the Hopf algebra of rooted trees. *Journal of Algebra*, 255(1):89–120, September 2002. doi:10.1016/S0021-8693(02)00110-2. 30

[Foi13] Loïc Foissy. An introduction to Hopf algebras of trees, 2013. http://www2.mathematik.hu-berlin.de/~kreimer/wp-content/uploads/Foissy.pdf. 11

[FP13] Loïc Foissy and Frédéric Patras. Natural endomorphisms of shuffle algebras. *International Journal of Algebra and Computation*, 23(4):989–1009, 2013. doi:10.1142/S0218196713400183. 75, 76, 91

[FP20] Loïc Foissy and Frédéric Patras. Lie Theory for Quasi-Shuffle Bialgebras. In José Ignacio Burgos Gil, Kurusch Ebrahimi-Fard, and Herbert Gangl, editors, *Periods in Quantum Field Theory and Arithmetic. ICMAT-MZV 2014*, volume 314 of *Springer Proceedings in Mathematics & Statistics*, pages 483–540. Springer Nature Switzerland, 2020. doi:10.1007/978-3-030-37031-2_19. 10

[FS82] Gerald B. Folland and Elias M. Stein. *Hardy spaces on homogeneous groups*, volume 28 of *Mathematical Notes*. Princeton University Press, Princeton, N.J. and University of Tokyo Press, Tokyo, 1982. doi:10.2307/j.ctv17db3q0. 36

[FS07] Robert L. Foote and Ed Sandifer. Area Without Integration: Make Your Own Planimeter. In Amy Shell-Gellasch, editor, *Hands on History: A Resource for Teaching Mathematics*, pages 71–88. Mathematical Association of America, 2007. doi:10.5948/UPO9780883859766.010. 87

[FS17] Peter Friz and Atul Shekhar. General rough integration, Lévy Rough paths and a Lévy–Kintchine type formula. *The Annals of Probability*, 45(4):2707–2765, July 2017. doi:10.1214/16-AOP1123. 19, 36

[FV10] Peter K. Friz and Nicolas B. Victoir. *Multidimensional Stochastic Processes as Rough paths: Theory and Applications*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. doi:10.1017/CBO9780511845079. 19, 23, 40, 42, 43, 45, 65, 122

[Gal19] Francesco Galuppi. The rough Veronese variety. *Linear Algebra and Its Applications*, 583:282–299, December 2019. doi:10.1016/j.laa.2019.08.029. 5, 65

[Gar90] Adriano M. Garsia. Combinatorics of the free Lie algebra and the symmetric group. In Paul H. Rabinowitz and Eduard Zehnder, editors, *Analysis, et cetera: Research Papers Published in Honor of Jürgen Moser's 60th Birthday*, pages 309–382. Elsevier, 1990. doi:10.1016/B978-0-12-574249-8.50021-3. 94

[GIP15] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics, Pi*, 3:e6, 75 pages, 2015. doi:10.1017/fmp.2015.2. 17

[GK08] Eric Gehrig and Matthias Kawski. A hopf-algebraic formula for compositions of noncommuting flows. In *2008 47th IEEE Conference on Decision and Control*, pages 1569–1574, 2008. doi:10.1109/CDC.2008.4738914. 67, 79, 107

[GL89] Robert Grossman and Richard G. Larson. Hopf-Algebraic Structure of Families of Trees. *Journal of Algebra*, 126(1):184–210, October 1989. doi:10.1016/0021-8693(89)90328-1. 11

[Gub04] M. Gubinelli. Controlling rough paths. *Journal of Functional Analysis*, 216(1):86–140, November 2004. doi:10.1016/j.jfa.2004.01.002. 16

[Gub10] Massimiliano Gubinelli. Ramification of rough paths. *Journal of Differential Equations*, 248(4):693–721, February 2010. doi:10.1016/j.jde.2009.11.015. 6, 16, 18, 19, 22, 42

[GVF01] José M. Gracia-Bondía, Joseph C. Várilly, and Héctor Figueroa. *Elements of Noncommutative Geometry*. Birkhäuser Advanced Texts: Basler Lehrbücher. Birkhäuser Boston Inc., Boston, MA, 2001. doi:10.1007/978-1-4612-0005-5. 28

[Hai02] Richard Hain. Iterated Integrals and Algebraic Cycles: Examples and Prospects. In Shiing-Shen Chern, Lei Fu, and Richard Hain, editors, *Contemporary trends in algebraic geometry and algebraic topology*, volume 5 of *Nankai Tracts in Mathematics*, pages 55–118. World Scientific Publishing Co., Singapore, August 2002. `doi:10.1142/9789812777416_0004`. 5, 65

[Hai13] Martin Hairer. Solving the KPZ equation. *Annals of Mathematics, Second Series*, 178(2):559–664, September 2013. `doi:10.4007/annals.2013.178.2.4`. 18

[Hai14] M. Hairer. A theory of regularity structures. *Inventiones mathematicae*, 198(2):269–504, November 2014. `doi:10.1007/s00222-014-0505-4`. 7, 17, 19, 25, 50, 51, 53, 54, 58

[Hai16] M. Hairer. The motion of a random string, June 2016. Version 2, `arXiv:1605.02192v2 [math.PR]`. 17, 18, 49, 51, 52

[Hal17] Paul R. Halmos. *Finite-Dimensional Vector Spaces.* Dover Publications, Inc., Mineola, New York, second edition, 2017. 118

[HK15] Martin Hairer and David Kelly. Geometric versus non-geometric rough paths. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 51(1):207–251, February 2015. `doi:10.1214/13-AIHP564`. 11, 15, 16, 18, 19, 20, 22, 28, 29, 36, 42, 43, 45, 47

[HL08] Ben Hambly and Terry Lyons. Some notes on trees and paths, September 2008. `arXiv:0809.1365v1 [math.CA]`. 4

[HL10] Ben Hambly and Terry Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Annals of Mathematics Second Series*, 171(1):109–167, January 2010. `doi:10.4007/annals.2010.171.109`. 4, 82

[Hof00] Michael E. Hoffman. Quasi-Shuffle Products. *Journal of Algebraic Combinatorics*, 11(1):49–68, January 2000. `doi:10.1023/A:1008791603281`. 6

[HP11] Patrick Hilger and Norbert Poncin. Lectures on algebraic operads. University of Luxemburg, 2011. `https://orbilu.uni.lu/bitstream/10993/14381/1/LecturesAlgebraicOperads.pdf`. 9

[IW89] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library.* North-Holland Publishing Co., Amsterdam and Kodansha, Ltd., Tokyo, second edition, 1989. 16, 123

[Kap09] Mikhail Kapranov. Noncommutative Geometry and Path Integrals. In Yuri Tschinkel and Yuri Zarhin, editors, *Algebra, Arithmetic, and Geometry: In Honor of Yu.Ĩ. Manin. Vol. II*, volume 270 of *Progress in Mathematics*, pages 49–87. Birkhäuser Boston, 2009. `doi:10.1007/978-0-8176-4747-6_3`. 5, 65

[Kaw09] Matthias Kawski. Chronological Calculus in Systems and Control Theory. In Robert A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*, pages 1027–1041. Springer, New York, NY, 2009. `doi:10.1007/978-0-387-30440-3_68`. 105

[KH13] David Kelly and Martin Hairer. Itô's formula via rough paths. Presentation slides, Stochastic Analysis Seminar 2013, Oxford, April 2013. 6

[KM21] Iryna Kashuba and Olivier Mathieu. On the free Jordan algebras. *Advances in Mathematics*, 383:107690, June 2021. `doi:10.1016/j.aim.2021.107690`. 9

[KO19] Franz J. Király and Harald Oberhauser. Kernels for sequentially ordered data. *Journal of Machine Learning Research*, 20(31):1–45, January 2019. `http://jmlr.org/papers/v20/16-314.html`. 5

[Kre98] Dirk Kreimer. On the Hopf algebra structure of perturbative quantum field theories. *Advances in Theoretical and Mathematical Physics*, 2(2):303–334, January 1998. `doi:10.4310/ATMP.1998.v2.n2.a4`. 20

[LCL07] Terry J. Lyons, Michael Caruana, and Thierry Lévy. *Differential Equations Driven by Rough Paths*, volume 1908 of *Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg, 2007. `doi:10.1007/978-3-540-71285-5`. 85

[Lia04] Ming Liao. *Lévy processes in Lie groups*, volume 162 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2004. `doi:10.1017/CBO9780511546624`. 36

[LM13] Alexander Lundervold and Hans Munthe-Kaas. Backward Error Analysis and the Substitution Law for Lie Group Integrators. *Foundations of Computational Mathematics*, 13(2):161–186, April 2013. `doi:10.1007/s10208-012-9130-z`. 94

[Lod95] Jean-Louis Loday. CUP-Product for Leibniz Cohomology and Dual Leibniz Algebras. *Mathematica Scandinavica*, 77(2):189–196, December 1995. `doi:10.7146/math.scand.a-12560`. 76, 77

[Lod01] Jean-Louis Loday. Dialgebras. In *Dialgebras and related operads*, pages 7–66. Springer, 2001. 10, 97

[Lod07] Jean-Louis Loday. On the algebra of quasi-shuffles. *manuscripta mathematica*, 123(1):79–93, May 2007. `doi:10.1007/s00229-007-0086-2`. 10

[LP93] Jean-Louis Loday and Teimuraz Pirashvili. Universal enveloping algebras of Leibniz algebras and (co)homology. *Mathematische Annalen*, 296(1):139–158, December 1993. `doi:10.1007/BF01445099`. 8

[LV12] Jean-Louis Loday and Bruno Vallette. *Algebraic Operads*, volume 346 of *Grundlehren der mathematischen Wissenschaften*. Springer Verlag Berlin-Heidelberg, 2012. `doi:10.1007/978-3-642-30362-3`. 10

[LY06] Terry J. Lyons and Phillip S. C. Yam. On Gauss–Green theorem and boundaries of a class of Hölder domains. *Journal de Mathématiques Pures et Appliquées*, 85(1):38–53, January 2006. `doi:10.1016/j.matpur.2005.10.005`. 86

[Lyo94] Terry Lyons. Differential equations driven by rough signals (I): An extension of an inequality of L. C. Young. *Mathematical Research Letters*, 1(4):451–464, January 1994. `doi:10.4310/MRL.1994.v1.n4.a5`. 15

[Lyo95] Terry J. Lyons. The interpretation and solution of ordinary differential equations driven by rough signals. In Michael C. Cranston and Mark A. Pinsky, editors, *Stochastic Analysis*, volume 57 of *Proceedings of Symposia in Pure Mathematics*, pages 115–128, 1995. `doi:10.1090/pspum/057/1335466`. 15

[Lyo98] Terry J. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998. https://eudml.org/doc/39555. 6, 15, 16, 42

[Man11] Dominique Manchon. A short survey on pre-Lie algebras. In Alan L. Carey, editor, *Noncommutative geometry and physics: renormalisation, motives, index theory*, ESI Lectures in Mathematics and Physics, pages 89–102. European Mathematical Society Publishing House, Zürich, 2011. doi:10.4171/008-1/3. 12, 18, 30

[McC78] Kevin McCrimmon. Jordan algebras and their applications. *Bulletin of the American Mathematical Society*, 84(4):612–627, July 1978. doi:10.1090/S0002-9904-1978-14503-0. 9

[MNT13] Frédéric Menous, Jean-Christophe Novelli, and Jean-Yves Thibon. Mould calculus, polyhedral cones, and characters of combinatorial Hopf algebras. *Advances in Applied Mathematics*, 51(2):177–227, July 2013. doi:10.1016/j.aam.2013.02.003. 107

[MP13] Frédéric Menous and Frédéric Patras. Logarithmic derivatives and generalized Dynkin operators. *Journal of Algebraic Combinatorics*, 38(4):901–913, December 2013. doi:10.1007/s10801-013-0431-3. 94

[MW08] H. Z. Munthe-Kaas and W. M. Wright. On the Hopf Algebraic Structure of Lie Group Integrators. *Foundations of Computational Mathematics*, 8(2):227–257, April 2008. doi:10.1007/s10208-006-0222-5. 6

[PR02a] Frédéric Patras and Christophe Reutenauer. Lie Representations and an Algebra Containing Solomon's. *Journal of Algebraic Combinatorics*, 16(3):301–314, November 2002. doi:10.1023/A:1021856522624. 119

[PR02b] Frédéric Patras and Christophe Reutenauer. On Dynkin and Klyachko idempotents in graded bialgebras. *Advances in Applied Mathematics*, 28(3-4):560–579, April 2002. doi:10.1006/aama.2001.0795. 94

[Pre16] Rosa Preiß. From Hopf algebras to rough paths and regularity structures. Master's thesis, Technische Universität Berlin, July 2016. http://page.math.tu-berlin.de/~preiss/files/masters.pdf. xix, 26, 56, 58, 59

[Pre21] Rosa Preiß. Areas of areas generate the shuffle algebra. Extended abstract, based on joint work with Joscha Diehl, Terry Lyons and Jeremy Reizenstein. In *Report No. 40/2020. New Directions in Rough Path Theory (online meeting). Organized by Thomas Cass, London; Dan Crisan, London; Peter Friz, Berlin; Massimiliano Gubinelli, Bonn. 6 December – 12 December 2020.* Mathematisches Forschungsinstitut Oberwolfach, April 2021. doi:10.14760/OWR-2020-40. xiii, 13

[PRS05] Frédéric Patras, Christophe Reutenauer, and Manfred Schocker. On the Garsia Lie Idempotent. *Canadian Mathematical Bulletin*, 48(3):445–454, September 2005. doi:10.4153/CMB-2005-041-x. 119

[PSS19] Max Pfeffer, Anna Seigal, and Bernd Sturmfels. Learning Paths from Signature Tensors. *SIAM Journal on Matrix Analysis and Applications*, 40(2):394–416, 2019. doi:10.1137/18M1212331. 65, 83

[Rah21] Ludwig Rahm. Translations of rough paths in combinatorial Hopf algebras, November 2021. arXiv:2111.02876 [math.CO]. 137, 138

[Ree58] Rimhak Ree. Lie Elements and an Algebra Associated With Shuffles. *Annals of Mathematics Second Series*, 68(2):210–220, September 1958. doi:10.2307/1970243. 3, 4, 69

[Rei19] Jeremy Francis Reizenstein. *Iterated-integral signatures in machine learning*. PhD thesis, University of Warwick, April 2019. http://webcat.warwick.ac.uk/record=b3431133~S15. 5, 14, 93, 128, 129

[Rei21] Jeremy F. Reizenstein. Python free_lie_algebra library. GitHub repository, 2021. https://github.com/bottler/free-lie-algebra-py. 88

[Reu93] Christophe Reutenauer. *Free Lie algebras*, volume 7 of *London Mathematical Society Monographs, New Series*. The Clarendon Press, Oxford University Press, New York, 1993. 9, 20, 22, 23, 66, 82, 88, 90, 91, 92, 94, 95, 98, 105, 107, 110, 117, 118, 119

[Roc03a] Eugénio Alexandre Miguel Rocha. *Uma abordagem algébrica à teoria de controlo não linear*. PhD thesis, Universidade de Aveiro, 2003. https://hdl.handle.net/10773/21444. 86, 88, 97, 99, 100, 101, 107

[Roc03b] Eugénio M. Rocha. On computation of the logarithm of the Chen-Fliess series for nonlinear systems. In Alan Zinober and David Owens, editors, *Nonlinear and Adaptive Control*, volume 281 of *Lecture Notes in Control and Information Sciences*, pages 317–326. Springer, Berlin, Heidelberg, 2003. doi:10.1007/3-540-45802-6_25. 86, 88, 98, 101, 107, 112, 115

[Roc05] Eugénio M. Rocha. A Lie-relation between noncommutative power sums symmetric functions, 2005. http://opac.ua.pt/cgi-bin/koha/opac-detail.pl?biblionumber=184776. 86, 88

[Row88] Louis H. Rowen. *Ring theory: Volume I*, volume 127 of *Pure and applied mathematics*. Academic Press, Inc., San Diego, CA, 1988. https://www.sciencedirect.com/bookseries/pure-and-applied-mathematics/vol/127/part/P1. 115

[Sal21] Cristopher Salvi. *Rough paths, kernels, differential equations and an algebra of functions on streams*. PhD thesis, University of Oxford, 2021. https://ora.ox.ac.uk/objects/uuid:7bcc0201-f8b2-4da6-8611-ea3f4536cad5. 88, 138

[SCF+21] Cristopher Salvi, Thomas Cass, James Foster, Terry Lyons, and Weixin Yang. The Signature Kernel is the solution of a Goursat PDE, March 2021. Version 9, arXiv:2006.14794v9 [math.AP]. 5

[Sch59] Marcel P. Schützenberger. Sur une propriété combinatoire des algèbres de Lie libres pouvant être utilisée dans un problème de mathématiques appliquées. Séminaire Dubreil. Algèbres et théorie des nombres, Volume 12 (1958-1959) no. 1, Talk no. 1, pages 1-23, 1958–1959. http://www.numdam.org/item/SD_1958-1959__12_1_A1_0/. 75, 76

[Sus91] Héctor J. Sussmann. Limits of the Wong-Zakai type with a modified drift term. In Eddy Mayer-Wolf, Ely Merzbach, and Adam Shwartz, editors, *Stochastic analysis: Liber Amicorum for Moshe Zakai*, pages 475–493. Academic Press, Boston, MA, 1991. doi:10.1016/B978-0-12-481005-1.50031-4. 16, 46

[TZ20] Nikolas Tapia and Lorenzo Zambotti. The geometry of the space of branched rough paths. *Proceedings of the London Mathematical Society*, 121(2):220–251, August 2020. `doi:10.1112/plms.12311`. 137

[Unt13] Jérémie Unterberger. A renormalized rough path over fractional Brownian motion. *Communications in Mathematical Physics*, 320(3):603–636, June 2013. `doi:10.1007/s00220-013-1707-0`. 19

[Yiz22] Yuan Yizheng. *Schramm-Loewner evolution and path regularity.* PhD thesis, Technische Universität Berlin, April 2022. `doi:10.14279/depositonce-15455`. xxi

[You36] L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Mathematica*, 67(1):251–282, 1936. `doi:10.1007/BF02401743`. 5

[Zam16] Lorenzo Zambotti. Renormalisation in regularity structures. Presentation slides, Workshop *Paths to, from and in renormalisation*, Potsdam, February 2016. `https://lmbp.uca.fr/~manchon/Potsdam2016/Slides/Zambotti-Potsdam2016.pdf`. 18

Frequently used notation

$\star$	The non-commutative associative Grossman-Larson product on $\mathcal{H}^*$, 28
$\bullet$	The non-commutative associative concatenation product of words, 3
$\blacksquare$	The associative shuffle-concatenation product on $\mathfrak{W}$, 89
$\trianglerighteq$	The left pre-Lie product on $\mathfrak{W}$ given as the tensor product of the halfshuffle $\succ$ and the Lie bracket $[\cdot,\cdot]$, 97
$\trianglerighteq_{\mathrm{Sym}}$	The symmetrization of the left pre-Lie product $\trianglerighteq$, 99
$\succ$	The non-associative right half-shuffle product of words satisfying the Zinbiel identity, 75
$\blacktriangleright$	The pre Lie product of smooth vector fields on $\mathbb{R}^e$, 31
$\curvearrowright$	The pre-Lie product on trees of the free pre-Lie algebra $\mathcal{B}(V)$, 31
$\mathscr{A}$	The area Tortkara algebra generated by the letters, 87
$\mathfrak{a}$	The universal homomorphism from the free associative algebra to the Grossman-Larson algebra with $\mathfrak{a}(\mathtt{i}) = \bullet_i$, 29
Ad_v	The linear map given by $\mathrm{ad}_{\mathtt{i}}\, w = [\mathtt{i}, w]$, $\mathrm{ad}_{\mathtt{i}v}\, w = [\mathtt{i}, \mathrm{ad}_v\, w]$, 98
ad_v	The linear map given by $\mathrm{ad}_v\, w = [v, w]$, 98
$\mathrm{ad}_{\blacksquare;x}$	The linear map given by $\mathrm{ad}_{\blacksquare;x}\, y = [x, y]_{\blacksquare}$, 110
Area	The signed area between two functions, 86
area	The algebraic area product given by $\mathsf{area}(x, y) = x \succ y - y \succ x$, 86
$\overleftarrow{\mathsf{area}}$	The left bracketing of area, 129
$\mathsf{area}_{\bullet}(\tau)$	The area bracketing according to the binary planar tree τ, 100
$\widetilde{\mathsf{area}_{\bullet}}(\tau)$	The area and $\sqcup$ bracketing according to the tree $\tau \in \widetilde{\mathsf{BPT}}$, 113
$\mathcal{B}$	The space of unordered rooted trees with vertices labelled from $\{0, \ldots, d\}$, 28
$\mathfrak{b}$	The linear map from $T(\mathbb{R}^{d+1})$ to $\mathcal{B}(\mathbb{R}^{d+1})$ given by $\mathfrak{b}(\mathsf{e}) := 0, \mathfrak{b}(\mathtt{i}) := \bullet_i, \mathfrak{b}(\mathtt{i}w) := \bullet_i \curvearrowright \mathfrak{b}(w)$, 45
$\mathcal{B}_-$	The subspace of $\mathcal{B}$ spanned by trees with no label f and with at most $\lfloor 1/\alpha \rfloor$ nodes, 57
BPT_n	The set of binary planar trees with n leaves, 100
$\widetilde{\mathsf{BPT}}_n$	The set of binary planar trees with nodes $\blacksquare$ and leaves in BPT, 112
$[.,.]_{\blacksquare}$	The Lie bracket given as the antisymmetrization of $\blacksquare$, 110

c	A coefficient function on BPT with $c(\tau) = 2c(\tau')c(\tau'')(\lvert\tau'\rvert_{\mathsf{leaves}} + \lvert\tau''\rvert_{\mathsf{leaves}} - 1)$, 100
D	The unique derivation of the $\bullet$ product with $D\mathtt{i} = \mathtt{i}$, 94
D^{-1}	The inverse of D on $T_{\geq 1}((\mathbb{R}^d))$, 94
$\underline{D}^{-1}$	The lift of D^{-1} to $\mathfrak{W}$ given by $\underline{D}^{-1}(p \otimes q) = p \otimes D^{-1}(q)$, 95
$\underline{D}$	The lift of D to $\mathfrak{W}$ given by $\underline{D}(p \otimes q) = p \otimes D(q)$, 95
δ	The algebra homomorphism from $\mathcal{H}$ to $\mathcal{A} \otimes \mathcal{H}$ defined on trees by extraction, 33
$\Delta_\star$	The Butcher-Connes-Kreimer coproduct dual to $\star$, 28
Δ^+	Coaction of $\mathcal{T}_+$ on $\mathcal{T}$, also coproduct on $\mathcal{T}_+$, later redefined as coaction of $\tilde{\mathcal{T}}_+$ on $\tilde{\mathcal{T}}$, and coproduct on $\tilde{\mathcal{T}}_+$, 52, 56
Δ^-	Coaction of $\mathcal{T}_-$ on $\mathcal{T}$, later restricted to a coaction of $\tilde{\mathcal{T}}_-$ on $\mathcal{T}$, also coproduct of $\mathcal{T}_-$, 51, 55
δ^-	The linear map $(\pi_- \otimes \mathrm{id})\delta : \tilde{\mathcal{H}} \to \mathcal{H}_- \otimes \tilde{\mathcal{H}}$, where π_- is the algebra homomorphism given by projecting elements of $\mathcal{B}$ onto terms of negative degree, 61
$\Delta_\bullet$	The deconcatenation coproduct on the tensor algebra dual to $\bullet$, 3
$\Delta_\odot$	The coproduct dual to $\odot$, 28
$\Delta_{\sqcup\!\sqcup}$	The deshuffle coproduct dual to $\sqcup\!\sqcup$, 23
$\mathcal{G}_-$	The group of characters on $\mathcal{T}_-$, 57
$\tilde{\mathcal{G}}_+$	The group of characters on $\mathcal{T}_+$, 58
$\mathcal{H}$	The space of forests composed of trees from $\mathcal{B}$, 28
$\mathcal{H}_-$	The free commutative algebra generated by the space $\mathcal{B}_-$, 57
$\tilde{\mathcal{H}}$	$\mathcal{H} \oplus \mathcal{H}\Xi_1 \oplus .... \oplus \mathcal{H}\Xi_d$, 56
I_w	The linear map given by $I_w x = w \succ x$, 98
$\mathsf{lie}_\bullet$	The $[\cdot,\cdot]$ bracketing according to the binary planar tree τ, 100
M_p	The Zinbiel homomorphism associated to a polynomial map p (do not confuse with M_v), 71
M_v	The branched translation map (do not confuse with M_p), 31
$\odot$	The commutative associative forest product on $\mathcal{H}$, 28
$(P_h)_{h\in H}$	A basis for the free Lie algebra $\mathfrak{g}(\mathbb{R}^d)$, 105
ϕ	The canonical bijective linear map from $\mathcal{B}_-$ to $\mathrm{span}_{\mathbb{R}}\,\mathcal{W}_-$, later extended to a linear bijection from $\mathcal{B}$ to $\tilde{\mathcal{H}}$, 57
π_1	The unique linear projection onto the Lie algebra $\mathfrak{g}(\mathbb{R}^d)$ whose series extension linearizes the logarithm $\log_\bullet : G \to \mathfrak{g}((\mathbb{R}^d))$, 107
$\pi_1^\top$	The linear map dual to π_1, 107
R	The primitive element $R \in T\langle\langle \mathcal{R}^d \rangle\rangle$ given by $R = \underline{r}(S)$, 95
r	The Dynkin operator, 94
$\underline{r}$	The lift of r to $\mathfrak{W}$ given by $\underline{r}(p \otimes q) = p \otimes r(q)$, 95
ρ	The linear map dual to the Dynkin operator r, 95

R_n	The n-th level of R, 99
S	The grouplike element $S \in T\langle\langle \mathcal{R}^d \rangle\rangle$ given by $S = \sum_w w \otimes w$, 95
$\mathfrak{s}$	A scaling of $\mathbb{R}^d$ associated to an SPDE, 7
$(S_h)_{h\in H}$	A collection of $S_h \in T(\mathbb{R}^d)$ such that $\langle S_h, P_{h'} \rangle = \delta_{h,h'}$, 107
shuff	The linear space of shuffles from below, 117
$⧢$	The commutative associative shuffle product of words, 68
$\sigma(X)$	The iterated-integral signature of the path X, 2
$\mathcal{T}$	The vector space spanned by the trees in $\mathcal{W}$, 50
$\mathcal{T}_+$	The free unital commutative algebra generated by $\mathcal{W}_+$, 52
$\mathcal{T}_-$	The free unital commutative algebra generated by the negative degree elements of $\mathcal{W}$, 51
$\dot{\tau}$	$\phi(\tau)$, 57
$\mathfrak{W}$	The product space of tensor products $\mathfrak{W} = \prod_{n=0}^{\infty} T(\mathbb{R}^d) \otimes T_n(\mathbb{R}^d)$, 89
T_v	The geometric translation map, 25
$\tilde{\mathcal{T}}$	The subspace of $\mathcal{T}$ spanned by $\tilde{\mathcal{W}}$, 54
$\tilde{\mathcal{T}}_+$	The subalgebra of $\mathcal{T}_+$ given as the free unital commutative algebra generated by $\tilde{\mathcal{W}}_+$, 55
vol	The algebraic volume map given by $\mathsf{vol}(x,y,z) := \mathsf{area}(\mathsf{area}(x,y),z) + \mathsf{area}(\mathsf{area}(y,z),x) + \mathsf{area}(\mathsf{area}(z,x),y)$, 126
$\mathcal{W}$	The set of all rooted trees with polynomial decorations on the nodes and where edges ending in leaves may carry a type $\mathfrak{t}_{\Xi_i}$, $i \in \{1,\dots,d\}$, 50
$\mathcal{W}_+$	The set containing the abstract variable X and $\mathcal{J}_k(\tau)$, $\tau \in \mathcal{W}$, 52
$\mathcal{W}_-$	The negative degree elements of $\mathcal{W}$ (or $\tilde{\mathcal{W}}$), 54
$\tilde{\mathcal{W}}$	The subset of $\mathcal{W}$ of trees without node decorations, 54
$\tilde{\mathcal{W}}_+$	The subset of $\mathcal{W}_+$ spanned by all $\mathcal{J}\tau$ for $\tau \in \mathcal{W}$, 55
$\overleftarrow{X}$	The path X parametrized backwards (time inversion), 1
Ξ_i	The abstract symbol corresponding to the driving noise ξ_i, 50
$(\zeta_h)_{h\in H}$	Coordinates of the first kind corresponding to $(P_h)_{h\in H}$, 105
$\sqcup$	Concatenation of paths, 1

Index