Directed polymers and rough paths

Nikolas Tapia

Sorbonne Université
Universidad de Chile
École Doctorale de Sciences Mathématiques de Paris Centre
Laboratoire de Probabilités, Statistique et Modélisation

Directed Polymers and Rough Paths

Par Nikolas Esteban Tapia Muñoz

Thèse de doctorat de mathématiques

Dirigée par Daniel Remenik Z. et Lorenzo Zambotti

Présentée et soutenue publiquement le 26 septembre 2018
Devant un jury composé de :
Cédric BOUTILLIER (Maître de Conférences)
Joaquín FONTBONA (Professeur)
Dominique MANCHON (Rapporteur)
Maria RONCO (Professeure)

To my wife Geraldine.
To my parents Jorge and Jeannette,
and to my brother Felipe.
I love you ♡.

Einstein was a giant.
His head was in the clouds,
but his feet were on the ground.
But those of us who are not that tall
have to choose!

— Richard Feynman.

Les gens savent ce qu'ils font :
souvent ils savent pourquoi ils font ce qu'ils font ;
mais ce qu'ils ignorent, c'est l'effet produit par ce qu'ils font.

— Michel Foucault

Acknowledgements

I would like to start by thanking both my advisors, Daniel and Lorenzo, for supporting me through this thesis. Without them none of this work would have been possible. Both of you gave me interesting problems to work on and guided me with great care and patience always in the right direction. Your guidance served as the main source of inspiration both for academic work and life itself. I am very grateful for all of the precious time you put into this endeavour during these three and something years which I cannot possibly repay, and for your aid during difficult times. I hope I have lived up to your expectations and I hope we can continue to work together in the future.

I would also like to thank my wife, Geraldine. This has not been an easy task, but we finally made it. You were a fundamental pillar all the way through, and if I managed to take this journey to a good end it was also because of you. Please know that you played an important part in this and that I feel very humbled to have had you have my back when I needed it. Now we start a new phase with new challenges and new horizons open up before us, and I know we will continue to conquer whatever difficulties may come our way, together.

My family also played an important role during the development of this thesis. Although we were apart for almost two years, I always felt supported and loved. I thank you for all the times you, knowingly or not, supported and pushed me to continue. I am here thanks to all of the opportunities you gave me during my life and for that I cannot be more grateful. The result of this work is partly yours in spirit.

Along the years I've met people who have helped me finish this work in one way or another, be it by discussing some of the problems or by providing me with much needed distraction in times of stress. I would specially like to thank, in no particular order, Amitai Linker, Felipe Muñoz, Avelio Sepúlveda, Nicolás Zalduendo, Felipe Subiabre, J. T. Méndez, Sebastián Riffo, Rodolfo Gutiérrez, Nicolás Hernández and Jocelyn Naranjo, Nicolás Zalduendo, Carlo Bellingeri, Heri Elad-Altmann, Sarah Kaakai, Paul Melotti, Nicolas Gilliers, Michel Pain, Carlos Román and Constanza Rojas, Matías Pavez, Mauricio Duarte, Manuel Cabezas and Claire Delplancke. I could not possible mention everyone but if I left you out of this list please be assured that it is because of space constraints and it does not mean in any way that I forgot about you. You are all in my mind. This goes to all of the many friends I have at DIM in Santiago and to the PhD students at the LPSM in Paris. I also thank the administrative staff at both departments: Éterin Jaña, Sylvia Mariano, Oscar Mori, Luis Mella, Gladys Cavallone and María Cecilia Cea at DIM and Florence Deschamps and Serena Bennasù from LPSM. Last but not least, I would like to thank some of my colleagues and professors with whom I've had many fruitful discussions that helped the development of this work in some form of the other: Prof. Joaquín Fontbona, Yvain Bruned, Ajay Chandra, Vlad Margarint, Rosa Preiß, Prof. Kurusch Ebrahimi-Fard, Prof.

Frédéric Patras, Prof. Dominique Manchon and Tom Klose.

I acknowledge financial support from the CONICYT/Doctorado Nacional/2013-21130733 doctoral scholarship, the Fondation de Sciences Mathématiques de Paris and Núcleo Milenio Modelos Estocásticos de Sistemas Complejos y Desordenados.

Contents

Part I

Preliminaries

Chapter 1

Introduction

In this chapter we review the main ideas of this thesis and informally explain some (if not all) of the results. We address each of the chapters in order, and suitable references to the actual results and formal presentation of the concepts are provided along the text.

One of the central ideas of this work is that of *renormalisation*. Roughly speaking, quantum physical systems often produce ill-defined quantities arising from infinities which appears when considering self-interactions at various scales. Even though at a fixed scale these interaction might be finite computed quantities it is possible that adding up an infinite number of scales produces infinities. Renormalisation is a way of dealing with these infinities by specifying relationships between the parameters of the theory in such a way to produce (a finite number of) what is called *counter-terms*. These terms are also diverging but in such a way that they exactly cancel the diverging quantities.

Even directly at the level of the continuum description of some Quantum Field Theory models the objects involved are ill-defined, and the naïve attempt at defining them as the continuum limit of some regularisation of the problem has to be carefully done. In general some form of renormalisation has to be performed. Ideally the limiting object, and hence the physical predictions of the theory, will not depend on the specific form of this prescription.

A wide amount of renormalisation procedures have been proposed by physicists throughout the years. Among these there is the Bogoliugov–Parasiuk–Hepp–Zimmerman (BPHZ for short) renormalisation procedure, which has been recently studied from a mathematical point of view by M. Hairer and others by means of the theory of Regularity Structures.

1.1 Hopf-algebraic deformations of products

A key step in various of the existing renormalisation procedures for perturbative QFTs is taking *Wick products*. In its simplest form, a Wick product associates to a pair of random variables X, Y with finite moments another random variable

$$:XY: \;\coloneqq XY - X\mathbb{E}Y - Y\mathbb{E}X + 2\mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[XY].$$

Although this definition might seem arbitrary at first it enjoys a rich structure intimately related to the *Leonov–Shiryaev relation* (eq. (1.5) below, see also Section 4.3.3) between moments and cumulants, and thus to moment generating functions. The first thing we remark is that $\mathbb{E}:XY: \;= 0$, that is, the Wick product is always a centred random variable independently of the specific distribution of X and Y provided that the involved expectations are finite, and this is in fact one of the defining properties of general Wick products. We also observe that $:XY:$ is a degree 2 polynomial in X and Y. Of particular interest is the Wick square $:X^2:$ and more in general the Wick powers of a single random variable with an appropriate amount of finite moments, which are then given by a polynomial in X of appropriate degree.

Definition 1.1.1. *Given a collection $(X_k : k \in \mathbb{N})$ of random variables with finite moments of all orders. For any multi-index α, the* Wick polynomial $:X_\alpha:$ *is a random variable recursively defined by setting $:X_0: \;\coloneqq 1$ and*

$$X^\alpha = \sum_{\beta+\gamma=\alpha} \binom{\alpha}{\beta, \gamma} \mathbb{E}[X^\beta] :X_\gamma: \;. \tag{1.1}$$

In the above definition by a multi-index we mean a function $\alpha : \mathbb{N} \to \mathbb{N}$ with finite support, and we have set up the notations

$$X^\alpha = \sum_{n\in\mathbb{N}} X_n^{\alpha_n}, \quad X_\alpha = \prod_{n\in\mathbb{N}} X_{\alpha_n}.$$

The Wick power $:X^n:$ is a polynomial of degree n in X. The specific form of this polynomial depends on the distribution of X and they are known explicitly in some specific cases. For example, if X has a standard Gaussian distribution then $:X^n: \;= h_n(X)$ and the polynomials h_n are known as Hermite polynomials. They are an orthogonal family in $L^2(\omega)$ where $\omega(x) = e^{-x^2/2}$ is the Gaussian weight, that is, they satisfy the relation

$$\int_{\mathbb{R}} h_n(x) h_m(x) \omega(x) \, \mathrm{d}x = \delta_{nm}.$$

An interesting as well as delicate example of an application of Wick products is the one appearing in Regularity Structures in the papers [15, 20, 56]. In this setting, the variable X

is space-time white noise which is a random distribution living in a Besov space of negative regularity. Usually one is interested in taking products of X with itself which are ill-defined. Then, one tries to regularise X by performing convolution with a smooth kernel depending on a small parameter $\varepsilon > 0$ and then trying to take the limit as $\varepsilon \to 0$. Naturally this limit cannot be well defined and some kind of renormalisation has to be performed. It turns out that altough X_ε will not converge in a suitable space since $\mathbb{E}X_\varepsilon^2$ will be proportional to some negative power of ε, the Wick square $:X_\varepsilon^2:$ is centered so the sequence is bounded in L^2.

In Chapter 4 we construct a Hopf algebra describing Wick products as the action of a linear functional on it. Thus we abstract the combinatorial properties of Wick polynomials into a combinatorial setting. In particular, we generalise Wick polynomials to arbitrary multivariate distributions with finite moments of all orders. We then use this Hopf-algebraic presentation of Wick polynomials to relate Wick products with Hopf-algebraic deformations of the classical product of polynomials. More precisely, given a collection $(X_a : a \in A)$ of random variables we let $H = \mathbb{R}[x_a : a \in A]$ be the commutative polynomial algebra on formal variables $(x_a : a \in A)$, i.e. H is the vector space spanned by monomials x^α and product is given by $x^\alpha \cdot x^\beta = x^{\alpha+\beta}$. The coproduct $\Delta : H \to H \otimes H$ is given by

$$\Delta x^\alpha = \sum_{\beta+\gamma=\alpha} \binom{\alpha}{\beta, \gamma} x^\beta \otimes x^\gamma.$$

Then, we define linear maps $\mu : H \to \mathbb{R}$ and $W : H \to H$ by letting $\langle \mu, x^\alpha \rangle := \mathbb{E}[X^\alpha]$ and $W(x^\alpha) := \, :x^\alpha:$. With these notations, eq. (1.1) can be rewritten as an identity of linear maps on H as

$$\mu \star W = \mathrm{id}$$

where $\star$ is the convolution product $(\phi \star \psi)(x^\alpha) = (\phi \star \psi)\Delta x^\alpha$. This gives a way to write the Wick product $:X^\alpha:$ in terms of μ.

Theorem 1.1.2. *For all multi-indices α we have*

$$W(x^\alpha) = (\mu^{-1} \star \mathrm{id})(x^\alpha) \tag{1.2}$$

where μ^{-1} is the convolutional inverse *of the linear map μ.*

Observe that although it is not apparent from the notation, this definition depends heavily on μ. We call $W(x^\alpha)$ the generalised Wick polynomial associated to the distribution of $(X_a : a \in A)$. In particular, using that the inverse of μ can be explicitly computed we obtain the following formula for $:x^\alpha:$.

Theorem 1.1.3. *Wick polynomials have the explicit expression*

$$:x^\alpha: \;= x^\alpha + \sum_{n\geq 1}(-1)^n \sum_{\beta+\gamma_1+\ldots+\gamma_n=\alpha} \binom{\alpha}{\gamma_1,\ldots,\gamma_n,\beta} \mu(x^{\gamma_1})\cdots\mu(x^{\gamma_n})x^\beta. \tag{1.3}$$

The group $\mathscr{G}(H) = \{\lambda \in H^* : \lambda(\mathbf{1}) = 1\}$ acts canonically on H given by means of the map $\phi_\lambda : H \to H$ defined as

$$\phi_\lambda(x^\alpha) := (\lambda \otimes \mathrm{id})\Delta x^\alpha = (\lambda \star \mathrm{id})(x^\alpha).$$

This is a group action since one can verify that $\phi_{\lambda_1\star\lambda_2} = \phi_{\lambda_1}\circ\phi_{\lambda_2}$ so that in particular $(\phi_\lambda)^{-1} = \phi_{\lambda^{-1}}$. These maps can be used to induce a *deformation* of the Hopf algebra structure of H by setting

$$\begin{aligned} x^\alpha \cdot_\lambda x^\beta &:= \phi_{\lambda^{-1}}(\phi_\lambda(x^\alpha)\phi_\lambda(x^\beta)),\\ \Delta_\lambda x^\alpha &:= (\phi_\lambda^{-1} \otimes \phi_\lambda^{-1})\Delta\phi_\lambda(x^\alpha),\\ \langle \varepsilon_\lambda, x^\alpha\rangle &:= \langle \varepsilon, \phi_\lambda(x^\alpha)\rangle = \langle \lambda, x^\alpha\rangle. \end{aligned}$$

We then have

Theorem 1.1.4. *For any $\lambda \in \mathscr{G}(H)$ the quintuple $(H, \cdot_\lambda, \mathbf{1}, \Delta_\lambda, \varepsilon_\lambda)$ defines a Hopf algebra, and the map ϕ_λ^{-1} is a Hopf algebra isomorphism from the initial Hopf algebra H onto the deformed one.*

In the particular case where $\lambda = \mu$ is the map representing the moments of the collection $(X_a : a \in A)$ defined above and Theorem 1.1.2 we obtain

Theorem 1.1.5. *The Wick product map $W : H \to H$ is equal to $\phi_{\mu^{-1}}$. In particular $W : (H, \cdot, \Delta) \to (H, \cdot_\mu, \Delta_\mu)$ is a Hopf algebra isomorphism and*

$$:x^\alpha \cdot x^\beta: \;= \;:x^\alpha: \;\cdot_\mu\; :x^\beta:$$

for all monomials $x^\alpha, x^\beta \in H$.

One of the main difficulties for the application of Theorem 1.1.2 is the computation of the inverse μ^{-1} appearing in eq. (1.2). For a general element $\mu \in \mathscr{G}(H)$ its inverse is given by the formal series

$$\mu^{-1} = \sum_{n\geq 0}(\varepsilon - \mu)^{\star n}.$$

We prove that in fact this computation can be simplified by introducing a larger Hopf algebra structure $\hat{H}$ such that H is a left comodule over it, and such that every element in $\mathscr{G}(H)$ lifts to a unique character $\hat{\mu}$ over $\hat{H}$. In particular, the convolutional inverse $\hat{\mu}^{-1}$ can be computed by means of the extended antipode $\hat{S}$ on $\hat{H}$, i.e. we have that $\hat{\mu}^{-1} = \hat{\mu} \circ \hat{S}$. The advantage in this situation is that $\hat{H}$ has a graded structure and thus the antipode $\hat{S}$ has a recursive formula when applied to an arbitrary monomial in $\hat{H}$. Concretely, we define $\hat{H}$ to be the free commutative polynomial algebra over H, so that as a vector space $\hat{H}$ is also spanned by the monomials x^α but product is just juxtaposition $x^\alpha \bullet x^\beta$. That is, we consider each monomial in H to be a different variable in $\hat{H}$. For example, for any single variable we have $x_a \cdot x_a = x_a^2 \neq x_a \bullet x_a$ in $\hat{H}$. There is a canonical way to extend the coproduct $\Delta : H \to H \otimes H$ to a coproduct $\hat{\Delta} : \hat{H} \to \hat{H} \otimes \hat{H}$ so that $(\hat{H}, \bullet, \hat{\Delta})$ is a graded connected Hopf algebra with antipode $\hat{S}$ given explicitly by the formula

$$\hat{S}x^\alpha = -x^\alpha + \sum_{n\geq 2}(-1)^n \sum_{\beta_1+\cdots+\beta_n=\alpha} \binom{\alpha}{\beta_1, \dots, \beta_n} x^{\beta_1} \bullet \cdots \bullet x^{\beta_n}$$

whence we deduce eq. (1.3) by restriction.

Finally, we generalise this construction to a larger class of *monomials* appearing in the theory of Regularity Structures. In this case, the underlying Hopf algebra is the Butcher–Connes–Kreimer Hopf algebra over the collection of non-planar decorated rooted trees. In [56] these trees are used to represent generalised Taylor expansions, and related to the problem of multiplication of distributions. Our setting then allows us to interpret the renormalised products in [20, 56] as Hopf algebraic deformations of the pointwise product of smooth functions by using the Hopf algebra structure over trees in a similar way, thus connecting Regularity Structures with Wick renormalisation. In particular, we identify the BPHZ renormalisation character appearing in [15, 20] as a Wick product of trees in the sense of Theorem 1.1.2.

Concretely, in the language of Regularity Structures a random distribution can be considered as a random linear map X over trees taking values in the space of continuous functions on $\mathbb{R}^d$, such that $\langle X, \mathbf{1}\rangle = 1$. If we suppose that $\langle X, \tau\rangle(0)$ has finite moments of all orders so that we may define a linear map on trees by setting

$$\mu(\tau) \coloneqq \mathbb{E}[\langle X, \tau\rangle(0)]$$

Theorem 1.1.6. *Suppose that X is stationary in the sense that $\langle X, \tau\rangle(x + \cdot)$ has the same law as $\langle X, \tau\rangle$ for all trees τ and $x \in \mathbb{R}^d$. The only character λ on forests such that*

$$\mathbb{E}[\langle \hat{X}, \psi_\lambda \tau\rangle(0)] = 0$$

for all non-empty forests τ coincides with $\hat{\mu}^{-1}$, where $\hat{X}$ is the unique character extension to forests of X.

We also use this Hopf-algebraic language to describe generalised versions of the Leonov-Shiryaev relation between moments and cumulants. Moment–cumulant relations appear naturally in the context of stochastic integration with respect to random measures, the Itô integral being an specific example of this broader theory. In particular they are related to chaos expansions.

See for example G. Peccati and M. Taqqu's book [100]. In the case of a univariate distribution X with finite moments of all orders, if we denote its moments by $\mu_n := \mathbb{E}X^n$, the sequence of *cumulants* (κ_n) is defined by the relation

$$\exp\left(\sum_{n=1}^{\infty}\frac{t^n}{n!}\kappa_n\right) = \sum_{n=0}^{\infty}\frac{t^n}{n!}\mu_n \tag{1.4}$$

and in fact the sequence (κ_n) characterises the distribution of X. Some classical examples include the standard Gaussian distribution whose cumulants are $\kappa_1 = 0$, $\kappa_2 = 1$ and $\kappa_n = 0$ for all $n \geq 3$. The Poisson distribution of parameter $\lambda > 0$ is characterised by the constant sequence $\kappa_n = \kappa$ for all $n \geq 1$. Equation eq. (1.4) is actually equivalent to the Leonov–Shiryaev relations

$$\mu_n = \sum_{m=1}^{n}\binom{n-1}{m-1}\kappa_m\mu_{n-m}$$

which are closely related to the Bell polynomials and Faà di Bruno's formula.

With the same notations as in eq. (1.1) we can define joint cumulants $(\mathbf{E}_c[X_\alpha])$ in terms of joint moments of a collection $(X_a : a \in A)$ of random variables via the generalised Leonov-Shiryaev relation

$$\mathbb{E}[X^\alpha] = \sum_{n=1}^{|\alpha|}\frac{1}{n!}\sum_{\beta_1+\cdots+\beta_n=\alpha}\binom{\alpha}{\beta_1,\ldots,\beta_n}\prod_{i=1}^{n}\mathbf{E}_c[X_{\beta_i}]. \tag{1.5}$$

This equation gives a recursive definition of $\mathbf{E}_c[X_\alpha]$ by induction over the sum

$$|\alpha| := \sum_{a\in A}\alpha_a.$$

If we define the linear functionals $\mu : H \to \mathbb{R}$ and $\kappa : H \to \mathbb{R}$ by

$$\mu(x^\alpha) := \mathbb{E}[X^\alpha], \quad \kappa(x^\alpha) := \mathbf{E}_c[X_\alpha]$$

as before, we can lift eq. (1.5) into our abstract setting.

Proposition 1.1.7. *We have*

$$\mu = \exp_\star(\kappa) = \varepsilon + \sum_{n\geq 1}\frac{1}{n!}\kappa^{\star n}.$$

We can invert this relation and obtain

Proposition 1.1.8. *We have*

$$\kappa = \log_\star(\mu) = \sum_{n\geq 1}\frac{(-1)^n}{n}(\mu-\varepsilon)^{\star n}.$$

In particular,

$$\mathbf{E}_c[X_\alpha] = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} \sum_{\beta_1+\cdots+\beta_n=\alpha} \binom{\alpha}{\beta_1,\ldots,\beta_n} \prod_{i=1}^{n} \mathbb{E}[X^{\beta_i}].$$

1.2 Modifying Rough Paths

Rough paths theory was initiated by T. Lyons twenty years ago in the seminal paper [86]. The main goal of his approach is to systematically treat differential equations of the form

$$\mathrm{d}y_t = \sum_{a\in A} f_a(y_t)\,\mathrm{d}x_t^a, \quad y_0 = \xi \tag{1.6}$$

where A is a finite index set with $|A| = d$, $f_a : \mathbb{R} \to \mathbb{R}$ are smooth functions and x^a are some irregular paths defined over a compact time interval. In some situations the path x^a will be so irregular that the derivatives appearing in eq. (1.6) might only be defined as distributions. In this case, the classical theory of Differential Equations will not be able to provide an answer to this problem. This kind of equations are common for example when dealing with Stochastic Differential Equations (SDEs) driven by Brownian paths.

The classical approach is to assume some extra structure on the paths, such as in K. Itô's theory of stochastic integration. In this approach the path x is assume to come from a realisation of a stochastic process, adapted to an appropriate filtration, having a semi-martingale property and finite quadratic variation. Solutions are then only defined outside a set of probability zero, and are never defined at a given path. Moreover, the solution map $x \mapsto y$ where y is the Itô solution to eq. (1.6) is not continuous: this is illustrated by the Wong–Zakai theorem [119, 120], stated here in a somewhat loose version –in particular, the exact hypothesis have been relaxed.

Theorem 1.2.1. *Let B denote a standard Brownian motion and let $(B^\varepsilon)_{\varepsilon>0}$ be a sequence of smooth approximations. Let y^ε be the classical solution to the ODE*

$$\dot{y}_t^\varepsilon = f(y_t^\varepsilon)\dot{B}_t^\varepsilon$$

where f is a smooth function. Then, as $\varepsilon \to 0$, the process y^ε converges a.s. to the solution y of the Stratonovich *SDE*

$$\mathrm{d}y_t = f(y_t) \circ \mathrm{d}B_t.$$

Therefore, in the Brownian case, the Wong–Zakai theorem [119, 120] gives an answer: for a class of "reasonable" approximations to Brownian motion, the solutions to eq. (1.6) converge to the stochastic differential equation interpreted in the Stratonovich sense. It is still possible to recover the Itô interpretation in the limit, but one has to apply some corrections to the approximations. This can be seen as an instance of *renormalisation*. Note however that there are other kinds of reasonable approximations to Brownian motion for which y^ε does not converge. In 1991, Lyons showed that in fact an even stronger statement is true: any subset $\mathscr{C} \subset C([0,1])$ such that the integral $\int_0^1 u \,\mathrm{d}x$ is defined for every $x, u \in \mathscr{C}$ must have Wiener measure zero. Hence, this probabilistic approach cannot provide a pathwise meaning to rough equations of the form eq. (1.6).

Lyons' novel insight was to postulate that the only missing information in order to make pathwise sense out of an driven equation such as eq. (1.6) are the iterated integrals of the path $\boldsymbol{x} = (x^1, \dots, x^d)$ against itself. This is motivated by the following observation: for simplicity assume that $d = 1$ and that the path x is of class C^1. Then, by a solution of eq. (1.6) one means a C^1 path y such that

$$y_t = \xi + \int_0^t f(y_s)\dot{x}_s \,\mathrm{d}s.$$

If we compute the increment $\delta y_{st} := y_t - y_s$, by Taylor expanding f up to first order we find that

$$\delta y_{st} = f(y_s)\delta x_{st} + R_{st} \tag{1.7}$$

where the remainder

$$R_{st} := \int_s^t \delta f(y)_{su}\dot{x}_u \,\mathrm{d}u$$

satisfies the bound $|R_{st}| \lesssim \|x\|_{C^1}\|y\|_\infty |t-s|^2$ and the identity $\delta R_{sut} = \delta f(y)_{su}\delta x_{ut}$, where δR is the second-order finite increment $\delta R_{sut} := R_{st} - R_{su} - R_{ut}$. Approximation (1.7) is also the basis for some numerical schemes extending the usual Euler's method to the controlled setting, and to more irregular paths. See [34, 46, 79].

A classical result of L. C. Young states that it is possible to continuously extend the integral operator

$$I(u,x)_t = \int_0^t u_s \dot{x}_s \,\mathrm{d}s$$

from $C^0 \times C^1 \to C^1$ to $C^\gamma \times C^\alpha \to C^\alpha$ as long as $\gamma + \alpha > 1$ [121]. Moreover, this extension is the unique C^α function $I(u,x)$ satisfying

$$\delta I(u,x)_{st} = u_s \delta x_{st} + o(|t-s|)$$

with the precise estimate on the remainder

$$|\delta I(u,x)_{st} - u_s\delta x_{st}| \lesssim |t-s|^{\gamma+\alpha}.$$

Therefore we can still make sense of eq. (1.6) in the case where the trajectories of the driving process x lie on the Hölder space C^γ, for any $\gamma > \frac{1}{2}$. Of course, almost surely Brownian motion does not belong to this space.

However, observe that in this case the solution y satisfies

$$\delta y_{st} = f(y_s)\delta x_{st} + \frac{1}{2} f(y_s) f'(y_s) \int_s^t \delta x_{su}\, \mathrm{d}x_u + R_{st}.$$

The two-parameter function $\mathbb{X}_{st} = \int_s^t \delta x_{su}\, \mathrm{d}x_u$ is such that $|\mathbb{X}_{st}| \lesssim |t-s|^{2\gamma}$ and its second-order increments are $\delta\mathbb{X}_{sut} = \delta x_{su}\delta x_{ut}$. Moreover, Young's estimate implies that the remainder

$$R_{st} \coloneqq \delta y_{st} - f(y_s)\delta x_{st} - \tfrac{1}{2}\mathbb{X}_{st}$$

satisfies $|R_{st}| \lesssim |t-s|^{3\gamma}$ and its finite increments are also know. Observe that in this scenario, the function $A_{st} \coloneqq f(y_s)\delta x_{st} + f(y_s)f'(y_s)\mathbb{X}_{st}$ satisfies the hypothesis of Gubinelli's Sewing Lemma (Theorem 3.1.3) as long as $\gamma \geq \frac{1}{3}$, hence we can find an *integral* $I : [0,1] \to \mathbb{R}$ such that

$$\delta I_{st} = f(y_s)\delta x_{st} + f(y_s)f'(y_s)\mathbb{X}_{st} + \mathrm{o}(|t-s|),$$

which is a possible reformulation of eq. (1.6) in this context. At this point, the existence of $\mathbb{X}$ does not depend on the ability to integrate x against itself but only on its defining properties, i.e. on the form of its second-order increment and its regularity at the diagonal. In the one-dimensional case it is easily seen that setting $\mathbb{X}_{st} = \frac{1}{2}(x_t - x_s)^2$ works for any γ-Hölder path as long as $\gamma \geq \frac{1}{3}$. One can iterate this idea and obtain that if $\gamma \geq \frac{1}{N+1}$ then it suffices to find appropriate objects $\mathbb{X}^k$, for $k = 2, \ldots, N$, satisfying similar analytical and algebraic conditions. In this way we arrive at our first definition of a *rough path*.

Definition 1.2.2. *A* γ-rough path *over x is a collection* $(\mathbb{X}^k : k = 1, \ldots, N)$ *such that* $\mathbb{X}^1_{st} = \delta x_{st}$,

$$\mathbb{X}^k_{st} = \sum_{j=1}^{k} \mathbb{X}^j_{su}\mathbb{X}^{k-j}_{ut}$$

and $|\mathbb{X}^k_{st}| \lesssim |t-s|^{k\gamma}$ *for all* $k \in \mathbb{N}$ *and* $s, u, t \in [0,1]$.

So, if $\mathbb{X}$ is a γ-rough path over x the germ $A_{st} = f(y_s)\delta x_{st} + f(y_s)f'(y_s)\mathbb{X}^2_{st} + \cdots$ will satisfy the hypothesis of Theorem 3.1.3 and so one can recast eq. (1.6) in a suitable way. For some choices of x, a "canonical" choice of a rough path is available: for example, if x is smooth then its *signature* [23, 86] serves this purpose. Other cases where geometric rough paths have been constructed are Brownian motion and fractional Brownian motion (see [27] for the $H > \frac{1}{4}$ case and [96] for the general case) among others. Still one might require additional properties from the integral given by the Sewing Lemma which might not be obvious. For example, it is not clear that integrals constructed this way should satisfy integration by parts for any choice of $\mathbb{X}^k$. In a stochastic integration context, the notion of integral so obtained coincides with **Stratonovich's** integral rather than Itô's version. If one would like to obtain Itô's stochastic integral, it is mandatory to let $\mathbb{X}_{st} = \frac{1}{2}((x_t - x_s)^2 - (t-s))$ as one would expect given the interpretation of $\mathbb{X}$ as an iterated integrals. In fact, the properties of the finite integral operator δ (see Section 3.1.1) entail that given a fixed second order process $\mathbb{X}$, setting $\mathbb{X}'_{st} = \mathbb{X}_{st} + \delta h_{st}$, for any 2γ-Hölder function h, yields another valid second order process, hence a different notion of integral. Therefore, modulo a shift by a 2γ-Hölder function all γ-rough paths are of the above form.

The case in dimension $d \geq 2$ is drastically different. For simplicity we only address the case $d = 2$, but the same remarks hold for any $d \geq 2$. When $\boldsymbol{x} = (x^1, x^2)$ has more than one component we have to deal with integrals of the components against themselves. As before, if $\boldsymbol{x}$ is smooth or γ-Hölder for $\gamma \geq \frac{1}{2}$ the integral

$$\mathbb{X}^{ij}_{st} = \int_s^t \delta x^i_{su} \, \mathrm{d}x^j_u \tag{1.8}$$

is well defined and it satisfies $\delta\mathbb{X}^{ij}_{sut} = \delta x^i_{su}\delta x^j_{ut}$ and $|\mathbb{X}^{ij}_{st}| \lesssim |t-s|^{2\gamma}$. For a general γ-Hölder path $\boldsymbol{x}$ we may proceed as before, postulating the existence of "iterated integrals" $\mathbb{X}^{ij}$, and we can recourse to Theorem 3.1.3 in order to build a notion of integral against $\boldsymbol{x}$, hence define a notion of solution to the corresponding multi-dimensional version of eq. (1.6). However, in this case the algebraic relations that these integrals ought to satisfy are not as simple, and in particular there is no canonical choice of the values of $\mathbb{X}^{ij}$ for an arbitrary path $x \in C^\gamma$. For example, the path $\mathbb{X}^{ij}_{st} = -\delta x^i_s \delta x^j_{st}$ satisfies $\delta\mathbb{X}^{ij}_{sut} = \delta x^i_{su}\delta x^j_{ut}$, but it behaves only as $|t-s|^\gamma$ when $t \sim s$, not as $|t-s|^{2\gamma}$ as required. We may think of the collection $\mathbb{X}^{ij}$ as a matrix (2-tensor) on the indices $i, j \in \{1, \ldots, d\}$ and as such we may decompose it as a sum of a symmetric matrix $\mathbf{S}$ and an anti-symmetric matrix $\mathbf{A}$ where

$$\mathbf{S}^{ij} := \frac{\mathbb{X}^{ij} + \mathbb{X}^{ji}}{2}, \quad \mathbf{A}^{ij} := \frac{\mathbb{X}^{ij} - \mathbb{X}^{ji}}{2}.$$

By using the properties of the finite increment δ (see Section 3.1.1, in particular eq. (3.4)) it is not hard to see that $\mathbf{S}^{ij}_{st} = \frac{1}{2}\delta x^i_{st}\delta x^j_{st}$ satisfies

$$\delta\mathbf{S}^{ij}_{sut} = \frac{1}{2}(\delta x^i_{su}\delta x^j_{ut} + \delta x^j_{su}\delta x^i_{ut}) \tag{1.9}$$

so the symmetric part can always be defined for an arbitrary path, but finding the anti-symmetric part is a non-trivial task. Indeed, in the iterated integral interpretation the matrix $\mathbf{A}$ has entries given by

$$\mathbf{A}^{ij}_{st} = \frac{1}{2}\int_s^t \delta x^i_{su} \, \mathrm{d}x^j_u - \delta x^j_{su} \, \mathrm{d}x^i_u \tag{1.10}$$

so this matrix is tightly related to the construction of the Lévy area process. This process is known to be discontinuous in the uniform topology, and even in the 2-variation topology, as a function of $\boldsymbol{x}$ so approximations by smooth functions are immediately ruled out as a valid strategy. In fact, we remark that since $\mathbf{A}$ is antisymmetric the couple $\delta x + \mathbf{A}$ actually lives in the vector space $\mathbb{R}^d \oplus \mathfrak{so}(d) := \mathfrak{g}^2$. This vector space is the *free 2 step nilpotent Lie algebra* where the Lie bracket is given by

$$[x + \mathbf{A}, y + \mathbf{B}] = x \otimes y - y \otimes x.$$

Moreover, there is an exponential $\exp_2 : \mathfrak{g}^2 \to G^2$ given by

$$\exp_2(x + \mathbf{A}) := \mathbf{1} + x + \mathbf{A} + \frac{1}{2} x \otimes x \tag{1.11}$$

where $G^2 := \mathbf{1} + \mathbb{R}^d \oplus (\mathbb{R}^d \otimes \mathbb{R}^d)$ is the *Lie group* associated to $\mathfrak{g}^2$ which is free 2 step nilpotent. The multiplication law on G^2 can be explicitly stated as

$$(\mathbf{1} + x + \mathbb{X}) \cdot (\mathbf{1} + y + \mathbb{Y}) = \mathbf{1} + x + y + \mathbb{X} + \mathbb{Y} + x \otimes y$$

but can also be described by means of the Baker–Campbell–Hausdorff formula (see eq. (5.7)). Observe from eq. (1.11) that the exponential actually adds the symmetric matrix $\frac{1}{2} x \otimes x$ to $\mathbf{A}$ thus obtaining a matrix $\mathbb{X} \in \mathbb{R}^d \otimes \mathbb{R}^d$ with symmetric part fixed by x in the fashion of $\mathbf{S}$. Thus, defining the antisymmetric matrix $\mathbf{A}$ in eq. (1.10) is equivalent to finding a matrix $\mathbb{X}$ satisfying the algebraic constraint below eq. (1.8) provided that we fix the symmetric part as in eq. (1.9). This choice of the symmetric part is clearly not unique but it is in some sense canonical; indeed, the identity $\mathbb{X}^{ij}_{st} + \mathbb{X}^{ji}_{st} = \delta x^i_{st} \delta x^j_{st}$ commes from the integration by parts rule

$$\int_s^t \delta x^i_{su} \, \mathrm{d}x^j_u + \int_s^t \delta x^j_{su} \, \mathrm{d}x^i_u = \delta x^i_{st} \delta x^j_{st} \tag{1.12}$$

for smooth paths $\boldsymbol{x} = (x^1, x^2)$. This identity is an example of a more general set of relations known as *shuffle relations*, in reference to the flipping of the indices appearing in the previous equation.

This can be formalised and generalised by taking the dual point of view. Consider the set $A = \{1, \ldots, d\}$ and let $\mathrm{M}(A)$ denote the free monoid over A. Elements of A can be regarded as words with letters from A and we let $\mathbf{1}$ denote the empty word which acts as the identity for concatenation. The *shuffle product* $\sqcup\!\sqcup : H \otimes H \to H$ is defined recursively by $a \sqcup\!\sqcup \mathbf{1} = \mathbf{1} \sqcup\!\sqcup a = a$ for all $a \in A$ and

$$au \sqcup\!\sqcup bv = a(u \sqcup\!\sqcup bv) + b(au \sqcup\!\sqcup v)$$

where $a, b \in A$ and $u, v \in \mathrm{M}(A)$ are non-empty words. The space H also carries a coproduct $\bar{\Delta} : H \to H \otimes H$ given by deconcatenation of words

$$\bar{\Delta}(a_1 \cdots a_n) = \sum_{k=0}^{n} a_1 \cdots a_k \otimes a_{k+1} \cdots a_n$$

in such a way that $(H, \sqcup\!\sqcup, \bar{\Delta})$ becomes a graded connected Hopf algebra. It is known that linear functionals $X \in H^*$ which are multiplicative in the sense that $\langle X, u \sqcup\!\sqcup v\rangle = \langle X, u\rangle\langle X, v\rangle$ form a group. Moreover, if we identify words of length greater than two with zero, the truncated group so obtained is isomorphic to G^2 defined above. We arrive at the following definition (see Section 5.4 for the general case).

Definition 1.2.3. *Let $\gamma \in [\frac{1}{3}, \frac{1}{2})$. A weakly γ-regular* geometric rough path *is a path $X : [0,1]^2 \to G^2$ such that $X_{st} = X_{su} \star X_{ut}$ and*

$$|\langle X_{st}, u\rangle| \lesssim |t-s|^{\gamma|v|}.$$

Notice that in this case the shuffle relation in eq. (1.12) might be rewritten as

$$\langle X_{st}, i \sqcup\!\sqcup j\rangle = \langle X_{st}, i\rangle\langle X_{st}, j\rangle$$

so the fact that X takes values in the group G^2 is just an expression of the fact that the integral represented by X satisfies integration by parts. Also, if we write Chen's rule $X_{st} = X_{su} \star X_{ut}$ in coordinates we see that

$$\mathbb{X}^{ij}_{st} = \langle X_{st}, ij\rangle = \langle X_{su} \star X_{ut}, ij\rangle = \mathbb{X}^{ij}_{su} + \mathbb{X}^{ij}_{ut} + \delta x^i_{su}\delta x^j_{ut}$$

i.e. we recover the algebraic constraint we had before, namely $\delta\mathbb{X}^{ij}_{sut} = \delta x^i_{su}\delta x^j_{ut}$.

A general theorem by T. Lyons and N. Victoir [88] gives the existence of a geometric rough path over any given γ-Hölder path. The precise result is as follows:

Theorem 1.2.4 (Lyons–Victoir extension)**.** *If $p \in [1, \infty) \setminus \mathbb{N}$, a continuous path of finite p-variation can be lifted to a geometric p-rough path. For any p, a continuous path of finite finite p-variation can be lifted to a geometric $(p+\varepsilon)$-rough path.*

Of course, this require extending Definition 1.2.3 to lower regularities by the addition of supplementary higher-order components, which can be seen in a similar way to represent iterated integrals and higher-order Lévy area processes. This theorem and its proof are very abstract and thus do not provide a concrete way of constructing the required iterated integrals. In Chapter 5 (see in particular Theorem 5.3.4), we provide an alternative approach, inspired by the Lyons–Victoir construction, to construct these integrals in an iterative manner. In its simplest form, namely when $\gamma \in [\frac{1}{3}, \frac{1}{2})$ and $d = 2$ we only need to construct one path since $\dim \mathfrak{so}(2) = 1$. This may be done by a technique akin to Gubinelli's *Sewing Lemma* (Theorem 3.1.3), by first defining the antisymmetric matrix $\mathbf{A}$ over a dyadic partition of the interval $[0,1]$ (say). We first define the values of $\mathbf{A}_{st}$ for *consecutive* dyadics $s = k2^{-m}, t = (k+1)2^{-m}$ by making some choices: for example, we start by setting $\mathbf{A}_{01} = 0$, and this choice is arbitrary. We may as well replace this initial value by any antisymmetric matrix of our choice and obtain a different rough path. Now, Chen's rule can be rephrased at the level of $\mathfrak{g}^2$ by using the BCH formula (see eq. (5.7) below): in this case we have that, since $\mathbb{X} = \exp_2(\delta x + \mathbf{A})$,

$$\begin{aligned}\exp_2(\delta x_{st} + \mathbf{A}_{st}) &= \exp_2(\delta x_{su} + \mathbf{A}_{su}) \cdot \exp_2(\delta x_{ut} + \mathbf{A}_{ut}) \\ &= \exp_2\left(\delta x_{st} + \mathbf{A}_{su} + \mathbf{A}_{ut} + \frac{1}{2}[\delta x_{su} + \mathbf{A}_{su}, \delta x_{ut} + \mathbf{A}_{ut}]\right)\end{aligned}$$

whence

$$\delta\mathbf{A}_{sut} = \mathbf{A}_{st} - \mathbf{A}_{su} - \mathbf{A}_{ut} = \frac{1}{2}(\delta x_{su} \otimes \delta x_{ut} - \delta x_{ut} \otimes \delta x_{su}). \tag{1.13}$$

Therefore, we also need to choose the values of $\mathbf{A}$ on the dyadics in such a way so that eq. (1.13) is preserved. Then, we use Chen's rule again in order to extend this definition to any pair of dyadics. At last, a continuity argument (Lemma 5.1.6) allows the passage to the limit and define the value of $\mathbf{A}$ for any $s, t \in [0, 1]$. For this last step we need the fact that there is a suitable metrisable topology in G^2 compatible in a nice way with the analytic bound in Definition 1.2.3.

Remark 1.2.5. The previous argument, i.e. the case $\gamma \in [\frac{1}{3}, \frac{1}{2}[$ and $d = 2$, already appears in [88]. △

Finally, a word about the non-geometric case. In some situations, the imposition of the usual rules of calculus might lead to an "undesirable" notion of integral; such is the case of Brownian motion where the Itô integral might be preferred over Stratonovich's since, even though both are adapted, the latter is the limit of non-adapted sums so it needs to "look into the future". Moreover the Itô integral enjoys a martingale property and it is an isometry into the space of square integrable functions; none of this is true for the Stratonovich integral. The shuffle relations mentioned in the above paragraph present the integration by parts rule in an algebraic language. For example, if x^1 and x^2 are smooth paths and we set $\mathbb{X}^{ij}_{st} = \int_s^t \delta x^i_{su}\,\mathrm{d}x^j_u$, the usual integration by parts implies that $\mathbb{X}^{ij}_{st} + \mathbb{X}^{ji}_{st} = \delta x^i_{st}\delta x^j_{st}$. In 1977, K. T. Chen showed [23] that a similar relation holds for higher-order iterated integrals

$$\mathbb{X}^{i_1\cdots i_n}_{st} = \int_{s<u_1<\cdots<u_n<t} \mathrm{d}x^{i_n}_{u_n}\mathrm{d}x^{i_{n-1}}_{u_{n-1}}\cdots\mathrm{d}x^{i_1}_{u_1}.$$

Considering the family of linear maps on H given by

$$\langle X_{st}, i_1\cdots i_n\rangle = \mathbb{X}^{i_1\cdots i_n}_{st}$$

we see that this can be rephrased as requiring that X_{st} be multiplicative with respect to the shuffle product. This is the context of Lyons' original theory and essentially it dealt only with this case. In 2010, M. Gubinelli further extended the theory to the non-geometric case in what he called *branched rough paths*. His approach is based on E. Hairer and G. Wanner's observation that solutions to the classical ODE $\dot{y} = f(y)$ can be expanded as a Butcher series

$$y_t = \xi + \sum_{\tau \in \mathfrak{T}} \frac{1}{\tau!} f_\tau(y_t) t^{|\tau|}$$

where the sum is over *non-planar* trees [16, 54].

In the branched case, the shuffle algebra is replaced by the Butcher–Connes–Kreimer Hopf algebra $\mathcal{H}$ over non-planar rooted trees decorated by the alphabet A. This is the commutative polynomial algebra over the space $\mathfrak{T}$ of decorated trees where the coproduct $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ can be represented by means of admissible cuts (see Section 2.5 for further details). As before, the set of linear maps $X \in \mathcal{H}^*$ which are multiplicative form a group, which we denote by $\mathcal{G}$.

Definition 1.2.6. *A branched γ-rough path is a path $X : [0,1]^2 \to \mathcal{G}$ such that $X_{st} = X_{su} \star X_{ut}$ and*

$$|\langle X_{st}, \tau \rangle| \lesssim |t - s|^{\gamma|\tau|}.$$

In this definition, $|\tau|$ denotes the number of nodes in the tree τ. We can think of a branched rough path as a family of paths indexed by forests, satisfying additional regularity and algebraic relations.

The connection between these two settings, namely geometric and branched rough paths, was extensively explored by M. Hairer and D. Kelly [57]. In a subsequent work, Y. Bruned, C. Curry and K. Ebrahimi-Fard examined the algebraic aspects of this relation [14]. The main contribution of Hairer and Kelly's article is to give an application ψ (latter dubbed the Hairer–Kelly map) converting branched rough paths to geometric rough paths. They also provide another map, which they denote by φ_g, converting geometric rough paths to branched rough paths. As a by-product of this conversion they obtain Itô–Stratonovich conversion formulas which generalise the classical formula converting Itô integrals into Stratonovich integrals to a more general class of driving paths. However, they remark that their construction is far from optimal in the general case since, even in the Brownian setting, the formula obtained contains redundant terms that have to be collected afterwards. In [14] this conversion is addressed in the context of semimartingale driving noises, interpreting the Hairer–Kelly map as the arborification of the Hoffman exponential.

We take up on this problem by constructing a translation map, similar to the Hairer–Kelly map, allowing to rewrite branched rough paths as a more general class of geometric rough paths, called *anisotropic geometric rough paths*, introduced in Section 5.5. Roughly, this new setting allows us to identify the redundant data constructed by the original Hairer–Kelly map as components having high regularity, thus permitting a finer control over what objects have to be actually constructed and what objects are already fixed by the other levels.

The idea of iteratively constructing the values of $\langle X, \tau \rangle$ for trees of increasing size by means of the finite second order increment δ already appears in Gubinelli's work on branched rough paths. He uses this technique in [51], together with his *Sewing map* (Theorem 3.1.3), to prove an extension theorem stating that branched rough paths depend effectively only on a finite number of trees. In other words, once one has the values of X on trees with at most $N := \lfloor \gamma^{-1} \rfloor$ nodes, the rest of the values are uniquely defined.

One of the main results of this chapter is the following:

Theorem 1.2.7. *Let $(x^a : a \in A)$ be a collection of real-valued γ-Hölder paths. There is a geometric γ-rough path X such that $\langle X_{st}, a\rangle = x^a_t - x^a_s$ for all $a \in A$.*

The proof consists of an iterative construction of a geometric rough path over the given Hölder paths, thus improving Theorem 1.2.4. In order to achieve our construction we have to keep a balance between the analytical and algebraic constraints imposed by the definition of a geometric rough path: on one hand, the algebraic conditions demand that at each step Chen's rule and the multiplicativity of the rough path are to be kept, and on the other hand the path so obtained should be γ-Hölder continuous with respect to a suitable metric, see Definition 5.3.1. Our approach heavily relies on an explicit form of the Baker–Campbell–Hausdorff formula by Reutenauer [103] (formula (5.9)), as well as on analytic techniques akin to Gubinelli's Sewing Lemma [50], exploiting the fact that the group of characters over the shuffle Hopf algebra is in fact a Lie group whose topology may be metrized in various ways.

In fact, our technique allows us to generalise Lyons and Victoir's construction to different cases and so we obtain

Theorem 1.2.8. *Let $(x^a : a \in A)$ be a collection of real-valued γ-Hölder paths. There is a branched γ-rough path X such that $\langle X_{st}, \bullet_a\rangle = x^a_t - x^a_s$ for all $a \in A$.*

Theorem 1.2.9. *Let $(x^a : a \in A)$ be a collection of real-valued paths and let $(\gamma_a : a \in A)$ be real numbers with $0 < \gamma_a < 1$, such that x^a is γ_a-Hölder. There is an anisotropic geometric rough path X such that $\langle X_{st}, a\rangle = x^a_t - x^a_s$ for all $a \in A$.*

In physical applications, it is interesting to understand the limiting behaviour of regularised (or discretised) solutions to eq. (1.6). Moreover, there are some reasonable approximations for which the corresponding solutions do not converge [85]. In this case, one has to perform some kind of modifications to the involved quantities in order to obtain a meaningful limiting object. For this reason we explore possible modifications to rough paths both in the geometric and in the branched case. In the branched case we prove that the abelian group $\mathscr{C}^\gamma$ of collections of functions $(g^\tau : \tau \in \mathfrak{T})$ such that g^τ is $\gamma|\tau|$-Hölder acts transitively over the space of branched rough paths.

Theorem 1.2.10. *The group $\mathscr{C}^\gamma$ acts transitively on branched rough paths. Given $g \in \mathscr{C}^\gamma$ and a branched rough path there is another branched rough path of the same regularity gX such that*

$g'(gX) = (g + g')X$ for all $g' \in \mathscr{C}^\gamma$. Moreover, if X and X' are any two branched rough paths of the same regularity, then there exists $g \in \mathscr{C}^\gamma$ such that $X' = gX$.

Our method works by iteratively constructing each level with the help of a general extension theorem (see Theorem 5.3.4). The main difficulty for the proof of Theorem 1.2.10 is that this successive construction of level has to respect not only Chen's rule and the analytical constraints imposed by the definition of rough path, but also the modification already done to the previous levels. In particular, the way in which $\mathscr{C}^\gamma$ acts on rough paths is by adding the increment of a $\gamma|\tau|$-Hölder function to each "component" $\langle X, \tau\rangle$. To see why this should be the case consider a $\gamma \in [\frac{1}{3}, \frac{1}{2}[$ so that $N = 2$, and assume that the first levels $\langle X_{st}, \bullet_a\rangle$ and $\langle X'_{st}, \bullet_a\rangle$ are equal and fixed. Chen's rule then implies that if we set $F \coloneqq \langle X_{st}, \substack{\bullet_b \\ | \\ \bullet_a}\rangle$ and define F' accordingly then

$$\delta F_{sut} = \langle X_{su}, \bullet_b\rangle\langle X_{ut}, \bullet_a\rangle = \langle X'_{su}, \bullet_b\rangle\langle X'_{ut}, \bullet_a\rangle = \delta F'_{sut}$$

so that there exists a 2γ-Hölder function g such that

$$\langle X'_{st}, \substack{\bullet_b \\ | \\ \bullet_a}\rangle = \langle X_{st}, \substack{\bullet_b \\ | \\ \bullet_a}\rangle + g_t - g_s.$$

Modulo some technical considerations this is the core of the argument and so our construction is fully explicit. A similar kind of modification was considered by Bruned, Chevyrev, Friz and Preiß in [13], but in the case where the modifications are given by constant multiples of an extra path X^0 which they assume to be regular enough so the translation[1] respects the definition of a rough path. The advantage is that they can describe the result of this translation as the action of a left comodule over the Butcher–Connes–Kreimer Hopf algebra by using a version of the extraction-contraction coproduct investigated in [17] and the cointeraction property it satisfies with respect to the usual Connes–Kreimer coproduct. For the moment we do not have a similar description of our modification in such an algebraic setting, but Theorem 1.2.10 is general enough to allow for modifications by arbitrary Hölder functions. Also, for the same reason, the "renormalisation group" $\mathscr{C}^\gamma$ we obtain is infinite-dimensional as opposed to the one found in [13]. Another instance of this modification theorem and its relation to Rough Differential Equations has been studied in [14] by Bruned, Curry and Ebrahimi-Fard, where they assume that the underlying Hopf algebra has a quasi-shuffle structure. This is supposed to represent the fact that the base paths are semi-martingales so that the Itô rule provides a deformed product where there's an extra term coming from the quadratic variation, hence the quasi-shuffle structure.

[1]In their terminology

1.3 Directed Polymers

One of the main motivations for this work was to study a coupled system of singular stochastic PDEs arising from as a scaling limit of a physical model called the *multilayer semi-discrete directed polymer* introduced by I. Corwin and A. Hammond [25]. This model, which we explain below, was introduced in order to understand a continuous model introduced by N. O'Connell and J. Warren known as the *multilayer Stochastic Heat Equation* (multilayer SHE for short) a few years earlier [98]. The multilayer SHE, for a fixed $n \in \mathbb{N}$ and $\beta > 0$, is defined by its chaos expansion

$$\mathscr{Z}_{\beta}^{(n)}(s,x;t,y) = p(s,x;t,y)^n \left(1 + \sum_{k=1}^{\infty} \int_{\Delta_k(s,t)} \int_{\mathbb{R}^k} R_k^{(n)}(\boldsymbol{u},\boldsymbol{z};s,x,t,y)\, \xi^{\otimes k}(\mathrm{d}\boldsymbol{u},\mathrm{d}z)\right)$$

where $p(s,x;t,y)$ is the standard heat kernel, $R_k^{(n)}$ is the correlation kernel of a collection of n non-intersecting Brownian motions and ξ is a Gaussian space-time white noise on $\mathbb{R}_+ \times \mathbb{R}$. The precise sense in which this iterated stochastic integral is defined is rigorously treated in Section 6.2. The random field for the case $n = 1$ first appeared in [3] as a scaling limit of the partition function of the (discrete) directed polymer model under intermediate disorder. Later on, it would be used in [2] to define a continuum random polymer for which $\mathscr{Z}_{\beta}^{(1)}$ acts as the partition function. This partition function can be interpreted as the mild solution to the classical multiplicative SHE

$$\partial_t \mathscr{Z} = \frac{1}{2}\Delta\mathscr{Z} + \beta\mathscr{Z}\xi \tag{1.14}$$

which corresponds to the Hopf-Cole solution to the KPZ equation. However, to our knowledge no similar description exists for the dynamics of the fields $\mathscr{Z}_{\beta}^{(n)}$ for $n > 1$ although some partials results in this direction have been obtained [83].

This family of stochastic processes has many properties shared by other related models such as the classical SHE, which have already been shown in the recent works [26, 83, 84, 95]. In particular, they bear a re

One way of thinking of $\mathscr{Z}_{\beta}^{(n)}$ is by formally interpreting it as the partition function for a n-layer continuum directed polymer. In the $n = 1$ case the value of the field can be formally seen as giving the average quenched energy of a continuum polymer

$$\mathscr{Z}_{\beta}^{(1)}(s,x;t,y) = p(s,x;t,y)\mathbf{E}_x\left[\exp\left(\beta \int_s^t \xi(u,X_u)\,\mathrm{d}u\right) \middle| X_t = y\right]$$

where X is a Brownian motion under $\mathbf{P}$ (and $\mathbf{E}$ is the associated expectation). Of course the above equality is only formal as the right-hand side is ill-defined as ξ is not a real function but a random distribution. This already makes clear that if one wants to formally define this field as the limit of discrete approximations some form of renormalisation has to be involved. In this interpretation, it is also clear by use of the Feynman–Kac formula that $\mathscr{Z}_\beta^{(1)}$ should satisfy eq. (1.14). It can be shown that $\mathscr{Z}_\beta^{(1)}$ appears as the scaling limit of the partition function for discrete and semi-discrete directed polymers in the intermediate disorder regime [3, 93], with convergence in the sense of chaos expansions in L^2.

A similar formal argument can be made for $n > 1$. In this case, $\mathscr{Z}_\beta^{(n)}$ can be now seen to be formally equal to the energy

$$\mathscr{Z}_\beta^{(n)} = p(s, x; t, y)^n \mathbf{E}_x^n \left[\exp\left(\beta \sum_{i=1}^{n} \int_s^t \xi(u, X_u^i)\, \mathrm{d}u \right) \middle| X_t^1 = \cdots = X_t^n = y \right]$$

where now $(X^1, \ldots, X^n)$ is a standard Brownian motion under $\mathbf{P}^n$. Some additional care has to be taken here as it is now immediately obvious that one can actually make direct sense of the measure $\mathbf{P}$ driving the processes $(X^1, \ldots, X^n)$ so this already adds some extra difficulty. Nonetheless, as in the previous case, one can also view $\mathscr{Z}_\beta^{(n)}$ as the limiting partition function of appropriately defined multilayer polymer models under the same regime [26, 95]. In Chapter 6 we show this for the multilayer semi-discrete directed polymer defined in [25]. See also [95] for an alternative proof.

As stated in a previous paragraph, the dynamics governing the behaviour of $\mathscr{Z}_\beta^{(n)}$ its not fully understood. In their seminal paper [98], O'Connell and Warren study this dynamic under a smoothness assumption for the noise potential ξ. Concretely, they replace ξ by a smooth function φ and derive a set of PDEs satisfied by the now smooth field $z_\beta^{(n)}$. They prove that the ratios $u_\beta^{(n)} = z_\beta^{(n)}/z_\beta^{(n-1)}$ satisfy the nested system of PDEs

$$\delta_t u_\beta^{(n)} = \frac{1}{2}\Delta u_\beta^{(n)} + \beta\left(\varphi + \Delta \log\left(\frac{z_\beta^{(n-1)}}{p^{n-1}} \right) \right) u_\beta^{(n)} \tag{1.15}$$

where $z_\beta^{(0)} \equiv 1$, subject to the initial conditions $u_\beta^{(n)}(s, x; s, y) = \delta(x - y)$. The irregularity of the initial condition also poses additional problems for the analysis of these equations in the stochastic setting. After showing this, O'Connell and Warren state that they were not able to show that if now φ is replaced by a suitable mollification ξ_ε of the space-time white noise, the corresponding solutions $u_{\beta,\varepsilon}^{(n)}$ converge to a meaningful limit in an appropriate topology as $\varepsilon \to 0$. We note that this is to be expected as this is also the case for $n = 1$, where already some renormalisation is needed in order to make sense of eq. (1.14) as the limit of smoothed-out versions as one removes the smoothing. In the single layer setting this result was proven by M. Hairer and C. Labbé by using the theory of Regularity Structures [58]. Specifically, they show that one has to consider instead the modified equation

$$\partial_t \mathscr{Z}_{\beta,\varepsilon} = \frac{1}{2}\Delta \mathscr{Z}_{\beta,\varepsilon} + \beta(\mathscr{Z}_{\beta,\varepsilon} - C_\varepsilon)\xi_\varepsilon$$

where C_ε is a function of ε which diverges as $\varepsilon \to 0$. They show this by essentially using the BPHZ renormalisation scheme built into Regularity Structures which uses Wick renormalisation as a crucial step.

To our knowledge there are no similar results in the case $n > 1$, and this was one of the main motivations for the work done on the previous chapters. We would like to have a similar kind of result expressing the field $\mathscr{Z}_\beta^{(n)}$ as the renormalised limit of the classical solutions to the system eq. (1.15) as in the single layer case. It could be also interesting to treat more general cases by letting, for instance, the distribution of the background noise to be different from a Gaussian distribution. Efforts in this direction directly at the level of the KPZ equation have been made by Hairer and Shen [59]. Again, a crucial step in order to be able to show this is to understand the finer structure of Wick products which play a more prominent role in this setup.

Chapter 2

Algebra

In this chapter we will introduce the main algebraic tools needed for the principal concepts appearing in the core of the work. We will provide suitable references at the beginning of each section.

2.1 Some category theory

Although not strictly necessary, category theory provides a unifying framework for some of the objects that will appear throughout this chapter, which are *universal* in a broad sense. We will just provide the main definitions and facts without going in too much detail, and focusing mainly on examples that will be used further in the text. This part is loosely based on Appendix B of the book by J.-L. Loday and B. Vallette [81].

A *category* is a collection of objects and *arrows* or maps between them, such that for any three objects C, C' and C'' and arrows $f : C \to C'$ and $g : C' \to C''$ there is a notion of composition where $g \circ f : C \to C''$ is again an arrow, and this operation is supposed to be associative. Each object comes with a special arrow $\mathrm{id}_C : C \to C$ such that for any $f : C \to C'$ and $g : C' \to C$ we have $f \circ \mathrm{id}_C = f$ and $\mathrm{id}_C \circ g = g$. Given two objects C and C' we denote by $\mathrm{Hom}_{\mathsf{C}}(C, C')$ the set of arrows between them. Categories will be written in sans-serif upright font and its objects will be denoted in serif italic style. The concept is best illustrated with examples. The category Set has sets as objects and functions as arrows; given a commutative ring R the category $R\mathsf{Mod}$ consists of (left) R modules with R-linear maps as arrows, and in the particular case where $R = k$ is a field this is the category Vect_k of vector spaces over k.

A *functor* between to categories C and D is a mapping sending object C of C to objects D of D and arrows to arrows, such that $F(\mathrm{id}_C) = \mathrm{id}_{F(C)}$, if $f : C \to C'$ is an arrow in C then $F(f) : F(C) \to F(C')$ is an arrow in D, and $F(g \circ f) = F(g) \circ F(f)$. Given a pair of functors

$L : \mathsf{C} \to \mathsf{D}$ and $R : \mathsf{D} \to \mathsf{C}$, we say that L is *left-adjoint* to R and R is *right-adjoint* to L if there is a "natural" isomorphism

$$\mathrm{Hom}_{\mathsf{D}}(L(C), D) \cong \mathrm{Hom}_{\mathsf{C}}(C, R(D))$$

for any two objects C and D in the respective categories. In some categories there is a special functor, called the *forgetful functor* and denoted by $\mathcal{U}$, obtained by "forgetting" some of the structure. A functor left-adjoint to the forgetful functor is called a *free functor*, denoted by $\mathcal{F}$; then, the above equality of hom-sets translates into the following property: for each pair of objects C in C and D in D, and an arrow $f \in \mathrm{Hom}_{\mathsf{C}}(C, \mathcal{U}(D))$, there is an $\hat{f} \in \mathrm{Hom}_{\mathsf{D}}(\mathcal{F}(C), D)$ such that $\mathcal{U}(\hat{f}) = f$.

2.2 Algebras

The next three sections are based of the standard references such as M. E. Sweedler's book [113], or the books [9, 10] by N. Bourbaki. We fix once and for all a field k of characteristic zero. An algebra is triple $\mathcal{A} = (A, m, u)$ where A is a k-vector space, and $m : A \otimes A \to A$ and $u : k \to A$ are k-linear maps such that the diagram

$$\begin{tikzcd}
A \otimes A \otimes A \arrow[rr, "m\otimes\mathrm{id}"] \arrow[dd, "\mathrm{id}\otimes m"'] & & A \otimes A \arrow[dd, "m"'] & \\
& & & k \otimes A \arrow[ul, "u\otimes\mathrm{id}"'] \arrow[dl] \\
A \otimes A \arrow[rr, "m"] & & A & \\
& A \otimes k \arrow[ul, "\mathrm{id}\otimes u"] \arrow[ur] & &
\end{tikzcd}$$

commutes. The main square expresses associativity of the product and the side triangles say that $u(1) \in A$ is the neutral element. Frequently we will not distinguish between the tuple and the vector space and so statements such as *let* $\mathcal{A}$ *be an algebra* will appear. Likewise, we will most usually denote the image of m on two elements $x, y \in \mathcal{A}$ simply by juxtaposition, that is, $m(x \otimes y) = xy$.

Let $\mathcal{A}$ and $\mathcal{A}'$ be two algebras. An *algebra map or homomorphism* is a map $\phi : \mathcal{A} \to \mathcal{A}'$

respecting the products and units, i. e. such that the diagram

$$\begin{tikzcd} \mathcal{A} \otimes \mathcal{A} \arrow[r, "m_{\mathcal{A}}"] \arrow[d, "\phi\otimes\phi"] & \mathcal{A} \arrow[d, "\phi"] & \\ & & k \arrow[ul, "u_{\mathcal{A}}"'] \arrow[dl, "u_{\mathcal{A}'}"] \\ \mathcal{A}' \otimes \mathcal{A}' \arrow[r, "m_{\mathcal{A}'}"] & \mathcal{A}' & \end{tikzcd}$$

commutes.

An *ideal* of an algebra $\mathcal{A}$ is a subspace $\mathfrak{a} \subset \mathcal{A}$ such that $m(\mathfrak{a} \otimes \mathcal{A} + \mathcal{A} \otimes \mathfrak{a}) \subset \mathfrak{a}$, that is, it absorbs left and right multiplication by elements of $\mathcal{A}$. A *subalgebra* of $\mathcal{A}$ is a subspace $\mathcal{S} \subset \mathcal{A}$ such that $m(\mathcal{S} \otimes \mathcal{S}) \subset \mathcal{S}$ and $u(k) \subset \mathcal{S}$. Ideals serve to form quotients as the next proposition shows.

Proposition 2.2.1. *Let $\mathcal{A}$ be an algebra and $\mathfrak{a} \subset \mathcal{A}$ an ideal. The quotient space $\mathcal{A}/\mathfrak{a}$ has an algebra structure and the canonical projection $\pi : \mathcal{A} \longrightarrow \mathcal{A}/\mathfrak{a}$ is an algebra map.*

Proof. It suffices to show that if $x - x' \in \mathfrak{a}$ and $y - y' \in \mathfrak{a}$ then $xy + \mathfrak{a} = x'y' + \mathfrak{a}$. But then there are $a_x, a_y \in \mathfrak{a}$ such that $x = x' + a_x$ and $y = y' + a_y$, so that

$$xy = (x' + a_x)(y' + a_y) = x'y' + x'a_y + y'a_x + a_x a_y = x'y' + a$$

where $a \in \mathfrak{a}$ by hypothesis. Hence, one can define a product on $\mathcal{A}/\mathfrak{a}$ unambiguously by setting $(x + \mathfrak{a})(y + \mathfrak{a}) = xy + \mathfrak{a}$. This also means that if $\pi(x) = x + \mathfrak{a}$ is the canonical projection then $\pi(xy) = xy + \mathfrak{a} = \pi(x)\pi(y)$. Obviously, $\bar{u} = \pi \circ u$ acts as the unit. □

Proposition 2.2.2. *Let $\mathcal{A}$ be an algebra $\mathfrak{a} \subset \mathcal{A}$ be an ideal. Suppose $\mathcal{A}'$ is another algebra and that $f : \mathcal{A} \longrightarrow \mathcal{A}'$ is an algebra homomorphism such that $\mathfrak{a} \subset \ker f$. Then, there exists a unique algebra morphism $\tilde{f} : \mathcal{A}/\mathfrak{a} \longrightarrow \mathcal{A}'$ such that $f = \tilde{f} \circ \pi$.*

Proof. Define $\tilde{f}(x + \mathfrak{a}) = f(x)$ for each $x \in \mathcal{A}$. This is well defined since if $x + \mathfrak{a} = y + \mathfrak{a}$ then there is $a \in \mathfrak{a}$ such that $x = y + a$ and so

$$\begin{aligned} \tilde{f}(x + \mathfrak{a}) &= f(x) \\ &= f(y + a) \\ &= f(y) + f(a) \\ &= f(y) \\ &= \tilde{f}(y + \mathfrak{a}). \end{aligned}$$

Moreover, for all $x, y \in \mathcal{A}$ we have

$$\begin{aligned}\tilde{f}((x+\mathfrak{a})(y+\mathfrak{a})) &= \tilde{f}(xy+\mathfrak{a}) \\ &= f(xy) \\ &= f(x)f(y) \\ &= \tilde{f}(x+\mathfrak{a})\tilde{f}(y+\mathfrak{a}).\end{aligned}$$

□

It is a simple exercise to show that an arbitrary intersection of ideals is again an ideal. Thus, given a subset $S \subset \mathcal{A}$ one can form the intersection of all the ideals containing S, which is the smallest ideal containing S, and will be denoted by $\langle S \rangle$. This ideal admits an *internal* description in terms of the elements of S.

Proposition 2.2.3. *Let $S \subset \mathcal{A}$. Then*

$$\langle S \rangle = \left\{ \sum_{i=1}^{n} \alpha_i a_i s_i b_i : n \in \mathbb{N}, \alpha_i \in k, s_i \in S, a_i, b_i \in \mathcal{A} \right\}. \tag{2.1}$$

Proof. Let I be the set appearing on the right-hand side of eq. (2.1). This set is clearly an ideal containing S, hence $\langle S \rangle \subset I$. □

A *left module* over an algebra $\mathcal{A}$, or left $\mathcal{A}$-module, is a tuple $\mathcal{M} = (M, \rho)$ where M is a k-vector space and $\rho : A \otimes M \to M$ is a linear map such that the diagram

$$\begin{tikzcd} A \otimes A \otimes M \arrow[r, "m\otimes\mathrm{id}"] \arrow[d, "\mathrm{id}\otimes\rho"] & A \otimes M \arrow[d, "\rho"] & \\ & & k \otimes M \arrow[ul, "u\otimes\mathrm{id}"'] \arrow[dl] \\ A \otimes M \arrow[r, "\rho"] & M & \end{tikzcd}$$

commutes. A right comodule is defined similarly.

Let $\mathcal{A}, \mathcal{A}'$ be two algebras. The tensor product $\mathcal{A} \otimes \mathcal{A}'$ also carries an algebra structure with product $m_{\mathcal{A}\otimes\mathcal{A}'} = (m_{\mathcal{A}} \otimes m_{\mathcal{A}'}) \circ \tau_{2,3}$, where $\tau_{2,3} : \mathcal{A} \otimes \mathcal{A}' \otimes \mathcal{A} \otimes \mathcal{A}' \to \mathcal{A} \otimes \mathcal{A} \otimes \mathcal{A}' \otimes \mathcal{A}'$ is the natural isomorphism. Thus, the product on $\mathcal{A} \otimes \mathcal{A}'$ is given as $(x \otimes x')(y \otimes y') = xy \otimes x'y'$. Clearly, the unit is the map $u_{\mathcal{A}\otimes\mathcal{A}'} = u_{\mathcal{A}} \otimes u_{\mathcal{A}'}$.

An algebra $\mathcal{A}$ is said to be *graded* if its underlying vector space can be decomposed as a direct sum

$$\mathscr{A} = \bigoplus_{p=0}^{\infty} \mathscr{A}_p$$

such that $m(\mathscr{A}_p \otimes \mathscr{A}_q) \subset \mathscr{A}_{p+q}$ and $u(k) \subset \mathscr{A}_0$. The elements of $\mathscr{A}_p$ are called *homogeneous* of degree p. Thus, every element in $\mathscr{A}$ can be write as a finite sum of homogeneous elements, and these terms are called its homogeneous components. If u is an isomorphism then $\mathscr{A}$ is said to be *connected*. We denote $\deg(x) = p$ –sometimes $|x|$ if this does not lead to confusion– if and only if $x \in \mathscr{A}_p$.

Proposition 2.2.4. *For every $q \geq 0$ the subspace*

$$\mathfrak{a}_q := \bigoplus_{p=q+1}^{\infty} \mathscr{A}_p$$

is an ideal in $\mathscr{A}$.

Proof. Let $x \in \mathscr{A}$ and $a \in \mathfrak{a}_q$. Since $\mathscr{A}$ is graded we have that for all $p \geq 0$ and $r > q$ the product $x_p a_r \in \mathscr{A}_{p+r} \subset \mathfrak{a}_q$. Therefore, xa and ax belong to $\mathfrak{a}_q$ since, for example,

$$(xa)_p = \sum_{r=0}^{p} x_r a_{p-r} \in \mathfrak{a}_q.$$

□

Corollary 2.2.5. *For all $q \geq 0$ the quotient algebra $\mathscr{A}/\mathfrak{a}_q$ is isomorphic to the algebra*

$$\mathscr{A}_{(q)} := \bigoplus_{p=0}^{q} \mathscr{A}_p$$

with product

$$x \cdot_q y = \begin{cases} xy & \deg(x) + \deg(y) \leq q, \\ 0 & \deg(x) + \deg(y) > q \end{cases}$$

2.2.1 The free algebra

Let V be a vector space. The p-fold tensor product of V with itself is the vector space $V^{\otimes p}$ with the convention that $V^{\otimes 0} \cong k$. The vector space

$$T(V) = \bigoplus_{p=0}^{\infty} V^{\otimes p} \tag{2.2}$$

carries a natural product structure given by the natural isomorphism $V^{\otimes p} \otimes V^{\otimes q} \cong V^{\otimes(p+q)}$. Denoting by $\mathbf{1}$ the unique vector spanning $V^{\otimes 0}$, the natural isomorphism $k \otimes W \cong W \otimes k \cong W$ –where W is an arbitrary real vector space– gives that $\mathbf{1}$ acts as the unit for this product. The algebra $(T(V), \otimes, \mathbf{1})$ is known in the literature as the *tensor algebra over* V, and it corresponds to the *free associative algebra* over V [81]. Observe that $T(V)$ is in fact a graded connected algebra, where the tensor products $V^{\otimes p}$ are the homogeneous components.

Remark 2.2.6. In order to avoid confusion, we reserve the tensor symbol $\otimes$ to denote the "internal" tensor product on $T(V)$, i.e. the one giving the algebra structure. When talking about the "external" tensor product of $T(V)$ with itself we will use the decorated tensor product $\tilde{\otimes}$. So for example the product on $T(V)$ can be written as $m(a \tilde{\otimes} b) = a \otimes b$. Likewise, the vectors $\mathbf{1} \tilde{\otimes} a$ and $\mathbf{1} \otimes a = a$ belong to $T(V) \tilde{\otimes} T(V)$ and $T(V)$, respectively. △

There is another equivalent construction of the free algebra. Let A be a set, considered as an alphabet, and let A^* denote the *free monoid* over A; the unit in A^* will be denoted by $\mathbf{1}$. As a set, A^* consists of all the finite sequences $w = (a_1, \ldots, a_p)$ of elements of A –called words–, and for such a word denote by $\ell(w) = p$ its *length*. By definition, we have an associative law $A^* \times A^* \longrightarrow A^*$, $(a, b) \mapsto a.b$ such that $\mathbf{1}.w = w.\mathbf{1} = w$ for all $w \in A^*$, and if $w = (a_1, \ldots, a_p), w' = (a'_1, \ldots, a_q) \in A^*$ then $w.w' = (a_1, \ldots, a_p, a'_1, \ldots, a'_q)$. Evidently $\ell(w.w') = \ell(w) + \ell(w')$. Every map $f : A \longrightarrow M$ where M is some monoid extends uniquely to a monoid morphism $f : A^* \longrightarrow M$ by $f(a_1 \ldots a_p) = f(a_1) \ldots f(a_p)$.

Denote by kA^* the free vector space over A^*. The bilinear extension of the monoid law on A^* induces a product on kA^*, where $\mathbf{1}$ remains as the unit, as is easy to check. Every map $f : A \longrightarrow \mathscr{A}$ where $\mathscr{A}$ in any algebra extends uniquely to an algebra morphism $f : kA^* \longrightarrow \mathscr{A}$, hence we can identify kA^* as the *free algebra over the set* A, and we denote it by $k\langle A \rangle$. Note that in particular all words can be written as a finite product $w = a_1.a_2 \ldots a_p$ of letters from A. We denote by A^p the collection of all words $w \in A^*$ with $\ell(w) = p$. Thus, $k\langle A \rangle$ is graded by word-length and the subspaces kA^p are the homogeneous components.

Remark 2.2.7. If V is a vector space of dimension the cardinality of A, there is a bijection between a basis for V and A. Thus, we can identify V as the free vector space over A and then the monoid operation on A^* will coincide with the tensor product on $T(kA)$. △

The algebra $k\langle A \rangle$ is sometimes called the *non-commutative polynomial algebra* and each letter from A plays the role of an indeterminate. Observe that there is no restriction on the cardinality of A. In the case where A is finite, $k\langle A \rangle$ is easily seen to be isomorphic to the non-commutative polynomial ring $k\langle x_a : a \in A \rangle$. For this reason, words in A^* –that is, the basis elements of the underlying vector space– are called monomials.

The free commutative algebra

There is a natural action of the symmetric group S_p on the vector space $V^{\otimes p}$, given by

$$\sigma.(v_1 \otimes \cdots \otimes v_p) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(p)}.$$

If we let $S^p(V) = V^{\otimes p}/S_p$ denote the collection of all orbits of elements of $V^{\otimes p}$ under the action of S_p, the vector space

$$S(V) = \bigoplus_{p=0}^{\infty} S^p(V)$$

becomes a commutative algebra with unit $\mathbf{1} \in S^0(V)$. This algebra is called the *symmetric algebra* over V.

Remark 2.2.8. An equivalent construction can be over an arbitrary set A by considering V to be the vector space kA spanned by A. The action of S_p is by permuting the letters in a word. △

We can obtain $S(V)$ as a quotient of $T(V)$ by a suitable ideal. Consider the set

$$S = \{u \otimes v - v \otimes u : u, v \in V\}$$

and let $\mathfrak{a} = \langle S \rangle$ be the ideal generated by this set. Then $S(V)$ is isomorphic as an algebra to the quotient $T(V)/\mathfrak{a}$, because every element in S_p can be decomposed into a finite number of transpositions.

The commutative algebra $S(V)$ enjoys the following universal property: if $f : V \to \mathcal{A}$ is any linear map from V to a commutative unital algebra $\mathcal{A}$, it uniquely extends to an algebra morphism $f : S(V) \to \mathcal{A}$. This is true in view of Proposition 2.2.2. Indeed, if $f : V \to \mathcal{A}$ is a linear map, it extends uniquely as an algebra homomorphism $f : T(V) \to \mathcal{A}$. Since $\mathcal{A}$ is commutative this map is symmetric, that is, $f(u \otimes v) = f(v \otimes u)$. Hence $\mathfrak{a} \subset \ker f$ and so Proposition 2.2.2 applies. In particular, if A is a finite alphabet, we can identify $k[A]$ with the usual (commutative) polynomial ring $k[x_a : a \in A]$.

A very special case, that will be of interest for latter developments is when $A = \{a\}$ –or V is 1-dimensional. In this case both $k\langle A \rangle$ and $k[A]$ are isomorphic to the polynomial ring in one variable $k[x]$. A basis for this space is given by the monomials $\{\mathbf{1}, x, x^2, \dots\}$ and the product is the usual product of powers $x^n \cdot x^m = x^{n+m}$.

2.2.2 The shuffle algebra

It is possible to endow the space $T(V)$ defined above with another commutative algebra structure which differs from the free commutative algebra constructed above. For the description of this product it is best to use the presentation of the free algebra with words. Construct a map bilinear map $\sqcup\!\sqcup : kA^* \otimes kA^* \to kA^*$ inductively by setting $\mathbf{1} \sqcup\!\sqcup a = a \sqcup\!\sqcup \mathbf{1} = a$ for all $a \in A$ and, if $v, w \in A^*$ are words and $a, b \in A$ are any two letters, then

$$av \sqcup\!\sqcup bw = a(v \sqcup\!\sqcup bw) + b(av \sqcup\!\sqcup w).$$

For example, we may compute

$$a \sqcup\!\sqcup b = ab + ba, \quad ab \sqcup\!\sqcup cd = abcd + acbd + cabd + cadb + cdab.$$

It can be shown that $(kA^*, \sqcup\!\sqcup, \mathbf{1})$ is a commutative unital algebra, which is also graded and connected. The homogeneous component of degree $p \geq 0$ is spanned by A^p, the set of words of length exactly p. In the case the set A is totally ordered, it can also be shown that the shuffle algebra is free over a special set of words $L(A) \subset A^*$, called *Lyndon words*. This means that, as an algebra, the shuffle algebra is isomorphic to the polynomial ring $k[x_w : w \in L(A)]$.

2.2.3 Universal enveloping algebras

Let $\mathfrak{L}$ be a Lie algebra with bracket $[\cdot,\cdot]$. Recall that this means that the map $[\cdot,\cdot] : \mathfrak{L} \times \mathfrak{L} \to \mathfrak{L}$ is bilinear, and satisfies $[x, x] = 0$ and the Jacobi identity

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \tag{2.3}$$

for all $x, y, z \in \mathfrak{L}$. If $\mathfrak{L}$ and $\mathfrak{L}'$ are two Lie algebras, a linear function $\phi : \mathfrak{L} \to \mathfrak{L}'$ satisfying $\phi([x, y]) = [\phi(x), \phi(y)]$ is said to be a *Lie algebra homomorphism or map*. There is a canonical way to associate, to any associative algebra $\mathscr{A}$ a Lie algebra $\mathfrak{L}_{\mathscr{A}}$ by setting $[x, y] = xy - yx$.

A *universal enveloping algebra* for $\mathfrak{L}$ is an algebra $\mathscr{A}$ such that for any other algebra $\mathscr{A}'$, every Lie algebra map $\phi : \mathfrak{L} \to \mathfrak{L}_{\mathscr{A}'}$ uniquely extends to an algebra map $\phi : \mathscr{A} \to \mathscr{A}'$. As with all objects satisfying universal properties, universal enveloping algebras are unique up to algebra isomorphism.

Given an arbitrary Lie algebra $\mathfrak{L}$, there is a canonical way of constructing its universal enveloping algebra. Consider the free algebra $T(\mathfrak{L})$ over the underlying vector space of $\mathfrak{L}$ as constructed on Section 2.2.1 and let $\mathfrak{a}$ be the ideal generated by the set

$$S = \{x \otimes y - y \otimes x - [x, y] : x, y \in \mathfrak{L}\}.$$

Proposition 2.2.9. *The quotient algebra $U(\mathfrak{L}) = T(\mathfrak{L})/\mathfrak{a}$ is a universal enveloping algebra for $\mathfrak{L}$.*

Proof. Let $\mathscr{A}'$ be an algebra and $f : \mathfrak{L} \to \mathfrak{L}_{\mathscr{A}'}$ a Lie algebra morphism. In particular, f is linear so it admits a unique extension $f : T(\mathfrak{L}) \to \mathfrak{L}_{\mathscr{A}'}$ as an algebra homomorphism. Now, by Proposition 2.2.3 every $a \in \mathfrak{a}$ is a linear combination of elements of the form

$$v_1 \otimes \cdots \otimes v_i \otimes (x \otimes y - y \otimes x - [x, y]) \otimes v_{i+1} \otimes \cdots \otimes v_p$$

where $v_1, \ldots, v_p, x, y \in \mathfrak{L}$ and $p \geq 1$. But the image under f of an element of this form is equal to

$$f(a) = f(v_1) \cdots f(v_i)(f(x)f(y) - f(y)f(x) - f([x, y]))f(v_{i+1}) \cdots f(v_p) = 0$$

hence $\mathfrak{a} \subset \ker f$. Thus, in view of Proposition 2.2.2 there exists a unique algebra map $\tilde{f} : U(\mathfrak{L}) \to \mathfrak{L}_{\mathscr{A}'}$ such that $f = \tilde{f} \circ \pi$. ◻

2.3 Coalgebras

A *coalgebra* is a triple $\mathscr{C} = (C, \Delta, \varepsilon)$ where C is a k-vector space, and $\Delta : C \to C \otimes C$ and $\varepsilon : C \to k$ are k-linear maps such that the diagram

$$\begin{tikzcd}
 & & k \otimes C & \\
 & C \arrow[ur] \arrow[rr, "\Delta"] \arrow[dl] \arrow[dd, "\Delta"] & & C \otimes C \arrow[ul, "\varepsilon\otimes\mathrm{id}"'] \arrow[dd, "\Delta\otimes\mathrm{id}"] \\
C \otimes k & & & \\
 & C \otimes C \arrow[ul, "\mathrm{id}\otimes\varepsilon"] \arrow[rr, "\mathrm{id}\otimes\Delta"] & & C \otimes C \otimes C
\end{tikzcd}$$

commutes. Observe that this is the same diagram as before but with "arrows reversed". We will use Sweedler's notation to denote the image of Δ as follows. For $x \in C$ we write

$$\Delta x = \sum_{(x)} x_1 \otimes x_2,$$

or in a more compact way, $\Delta x = x_1 \otimes x_2$. Sometimes we will also drop the limit on the sum. In this notation, the coassociativity property $(\mathrm{id} \otimes \Delta)\Delta = (\Delta \otimes \mathrm{id})\Delta$ reads

$$\sum x_1 \otimes (x_2)_1 \otimes (x_2)_2 = \sum (x_1)_1 \otimes (x_1)_2 \otimes x_2,$$

for all $x \in C$, hence there is no ambiguity in simply writing

$$(\mathrm{id} \otimes \Delta)\Delta x = (\Delta \otimes \mathrm{id})\Delta x = \sum x_1 \otimes x_2 \otimes x_3.$$

Likewise, the counit satisfies

$$x = \sum \langle \varepsilon, x_1 \rangle x_2 = \sum \langle \varepsilon, x_2 \rangle x_1 .$$

Let $\mathscr{C}$ and $\mathscr{C}'$ be two coalgberas. A *coalgebra homomorphism* is a map $\phi : \mathscr{C} \to \mathscr{C}'$ such that the diagram

$$\begin{tikzcd}
 & \mathscr{C} \arrow[r, "\Delta"] \arrow[d, "\phi"] & \mathscr{C} \otimes \mathscr{C} \arrow[d, "\phi\otimes\phi"] \\
k \arrow[ur, "\varepsilon"] \arrow[dr, "\varepsilon'"'] & & \\
 & \mathscr{C}' \arrow[r, "\Delta'"'] & \mathscr{C}' \otimes \mathscr{C}'
\end{tikzcd}$$

A *coideal* in a coalgebra $\mathscr{C}$ is a subspace such that $\Delta\mathfrak{c} \subset \mathscr{C} \otimes \mathfrak{c} + \mathfrak{c} \otimes \mathscr{C}$. A *subcoalgebra* of $\mathscr{C}$ is a suspace $\mathbb{S}$ such that $\Delta\mathbb{S} \subset \mathbb{S} \otimes \mathbb{S}$ and $\mathbb{S} \subset \ker \varepsilon$. Coideals serve, as ideals in an algebra, to form quotients.

Proposition 2.3.1. *Let $\mathscr{C}$ be a coalgebra and $\mathfrak{c} \subset \mathscr{C}$ a coideal. The quotient space $\mathscr{C}/\mathfrak{c}$ has a coalgebra structure and the canonical projection is a coalgebra map.*

Proof. Let $\pi : \mathscr{C} \to \mathscr{C}/\mathfrak{c}$ be the canonical projection, and define a coproduct $\Delta_q : \mathscr{C}/\mathfrak{c} \to \mathscr{C}/\mathfrak{c} \otimes \mathscr{C}/\mathfrak{c}$ by $\Delta_q \pi(x) = (\pi \otimes \pi)\Delta x$. To check that Δ_q is well defined it suffices to show that if $\pi(x) = \pi(x')$ then $\Delta_q \pi(x) = \Delta_q \pi(x')$. But in that case there is a $c \in \mathfrak{c}$ such that $x = x' + c$ and then $\Delta x = \Delta x' + \Delta c$, whence

$$\Delta_q \pi(x) = \Delta_q \pi(x') + (\pi \otimes \pi)\Delta c = \Delta_q \pi(x')$$

since $\mathscr{C} \otimes \mathfrak{c} + \mathfrak{c} \otimes \mathscr{C} \subset \ker(\pi \otimes \pi)$.

Now define $\bar{\varepsilon} : \mathscr{C}/\mathfrak{c} \to k$ by $\langle \bar{\varepsilon}, \pi(x) \rangle = \langle \varepsilon, x \rangle$. This is well-defined and

$$(\bar{\varepsilon} \otimes \mathrm{id})\Delta_q \pi(x) = (\varepsilon \otimes \pi)\Delta x = \pi(x).$$

The other identity can be checked in a similar way. □

A *left comodule* over a coalgebra $\mathscr{C}$, or left $\mathscr{C}$-comodule, is a tuple $\mathbb{M} = (M, \rho)$ where M is a vector space and $\rho : M \to C \otimes M$ is such that the diagram

$$\begin{tikzcd} & M \arrow[r, "\rho"] \arrow[d, "\rho"] & C \otimes M \arrow[d, "\mathrm{id}\otimes\rho"] \\ k \otimes M \arrow[ur] \arrow[dr, "\varepsilon\otimes\mathrm{id}"'] & & \\ & C \otimes M \arrow[r, "\Delta\otimes\mathrm{id}"'] & C \otimes C \otimes M \end{tikzcd}$$

commutes. A right comodule is defined in a similar fashion. Using Sweedler's notation, the image of ρ is sometimes written

$$\rho(x) = \sum x_1 \otimes x_0$$

so the diagram above imposes the identities

$$\sum x_1 \otimes (x_0)_1 \otimes (x_0)_0 = \sum (x_1)_1 \otimes (x_1)_2 \otimes x_0$$

and

$$\sum \langle \varepsilon, x_1 \rangle x_0 = x.$$

If $\mathscr{C}, \mathscr{C}'$ are two coalgebras, their tensor product $\mathscr{C} \otimes \mathscr{C}'$ is also a coalgebra with coproduct $\Delta_{\mathscr{C}\otimes\mathscr{C}'} = \tau_{2,3} \circ (\Delta_{\mathscr{C}} \otimes \Delta_{\mathscr{C}'})$ and counit $\varepsilon_{\mathscr{C}\otimes\mathscr{C}'} = \varepsilon_{\mathscr{C}} \otimes \varepsilon_{\mathscr{C}'}$, where $\tau_{2,3}$ is the flip map $\tau_{2,3}(a \otimes b \otimes c \otimes d) = a \otimes c \otimes b \otimes d$.

A coalgebra $\mathscr{C}$ is said to be *graded* if its underlying vector space admits a decomposition

$$\mathscr{C} = \bigoplus_{p=0}^{\infty} \mathscr{C}_p$$

such that $\ker \varepsilon \subset \mathscr{C}_1 \oplus \mathscr{C}_2 \oplus \cdots$ and

$$\Delta \mathscr{C}_p \subset \bigoplus_{q+r=p} \mathscr{C}_q \otimes \mathscr{C}_r.$$

As before, a graded coalgebra $\mathscr{C}$ is said to be *connected* if $\mathscr{C}_0$ is one-dimensional.

It is an easy exercise to prove that

Proposition 2.3.2. *Fix $q \geq 1$. The subspace*

$$\mathfrak{c}_q := \bigoplus_{p=0}^{q} \mathscr{C}_p$$

is a coideal in $\mathscr{C}$.

2.3.1 The cofree coalgebra

Similarly to the case of an algebra, there is an explicit construction of the cofree coalgebra over a vector space V. Let V be a vector space. As before, let

$$T^c(V) = \bigoplus_{p=0}^{\infty} V^{\otimes p}.$$

It is possible to define a coproduct $\Delta : T^c(V) \to T^c(V) \tilde{\otimes} T^c(V)$ in the following way: let $\Delta \mathbf{1} = \mathbf{1} \tilde{\otimes} \mathbf{1}$ and for an elementary tensor $x = v_1 \otimes \cdots \otimes v_p \in V^{\otimes p}$ set

$$\Delta x = \mathbf{1} \tilde{\otimes} x + x \tilde{\otimes} \mathbf{1} + \sum_{j=1}^{p-1} (v_1 \otimes \cdots \otimes v_j) \tilde{\otimes} (v_{j+1} \otimes \cdots \otimes v_p).$$

The counit is the map such that $\langle \varepsilon, \mathbf{1} \rangle = 1$ and $\langle \varepsilon, v \rangle = 0$ else. The coalgebra $(T^c(V), \Delta, \varepsilon)$ has the following property: if $\mathscr{C}$ is any coalgebra, each map $f : \mathscr{C} \to V$ extends uniquely to a coalgebra map $f : \mathscr{C} \to T^c(V)$. Note that $T^c(V)$ is a graded connected coalgebra where the tensor products $V^{\otimes p}$ play the role of the homogeneous components.

There is another, equivalent construction of the cofree coalgebra. Let A be a set, regarded as an alphabet, and let A^* be the free monoid over A; the unit in A^* is denoted by $\mathbf{1}$. There is a coproduct $\Delta : kA^* \to kA^* \otimes kA^*$ such that $\Delta \mathbf{1} = \mathbf{1} \otimes \mathbf{1}$ and if $w = a_1 \dots a_p$ is a non-empty word then

$$\Delta w = \mathbf{1} \otimes w + w \otimes \mathbf{1} + \sum_{j=1}^{p} a_1 \dots a_j \otimes a_{j+1} \dots a_p.$$

As before, both constructions are trivially isomorphic.

2.3.2 The convolution algebra

Let $\mathscr{C}$ be a coalgebra and $\mathscr{A}$ be an algebra. It is possible to obtain an algebra structure on the space of k-linear maps $\mathrm{Hom}_k(\mathscr{C}, \mathscr{A})$ by transposing the coalgebra structure of $\mathscr{C}$. Specifically, given $\phi, \psi \in \mathrm{Hom}_k(\mathscr{C}, \mathscr{A})$, their convolution product is the map $\phi \star \psi \in \mathrm{Hom}_k(\mathscr{C}, \mathscr{A})$ as

$$(\phi \star \psi)(x) = m(\phi \otimes \psi)\Delta x. \tag{2.4}$$

In Sweedler's notation this reads

$$(\phi \star \psi)(x) = \sum \phi(x_1)\psi(x_2).$$

Associativity of the convolution product follows from both the associativity of m and the coassociativity of Δ since

$$\begin{aligned}
[\phi \star (\psi \star \theta)](x) &= m(\phi \otimes \psi \star \theta)\Delta x \\
&= m(\mathrm{id} \otimes m)(\phi \otimes \psi \otimes \theta)(\mathrm{id} \otimes \Delta)\Delta x \\
&= m(m \otimes \mathrm{id})(\phi \otimes \psi \otimes \theta)(\Delta \otimes \mathrm{id})\Delta x \\
&= [(\phi \star \psi) \star \theta](x)
\end{aligned}$$

The map $\eta = u \circ \varepsilon$ acts as the unit, since

$$\begin{aligned}
(\eta \star \psi)(x) &= m(u \circ \varepsilon \otimes \psi)\Delta x \\
&= \sum \langle \varepsilon, x_1 \rangle u(1)\psi(x_2) \\
&= \sum \langle \varepsilon, x_1 \rangle \psi(x_2) \\
&= \psi\Big(\sum \langle \varepsilon, x_1 \rangle x_2\Big) \\
&= \psi(x).
\end{aligned}$$

The other identity is proven similarly.

This shows that in particular the linear dual $\mathscr{C}^* = \mathrm{Hom}_k(\mathscr{C}, k)$ of $\mathscr{C}$ possesses an associative algebra structure. In this case, the counit of $\mathscr{C}$ plays the role of the unit in $\mathscr{C}^*$. The associative algebra $(\mathscr{C}^*, \star, \varepsilon)$ is known as the *dual algebra* of the coalgebra $\mathscr{C}$.

There is a nice duality between subcoalgebras of $(\mathscr{C}, \Delta)$ and ideals of $(\mathscr{C}^*, \star)$. Recall that for a subset $J \subset \mathscr{C}$, its *anihilator* is the subspace

$$J^\perp := \{\psi \in \mathscr{C}^* : \langle \psi, x \rangle = 0, \forall x \in J\}.$$

Proposition 2.3.3. *Let $\mathcal{S}$ be a subspace of $\mathscr{C}$. Then*

1. *$\mathcal{S}$ is a subcoalgebra if and only if $\mathcal{S}^\perp$ is an ideal in $\mathscr{C}^*$.*

2. *$\mathcal{S}$ is a coideal if and only if $\mathcal{S}^\perp$ is a subalgebra of $\mathscr{C}^*$.*

2.3.3 The restricted and graded duals

We observe however that it is not always possible to turn the dual space of an arbitrary algebra into a coalgebra by dualising the product. Suppose $(\mathscr{A}, m, \mathbf{1})$ is an algebra and let $\mathscr{A}^* = \mathrm{Hom}_k(\mathscr{A}, k)$ denote the dual space of its underlying vector space. The product is a linear map $m\colon \mathscr{A} \otimes \mathscr{A} \to \mathscr{A}$ hence its transpose $m^*\colon \mathscr{A}^* \to (\mathscr{A} \otimes \mathscr{A})^*$. But in general $(\mathscr{A} \otimes \mathscr{A})^*$ is strictly bigger than $\mathscr{A}^* \otimes \mathscr{A}^*$, unless $\mathscr{A}$ is finite-dimensional, in which case they are isomoprhic.

Suppose there is a subspace $\mathscr{D} \subset \mathscr{A}^*$ such that $m^*(\mathscr{D}) \subset \mathscr{D} \otimes \mathscr{D}$. In this case, the restriction $\Delta = m^*|_{\mathscr{D}}$ defines a coproduct because, since m is associative, one has

$$\begin{aligned}
\langle (\Delta \otimes \mathrm{id})\Delta\psi, x \otimes y \otimes z \rangle &= \langle \Delta\psi, xy \otimes z \rangle \\
&= \langle \psi, (xy)z \rangle \\
&= \langle \psi, x(yz) \rangle \\
&= \langle \Delta\psi, x \otimes yz \rangle \\
&= \langle (\mathrm{id} \otimes \Delta)\Delta\psi, x \otimes y \otimes z \rangle.
\end{aligned}$$

Moreover, the linear map $\varepsilon \colon \mathscr{A}^* \to k$ defined by $\langle \varepsilon, \psi \rangle = \langle \psi, \mathbf{1} \rangle$ satisfies

$$\langle (\varepsilon \otimes \mathrm{id})\Delta\psi, x \rangle = \langle \Delta\psi, \mathbf{1} \otimes x \rangle = \langle \psi, x \rangle$$

so that $(\mathscr{D}, \Delta, \varepsilon)$ is a coalgebra, which is cocommutative if $\mathscr{A}$ is commutative.

One can consider the space $\mathscr{A}^\circ$ consisting of all linear maps $\psi \in \mathscr{A}^*$ such that $\ker\psi$ contains a cofinite ideal, that is, such that there is an ideal $\mathfrak{a} \subset \ker\psi$ with $\dim(\mathscr{A}/\mathfrak{a}) < \infty$. This space satisfies the identity $(\mathscr{A} \otimes \mathscr{A})^\circ \cong \mathscr{A}^\circ \otimes \mathscr{A}^\circ$ and $m^*(\mathscr{A}^\circ) \subset \mathscr{A}^\circ \otimes \mathscr{A}^\circ$. Hence the restriction $\Delta = m^*|_{\mathscr{A}^\circ}$ defines a coproduct, and the map $\langle \varepsilon, \psi \rangle = \langle \psi, \mathbf{1} \rangle$ is a counit so $(\mathscr{A}^\circ, \Delta, \varepsilon)$ becomes a coalgebra.

We obtain an alternate construction of the cofree coalgebra [30].

Proposition 2.3.4. *Let V be a vector space. Then $T(V^*)^\circ$ is the cofree coalgebra over V.*

In the case $\mathscr{A}$ is also graded one can do a little bit better. Suppose $\mathscr{A}$ is a graded algebra with homogeneous components $\mathscr{A}_p$. The *graded dual* of $\mathscr{A}$ is the space

$$\mathscr{A}^{\mathrm{gr}} := \bigoplus_{p=0}^{\infty} \mathscr{A}_p^*,$$

which is a subspace of the full dual $\mathscr{A}^*$. One has $(\mathscr{A} \otimes \mathscr{A})^{\mathrm{gr}} \cong \mathscr{A}^{\mathrm{gr}} \otimes \mathscr{A}^{\mathrm{gr}}$ and $m^*(\mathscr{A}^{\mathrm{gr}}) \subset \mathscr{A}^{\mathrm{gr}} \otimes \mathscr{A}^{\mathrm{gr}}$ and so the restriction $\Delta = m^*|_{\mathscr{A}^{\mathrm{gr}}}$ is a coproduct, hence $(\mathscr{A}^{\mathrm{gr}}, \Delta, \varepsilon)$ is also a coalgebra.

There is a relation between the graded a restricted duals.

Proposition 2.3.5. *Let $\mathscr{A}$ be a* locally finite *graded algebra, i.e. such that $\mathscr{A}_p$ is finite-dimensional for each $p \geq 0$. Then $\mathscr{A}^{\mathrm{gr}} \subset \mathscr{A}^\circ$.*

Proof. Suppose $\psi \in \mathscr{A}^{\mathrm{gr}}$. Then, ψ decomposes as a sum

$$\psi = \sum_{p=0}^{\infty} \alpha_p \psi_p$$

where $\psi_p \in \mathcal{A}_p^*$ and only a finite number of coefficients α_p is non-zero. Moreover, for all $x \in \mathcal{A}$ one has

$$\langle \psi, x \rangle = \sum_{p=0}^{\infty} \alpha_p \langle \psi_p, x_p \rangle$$

where $x_p \in \mathcal{A}_p$ are the homogeneous components of x.

Let $q \in \mathbb{N}$ be such that $\psi_p = 0$ for all $p \geq q$. Then, if $x \in \mathcal{A}_q \oplus \mathcal{A}_{q+1} \oplus \cdots$ we have $\langle \psi, x \rangle = 0$, that is, $x \in \ker \psi$. But the subspace

$$\mathfrak{a} = \bigoplus_{p=q}^{\infty} \mathcal{A}_p$$

is an ideal (Proposition 2.2.4) in $\mathcal{A}$ contained in $\ker \psi$ such that (Corollary 2.2.5)

$$\mathcal{A}/\mathfrak{a} \cong \bigoplus_{p=0}^{q-1} \mathcal{A}_p,$$

so in particular $\mathfrak{a}$ is cofinite, whence $\psi \in \mathcal{A}^\circ$. □

Remark 2.3.6. The full linear dual $\mathcal{A}^*$ of a graded algebra consists of formal series of the type

$$\psi = \sum_{p \geq 0} \psi_p$$

where $\psi_p \in \mathcal{A}_p^*$. In other words, we may identify

$$\mathcal{A}^* = \prod_{p=0}^{\infty} \mathcal{A}_p^*,$$

the direct product of the duals. In general this space is much larger than $\mathcal{A}^{\mathrm{gr}}$, except when $\dim \mathcal{A} < \infty$ in which case $\mathcal{A}^* \cong \mathcal{A}^\circ \cong \mathcal{A}^{\mathrm{gr}}$. In fact, when $\mathcal{A}$ is locally finite we can identify $\mathcal{A}^{\mathrm{gr}} \cong \mathcal{A}$ as vector spaces. △

2.4 Bialgebras and Hopf algebras

For this section, besides the references given at the beginning of Section 2.2 the reader is also referred to the more specific P. Cartier's *A Primer of Hopf algebras* [18] and the nice review by D. Manchon [89].

A bialgebra $\mathcal{B} = (B, m, u, \Delta, \varepsilon)$ is a 5-tuple such that (B, m, u) is an algebra and (B, Δ, ε) is a coalgebra, and the either Δ and ε are algebra morphisms, or m and u are coalgebra morphisms. Here, the tensor product $\mathcal{B} \otimes \mathcal{B}$ is supposed to carry the natural algebra or coalgebra structure introduced on the previous sections. A proof of the equivalence between these two statements can be found in [113].

A morphism of bialgebras is a map which is a morphism of algebras and coalgebras as the same time. A biideal of a bialgebra $\mathcal{B}$ is a subspace $\mathfrak{b} \subset \mathcal{B}$ which is an ideal of $(\mathcal{B}, m, \mathbf{1})$ and a coideal of $(\mathcal{B}, \Delta, \varepsilon)$. If $\mathfrak{b}$ is a biideal then $\mathcal{B}/\mathfrak{b}$ has a canonical bialgebra structure and the projection is a morphism of bialgebras.

A Hopf algebra $\mathcal{H}$ is a bialgebra together with a linear map $S : \mathcal{H} \to \mathcal{H}$, called the *antipode*, such that the diagram

$$
\begin{tikzcd}
& \mathcal{H} \otimes \mathcal{H} \arrow[rr, "S\otimes \mathrm{id}"] & & \mathcal{H} \otimes \mathcal{H} \arrow[dr, "m"] & \\
\mathcal{H} \arrow[ur, "\Delta"] \arrow[dr, "\Delta"'] \arrow[rr, "\varepsilon"] & & k \arrow[rr, "u"] & & \mathcal{H} \\
& \mathcal{H} \otimes \mathcal{H} \arrow[rr, "\mathrm{id}\otimes S"'] & & \mathcal{H} \otimes \mathcal{H} \arrow[ur, "m"'] &
\end{tikzcd}
$$

commutes. This means that the identity

$$m(S \otimes \mathrm{id})\Delta x = m(\mathrm{id} \otimes S)\Delta x = \varepsilon(x)\mathbf{1} \tag{2.5}$$

holds for all $x \in \mathcal{H}$. In other words, S is the convolutional inverse of id in $\mathrm{Hom}_k(\mathcal{H}, \mathcal{H})$, hence it is unique when it exists.

Proposition 2.4.1. *Let $\mathcal{H}$ and $\mathcal{H}'$ be two Hopf algebras and $f : \mathcal{H} \to \mathcal{H}'$ a bialgebra map. Then $f \circ S = S' \circ f$.*

The next proposition from [30, 113] gives some of the properties of the antipode.

Proposition 2.4.2. *Let $\mathcal{H}$ be a Hopf algebra with antipode S. Then*

1. $S(xy) = S(y)S(x)$,
2. $S(\mathbf{1}) = \mathbf{1}$,
3. $\Delta S(x) = \sum S(x_2) \otimes S(x_1)$ *and*
4. $\varepsilon \circ S = \varepsilon$.

If $\mathcal{H}$ is commutative or cocommutative then $S^2 = \mathrm{id}$.

The following theorem is of utter importance in what follows. We cite it here without proof, but one can be found in [89].

Theorem 2.4.3. *Let $\mathcal{H}$ be a graded connected bialgebra. Then $\mathcal{H}$ is a Hopf algebra, and its antipode satisfies the recursion*

$$S(x) = -x + \sum_{(x)} S(x_1)x_2 = -x + \sum_{(x)} x_1 S(x_2)$$

for all $x \in \ker \varepsilon$.

2.4.1 Primitive and group-like elements

Let $(\mathcal{H}, m, \Delta, S)$ be a Hopf algebra. We now single out two special classes of elements of $\mathcal{H}$. An element $x \in \mathcal{H}$ is said to be *group-like* if $\Delta x = x \otimes x$; we denote by $G(\mathcal{H})$ the collection of all group-like elements in $\mathcal{H}$. An element $x \in \mathcal{H}$ is said to be *primitive* if $\Delta x = x \otimes \mathbf{1} + \mathbf{1} \otimes x$; we denote by $\mathfrak{g}(\mathcal{H})$ the collection of all primitive elements in $\mathcal{H}$.

We have

Proposition 2.4.4. *The group-like elements form a group with unit $\mathbf{1} \in \mathcal{H}$ and inverses $x^{-1} = S(x)$. Furthermore $\langle \varepsilon, x \rangle = 1$ for all $x \in G(\mathcal{H})$.*

Proof. First note that by definition $\Delta \mathbf{1} = \mathbf{1} \otimes \mathbf{1}$ so $\mathbf{1} \in G(\mathcal{H})$. Let $x, y \in G(\mathcal{H})$. Since Δ is an algebra morphism then

$$\Delta(xy) = \Delta(x)\Delta(y) = xy \otimes xy$$

so $xy \in G(\mathcal{H})$. The counit property implies the identity $x = \langle \varepsilon, x \rangle x$ and so $\langle \varepsilon, x \rangle = 1$. Finally, the defining property of the antipode gives the identity

$$\mathbf{1} = m(S \otimes \mathrm{id})\Delta x = S(x)x = xS(x),$$

that is, $x^{-1} = S(x)$. ◻

Proposition 2.4.5. *The space of primitive elements $\mathfrak{g}(\mathcal{H})$ is a Lie algebra under the bracket $[x, y] = xy - yx$, and $\mathfrak{g}(\mathcal{H}) \subset \ker \varepsilon$.*

Proof. Clearly $\mathfrak{g}(\mathcal{H})$ is a subspace of $\mathcal{H}$. A simple computation yields

$$\begin{aligned}
\Delta[x, y] &= \Delta(x)\Delta(y) - \Delta(y)\Delta(x) \\
&= xy \otimes \mathbf{1} + x \otimes y + y \otimes x + \mathbf{1} \otimes xy - yx \otimes \mathbf{1} - y \otimes x - x \otimes y - \mathbf{1} \otimes yx \\
&= (xy - yx) \otimes \mathbf{1} + \mathbf{1} \otimes (xy - yx) \\
&= [x, y] \otimes \mathbf{1} + \mathbf{1} \otimes [x, y]
\end{aligned}$$

hence $[x, y] \in \mathfrak{g}(\mathcal{H})$. Finally, the properties of the counit give the identity

$$x = \langle \varepsilon, x \rangle \mathbf{1} + x$$

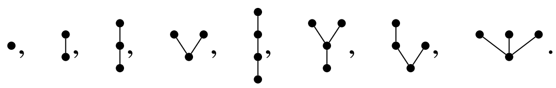
whence $\langle \varepsilon, x \rangle = 0$. $\square$

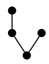
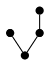
Finally we have the following

Proposition 2.4.6. *Let $\mathscr{A}$ be an algebra and $f \in \mathscr{A}^*$ be an algebra map, i.e. such that $\langle f, xy \rangle = \langle f, x \rangle \langle f, y \rangle$. Then if $(\mathscr{A}^\circ, \Delta, \varepsilon)$ denotes the dual coalgebra as described in Section 2.3.3 we have $f \in \mathscr{A}^\circ$, $\Delta f = f \otimes f$ and $\langle \varepsilon, f \rangle = 1$.*

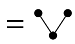
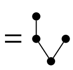
2.5 Algebraic structures on rooted trees

A non-planar rooted tree is a connected acyclic graph $t = (N_t, E_t)$ with one distinguished vertex, called the *root*. Graphically, we draw rooted trees with the root at the bottom and growing northwards; the root is drawn in slightly bigger size. As a convention we think of edges as being oriented away from the root, but this will not be reflected in the graphical representation; given an edge $e = (x, y) \in E_t$ we will call $x, y \in N_t$ its source and target, respectively. The first eight rooted trees are



Non planarity means that we do not distinguish the tree  from . Henceforth we shall drop the adjectives "non-planar" and "rooted" since this will be the only type of trees we will consider. Given a tree t we write $|t|$ for the number of its vertices or nodes. There is a natural partial order on N_t by declaring that $x \leq y$ if and only if there is a path in t from the root to $y \in N_t$ containing the vertex x. For example, the node set of a linear tree is linearly order with its root acting as the minimal element, and its unique leaf as the maximal element; the leaves on the fourth three in the above example are not comparable.

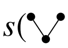
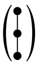
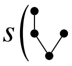
A *rooted forest* is a finite collection of rooted trees, irrespective of order; in particular, trees are a particular kind of forest. There is also the *empty forest* which we will denote by $\mathbf{1}$. We can union two forests two obtain a third forest, and this operation will be denoted simply by juxtaposition both symbolically and graphically. Since a forest can be thought of as the union of its constituent trees, these will be denoted by $t = t_1 t_2 \cdots t_k$, and we set $|t| = |t_1| + |t_2| + \cdots + |t_k|$. The trees $t_1, \ldots, t_k$ forming a forest t are said to be the *factors* of t. Finally, given a forest with k factors $s = t_1 \cdots t_k$, we denote by $B_+(t_1 \cdots t_k)$ the tree obtained by grafting each of the roots of the trees $t_1, \ldots, t_k$ to a new common root, e.g.

$B_+(\mathbf{1}) = \bullet$, $B_+(\bullet) =$ , $B_+(\bullet\bullet) =$ , $B_+($$\bullet) =$ , and it is fairly easy to see that all trees can be obtained from $\bullet$ by iteration of this operation and the union of forests. We also denote by B_- the inverse operation, so for a tree t, $B_-(t)$ is the forest formed by the children of its root.

Finally, we define the symmetry factor $s(t)$ of a tree t recursively by setting $s(\bullet) = 1$ and

$$s(B_+(t_1^{n_1} \cdots t_k^{n_k})) = n_1! \cdots n_k! s(t_1) \cdots s(t_k) \tag{2.6}$$

where $t_1, \dots, t_k$ are distinct trees and t_i is grafted n_1 times onto the root. For example

$s(\bullet) = 1$, $s($$) = 1$, $s($$) = 2$, $s\Big($$\Big) = 1$, $s\Big($$\Big) = 1.$

2.5.1 The Connes–Kreimer Hopf algebra

First, we will describe a Hopf algebra structure introduced in 1998 [24] by A. Connes and D. Kreimer in order to describe the renormalisation procedure in the perturbative expansion of Quantum Field Theories (QFTs for short). Without going into further details, this Hopf algebra describes the combinatorics behind various "renormalisation procedures" which serve to extract suitable counterterms from divergent Feynman graphs, which can be thought of as representing diverging integrals appearing in the Taylor expansion of Green's function of the interaction, with respect to a dimensionless coupling parameter.

We will now proceed to a mathematical description of this Hopf algebra, which we will denote by $\mathcal{H}_{CK}$. As an algebra, $\mathcal{H}_{CK}$ is constructed as the free commutative algebra over the real vector space spanned by trees. In particular, it can be seen to be isomorphic, as a vector space, to the real vector space spanned by forests, with forest union as product and $\mathbf{1}$ as the unit. The coproduct and counit of $\mathcal{H}_{CK}$ are constructed by using its freeness as a commutative algebra in the following way: the counit $\varepsilon : \mathcal{H}_{CK} \to \mathbb{R}$ is the unique algebra morphism such that $\langle \varepsilon, t \rangle = 0$ for all trees. The coproduct $\Delta : \mathcal{H}_{CK} \to \mathcal{H}_{CK} \otimes \mathcal{H}_{CK}$ is the unique algebra morphism such that

$$\Delta B_+(t) = B_+(t) \otimes \mathbf{1} + (\mathrm{id} \otimes B_+)\Delta t. \tag{2.7}$$

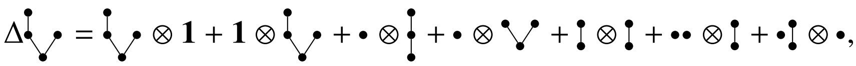
In [24] it is shown that this coproduct is coassociative, and that it admits a description in terms of *cuts*. Given a tree t, we call an arbitrary subset $C \subset E_t$ a *cut*; a cut is said to be *admissible* if any path from the root to any vertex of t contains at most one edge from C, and it is said to be elementary if it consists of a single edge. Any admissible cut C of t containing k edges maps t to a forest $C(t) = t_1 \cdots t_{k+1}$ obtained by deleting the edges from C. It is customary to denote by $R^C(t)$ the unique factor of $C(t)$ containing the original root, and by $P^C(t)$ the forest formed by the rest. This notion can be extended to forests $t = t_1 \cdots t_k$ in the obvious way by choosing an admissible cut C_j for each tree t_j and letting

$$C(t) = C_1(t_1)\cdots C_k(t_k), \quad P^C(t) = P^{C_1}(t_1)\cdots P^{C_k}(t_k), \quad R^C(t) = R^{C_1}(t_1)\cdots R^{C_k}(t_k).$$

If we denote the set of admissible cuts of t by $\mathfrak{A}(t)$ then

$$\Delta t = t \otimes \mathbf{1} + \mathbf{1} \otimes t + \sum_{C \in \mathfrak{A}(t)} P^C(t) \otimes R^C(t). \tag{2.8}$$

for all $t \in \mathcal{H}_{CK}$. As an example

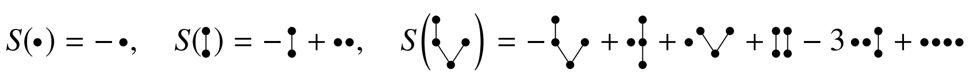


and this shows that Δ is not cocommutative.

Until now, we have defined a commutative bialgebra $\mathcal{H}_{CK}$ on rooted forests. This bialgebra is graded by the number of nodes, that is, its homogeneous component of degree $n \geq 0$ is spanned by forests with exactly n nodes. Since there is only one forest with zero nodes, namely the empty forest $\mathbf{1}$, we have that in fact $\mathcal{H}_{CK}$ is connected. Hence, appealing to a general result [89, Corollary 5] we obtain that $\mathcal{H}_{CK}$ is indeed a Hopf algebra. Moreover, we have the following recursive expression for its antipode $S : \mathcal{H}_{CK} \to \mathcal{H}_{CK}$ by $S(\mathbf{1}) = \mathbf{1}$ and

$$S(t) = -t - \sum_{C \in \mathfrak{A}(t)} S(P^C(t))R^C(t) = -t - \sum_{C \in \mathfrak{A}(t)} P^C(t)S(R^C(t)),$$

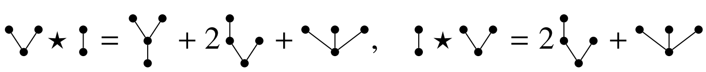
for all $t \in \mathcal{H}_{CK}$. For example,



where we have used the fact that S is an algebra morphism since $\mathcal{H}_{CK}$ is commutative –in general it is only an antihomomorphism.

2.5.2 The Grossman-Larson Hopf algebra

There is another Hopf algebra structure that can be defined on trees which was first defined by R. Grossman and R. Larson in 1989 [48], in order to study efficient ways to compute the action of certain differential operators. We will denote this Hopf algebra by $\mathcal{H}_{GL}$, and as a vector space is spanned by trees. The product of two trees $t = B_+(t_1 \cdots t_k)$ and t, denoted by $t \star t'$, is obtained as the sum of all the ways of grafting the trees $t_1, \ldots, t_k$ to a vertex of t, e.g.

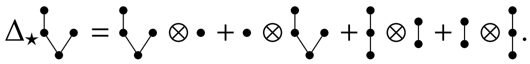


(2.9)

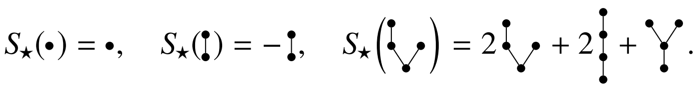
so in particular the product is not commutative. Remark that the tree consisting of a single node acts as the unit. As an algebra $\mathcal{H}_{GL}$ is graded by $\deg(t) = |t| - 1$ and therefore it is also connected. The coproduct $\Delta_\star : \mathcal{H}_{GL} \to \mathcal{H}_{GL} \otimes \mathcal{H}_{GL}$ is given, for $t = B_+(t_1 \cdots t_k)$, by

$$\Delta_\star t = \sum_{I,J} t_I \otimes t_J$$

where the sum ranges over all the partitions I, J of the finite set $\{1, \ldots, k\}$ and $t_I = B_+(t_{i_1} \cdots t_{i_p})$ if $I = \{i_1, \ldots, i_p\} \subset \{1, \ldots, k\}$; also the convention $t_\varnothing = \mathbf{1}$ is in force. For example



In general, $\Delta_\star$ is cocommutative and this endows $\mathcal{H}_{GL}$ with a graded connected cocommutative bialgebra structure. As before, this implies that $\mathcal{H}_{GL}$ is actually a Hopf algebra and we have a similar recursive expression for the antipode $S_\star$. As an example



There is an interesting relation between the Connes–Kreimer and the Grossman–Larson Hopf algebras which was shown by M. E. Hoffman [63]. Given two trees t, t' let $(t, t') = s(t)\delta_{t,t'}$ where $s(t)$ is the symmetry factor defined in eq. (2.6). This definition is extended to rooted forests $t, t' \in \mathcal{F}$ by $(t, t') = (B_+(t), B_+(t'))$; observe this is well defined since for rooted trees $s(t) = s(B_+(t))$.

Theorem 2.5.1. *The linear map* $\chi : \mathcal{H}_{GL} \to \mathcal{H}_{CK}^{\mathrm{gr}}$ *given by* $\langle \chi(t), u \rangle = (B_-(t), u) = (t, B_+(u))$ *is a Hopf algebra isomorphism.*

Thus, the $\star$ product in $\mathcal{H}_{GL}$ as defined above corresponds to the convolution product associated to the Connes–Kreimer coproduct, that is, $\langle \chi(t \star t'), u \rangle = \langle \chi(t) \otimes \chi(t'), \Delta u \rangle$. We will present a full proof of this statement in the next section.

2.5.3 Decorated rooted trees

We now introduce decorated versions of the Connes–Kreimer and Grossman–Larson Hopf algebras, as this is the right context in which Rough Paths theory is formulated. Fix, once and for the rest of this section, a finite alphabet A.

A (non-planar) rooted tree decorated by the alphabet A is a pair $\tau = (t, c)$ where t is a tree in the sense of the previous sections and $c : N_t \to A$ is a function. We denote the *underlying tree* of $\tau = (t, c)$ by $t(\tau) = t$; the collection of all decorated trees will be denoted by $\mathcal{T}$; for a positive integer $n \in \mathbb{N}$, we denote by $\mathcal{T}_n$ (resp. $\mathcal{T}_{(n)}$) the collection of trees having at most (resp. exactly)

n nodes. The vector spaces spanned by these collections will be denoted by $\mathcal{B}$, $\mathcal{B}_n$ and $\mathcal{B}_{(n)}$, respectively. Graphically, we put decorations besides the nodes, e.g.

$$\bullet_a,\quad \overset{\bullet_b}{\underset{\bullet_a}{|}},\quad \overset{\bullet_b\,\bullet_c}{\underset{\bullet_a}{\vee}}.$$

A decorated rooted forest is a forest of decorated trees and we use the notation $\tau = \tau_1 \cdots \tau_k$ where each factor $\tau_i \in \mathcal{T}$. The collection of all decorated rooted forests will be denoted by $\mathcal{F}$, and the empty forests will also be denoted by $\mathbf{1}$. For a natural number $n \in \mathbb{N}$, we denote by $\mathcal{F}_n$ (resp. $\mathcal{F}_{(n)}$) the collection of forests with at most (resp. exactly) n edges; the empty forest is the unique forest with no vertices. For each letter $a \in A$ we have maps $B_+^a : \mathcal{F} \to \mathcal{T}$ where $B_+^a(\tau_1 \cdots \tau_k)$ is the decorated tree obtained by grafting each of the trees $\tau_1, \ldots, \tau_k$ to a new root decorated by a. There is also, for each letter $a \in A$, a map $B_-^a : \mathcal{T} \to \mathcal{F}$ such that

$$B_-^a(B_+^b(\tau_1 \cdots \tau_k)) = \begin{cases} \tau_1 \cdots \tau_k & a = b \\ 0 & a \neq b \end{cases}.$$

All of the previous notions can be imported into this setting, applying them to the underlying tree taking care of the decorations since now decorated trees with identical underlying trees are considered to be different if their decorations do not match for all nodes. For example, if $b \neq c$ we have

$$s\Big(\overset{\bullet_b\,\bullet_c}{\underset{\bullet_a}{\vee}}\Big) = 1, \quad s\Big(\overset{\bullet_b\,\bullet_b}{\underset{\bullet_a}{\vee}}\Big) = 2.$$

The decorated version of the Connes–Kreimer Hopf algebra has been considered in [51]. We will denote this Hopf algebra by $\mathcal{H}_{CK}^A$. As an algebra $\mathcal{H}_{CK}^A$ is the free commutative algebra over $\mathcal{T}$; hence, as a vector space it is isomorphic to $\mathcal{F}$ and the product is given by the disjoint union of decorated forests. As before, it is graded by the number of nodes so in particular the homogeneous component of degree $n \in \mathbb{N}$ is spanned by $\mathcal{F}_{(n)}$. Of course, when the alphabet contains a single letter $\mathcal{H}_{CK}^A \cong \mathcal{H}_{CK}$. The coproduct $\Delta : \mathcal{H}_{CK}^A \to \mathcal{H}_{CK}^A \otimes \mathcal{H}_{CK}^A$ is the same as for $\mathcal{H}_{CK}$ by keeping the decorations in each of $P^C(\tau)$ and $R^C(\tau)$, e.g.

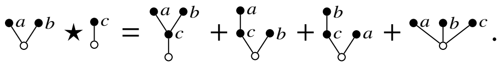
$$\Delta \overset{\bullet_b\,\bullet_c}{\underset{\bullet_a}{\vee}} = \overset{\bullet_b\,\bullet_c}{\underset{\bullet_a}{\vee}} \otimes \mathbf{1} + \mathbf{1} \otimes \overset{\bullet_b\,\bullet_c}{\underset{\bullet_a}{\vee}} + \bullet_b \otimes \overset{\bullet_c}{\underset{\bullet_a}{|}} + \bullet_c \otimes \overset{\bullet_b}{\underset{\bullet_a}{|}} + \bullet_b \bullet_c \otimes \bullet_a.$$

The antipode has a similar expression with the inclusion of the decorations so for example

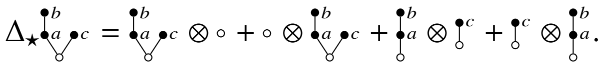
$$S(\bullet_a) = -\bullet_a, \quad S\Big(\overset{\bullet_b\,\bullet_c}{\underset{\bullet_a}{\vee}}\Big) = -\overset{\bullet_b\,\bullet_c}{\underset{\bullet_a}{\vee}} + \bullet_b \overset{\bullet_c}{\underset{\bullet_a}{|}} + \bullet_c \overset{\bullet_b}{\underset{\bullet_a}{|}} - \bullet_a \bullet_b \bullet_c.$$

We will now proceed to describe a decorated version of the Grossman–Larson Hopf algebra, which we denote by $\mathcal{H}_{GL}^A$. For this, we will need a new type of trees, in which the root is distinct in that it is always undecorated. We draw this special root as a blank node, and we have a map

B_+° attaching the trees in a forest to an undecorated root; hence all trees in the vector basis of $\mathcal{H}_{GL}^A$ can be written as $B_+^\circ(\tau_1 \cdots \tau_k)$ for some decorated forest. Given such a tree τ we denote by $\deg(\tau) = |\tau| - 1$ the number of its decorated nodes. The product and the coproduct are given by the same formula as before, so for example



Compare this with eq. (2.9), and notice that there the second tree has a factor of two while here we have two (potentially) different trees due to the decorations. As before, the tree $\circ$ acts as the unit. The coproduct $\Delta_\star : \mathcal{H}_{GL}^A \to \mathcal{H}_{GL}^A \otimes \mathcal{H}_{GL}^A$ is again given by partitions of the children of the root, e.g.



Remark 2.5.2. Of course, considering $\mathcal{H}_{GL}^A$ to be generated, as a vector space, by forests instead of trees with an undecorated root would have given an isomorphic construction. The only caveat is that in that case one has to be more careful in order to define the product and the coproduct. △

2.6 The free pre-Lie algebra

We start this section by recalling the definition of a pre-Lie algebra. A *left pre-Lie algebra* is a vector space $\mathfrak{L}$ together with a linear map $\triangleright : \mathfrak{L} \otimes \mathfrak{L} \to \mathfrak{L}$ such that the *associator* $a(x, y, z) = x \triangleright (y \triangleright z) - (x \triangleright y) \triangleright z$ is left symmetric, i,e.

$$x \triangleright (y \triangleright z) - (x \triangleright y) \triangleright z = y \triangleright (x \triangleright z) - (y \triangleright x) \triangleright z \tag{2.10}$$

for all $x, y, z \in \mathfrak{L}$. If the associator is right-symmetric instead of left-symmetric, $\mathfrak{L}$ is said to be a right pre-Lie algebra. The following proposition justifies the name.

Proposition 2.6.1. *Let* $(\mathfrak{L}, \triangleright)$ *be a pre-Lie algebra, and define* $[x, y]_\triangleright = x \triangleright y - y \triangleright x$. *Then* $(\mathfrak{L}, [\cdot, \cdot]_\triangleright)$ *is a Lie algebra.*

Proof. It is clear that $[\cdot, \cdot]_\triangleright$ is bilinear and $[x, x]_\triangleright = 0$ for all $x \in \mathfrak{L}$. Now we check the Jacobi identity. Let $x, y, z \in \mathfrak{L}$, then

$$\begin{aligned}
[x, [y, z]_\triangleright]_\triangleright &= x \triangleright [y, z]_\triangleright - [y, z]_\triangleright \triangleright x \\
&= x \triangleright (y \triangleright z - z \triangleright y) - (y \triangleright z - z \triangleright y) \triangleright x \\
&= x \triangleright (y \triangleright z) - x \triangleright (z \triangleright y) - (y \triangleright z) \triangleright x + (z \triangleright y) \triangleright x,
\end{aligned}$$

and

$$
\begin{aligned}
[z,[x,y]_\triangleright]_\triangleright &= z \triangleright [x,y]_\triangleright - [x,y]_\triangleright \triangleright z \\
&= z \triangleright (x \triangleright y - y \triangleright x) - (x \triangleright y - y \triangleright x) \triangleright z \\
&= z \triangleright (x \triangleright y) - z \triangleright (y \triangleright x) - (x \triangleright y) \triangleright z + (y \triangleright x) \triangleright z,
\end{aligned}
$$

and also

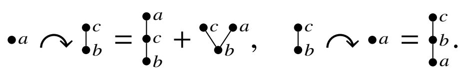
$$
\begin{aligned}
[y,[z,x]_\triangleright]_\triangleright &= y \triangleright [z,x]_\triangleright - [z,x]_\triangleright \triangleright y \\
&= y \triangleright (z \triangleright x - x \triangleright z) - (z \triangleright x - x \triangleright z) \triangleright y \\
&= y \triangleright (z \triangleright x) - y \triangleright (x \triangleright z) - (z \triangleright x) \triangleright y + (x \triangleright z) \triangleright y,
\end{aligned}
$$

hence the sum $[x,[y,z]_\triangleright]_\triangleright + [z,[x,y]_\triangleright]_\triangleright + [y,[z,x]_\triangleright]_\triangleright = 0$ by eq. (2.10). □

Now, let A be a finite alphabet and $\mathfrak{T}$ be the collection of all rooted trees decorated by A, as above. Given three trees $\tau, \sigma, \rho \in \mathfrak{T}$ denote by $n(\tau,\sigma;\rho)$ the number of elementary cuts on ρ such that $P^C(\rho) = \tau$ and $R^C(\rho) = \sigma$. Define a bilinear operation on $\mathfrak{T}$ by

$$
\tau \curvearrowright \sigma = \sum_{\rho \in \mathfrak{T}} \frac{s(\sigma)s(\tau)}{s(\rho)} n(\tau,\sigma;\rho)\rho.
$$

Observe that the sum is actually finite since $n(\tau,\sigma;\rho)$ is non-zero if and only if $|\rho| = |\tau| + |\sigma|$. For example



This operation makes $(\mathfrak{T}, \curvearrowright)$ a pre-Lie algebra as is shown in [22]. In fact, we have

Theorem 2.6.2 (Chapoton–Livernet)**.** *The pre-Lie algebra $(\mathfrak{T}, \curvearrowright)$ is the free pre-Lie algebra over the vector space spanned by A, once one considers the natural embedding $\mathbb{R}A \cong \mathfrak{T}_1 \subset \mathfrak{T}$.*

This operations bears an interesting relation with the decorated Grossman–Larson product defined in Section 2.5.3.

Proposition 2.6.3. *The identity $\tau \curvearrowright \sigma - \sigma \curvearrowright \tau = B_+^\circ(\tau) \star B_*^\circ(\sigma) - B_+^\circ(\sigma) \star B_+^\circ(\tau)$ holds for all $\tau, \sigma \in \mathfrak{T}$.*

This Proposition together with the Chapoton–Livernet theorem 2.6.2 provides us with a procedure for defining characters on $\mathcal{H}_{CK}^A$ as follows: let $\mathscr{A}$ be an associative algebra, pick $a_1, \ldots, a_d \in \mathscr{A}$ and define $\phi(\bullet_i) = a_i$ for all $i = 1, \ldots, d$. Observe that since $\mathscr{A}$ is associative its product endows it with a preLie algebra structure. Hence, ϕ extends uniquely to a preLie algebra morphism $\phi : \mathscr{B} \to \mathscr{A}$, and further to a Lie algebra morphism $\phi : \mathscr{B} \to \mathfrak{L}_{\mathscr{A}}$. Finally, we use the fact that $\mathcal{H}_{CK}^A$ is the universal enveloping algebra of its primitive elements and the previous proposition.

Chapter 3

Analysis

In this chapter we will describe the main analytical tools needed to understand –and prove– some of the results in this work. We will mainly follow M. Gubinelli's line of work, introduced in [50, 51]. There are also some classical notions such as Hölder and Besov regularity, for which the reader is referred to the classical textbooks [32, 90, 116].

3.1 Hölder spaces

Let $[0,T]$ be a fixed time interval, and fix once and for the rest of this section a normed vector space $(V, |\cdot|)$. For a function $f : [0,T] \to V$ and $\alpha \in (0,1)$ let

$$[f]_\alpha \coloneqq \sup_{(s,t)\in[0,T]^2} \frac{|f_t - f_s|}{|t-s|^\alpha}. \tag{3.1}$$

The classical α-Hölder space C^α is the vector space of functions such that $[f]_\alpha < \infty$. Note that eq. (3.1) defines only a seminorm, because it becomes zero for all constant functions. Hölder functions are necessarily continuous since we can bound the differences $|f_{t+h} - f_t| \leq [f]_\alpha |h|^\alpha$. The seminorm on eq. (3.1) can be turned into a norm in two different inequivalent ways: one can consider $\|f\|_\alpha = |f_0| + [f]_\alpha$ or the space $\dot{C}^\alpha = C^\alpha/\mathbb{R} = \{f \in C^\alpha : f_0 = 0\}$. We will also denote by C^0 the space of continuous functions.

Proposition 3.1.1. *Let $0 < \alpha < 1$. The Hölder space C^α is an algebra under pointwise product.*

Proof. Let $f, g \in C^\alpha$. Then

$$\begin{aligned}|f_t g_t - f_s g_s| &\leq |g_t(f_t - f_s)| + |f_s(g_t - g_s)| \\ &\leq (\|g\|_\infty [f]_\alpha + \|f\|_\infty [g]_\alpha)|t-s|^\alpha\end{aligned}$$

whence $fg \in C^\alpha$ and $[fg]_\alpha \leq \|g\|_\infty [f]_\alpha + \|f\|_\infty [g]_\alpha$. □

Since we will be working on a compact interval we have the following

Proposition 3.1.2. *There is an inclusion $C^1 \subset C^\beta \subset C^\alpha \subset C^0$ if $0 < \alpha < \beta < 1$. Moreover, this inclusion is continuous, that is, the canonical map $\iota : C^\beta \to C^\alpha$ is continuous.*

Proof. The inclusion $C^\alpha \subset C^0$ was shown in the previous paragraph. A simple computation gives that if $0 < \alpha < \beta < 1$ then $[f]_\alpha \leq T^{\beta-\alpha}[f]_\beta$. In particular the inclusion map $\iota : C^\beta \to C^\alpha$ is bounded and $\|\iota\|_\alpha \leq T^{\beta-\alpha}$. Moreover, if $f \in C^1$ then its derivative is bounded and so by the Mean Value Theorem $[f]_\beta \leq T^{1-\beta}\|f'\|_\infty$. □

It should be noted however that in general these inclusions are strict. As an example take $V = \mathbb{R}$ and consider the function $f_t = t^\alpha$ which is in C^α since the inequality $(t+s)^\alpha \leq t^\alpha + s^\alpha$ holds for all $s,t \in [0,T]$ and $\alpha \in (0,1)$. Therefore, $|f_t - f_s| \leq |t-s|^\alpha$ and so $[f]_\alpha \leq 1$. But if $\beta > \alpha$ then $\frac{|f_t|}{t^\beta} = t^{\alpha-\beta}$ is unbounded so f cannot lie in C^β. The same example shows that the inclusion of C^1 into C^α is strict for any $\alpha \in (0,1)$ since $f'_t \to \infty$ when $t \to 0$. Taking

$$f_t = \frac{1}{\log(2\frac{t}{T})}$$

with $f_0 = 0$ shows that the inclusion $C^\alpha \subset C$ is strict. In fact, if f were in C^α for some $\alpha \in (0,1)$ then we would have that

$$Ct^\alpha \log(t) \geq 1$$

for some positive constant $C > 0$ and all $t \in [0,1/2]$ which is impossible since $t^\alpha \log(t) \to 0$ as $t \to 0^+$.

3.1.1 Finite increments

These definitions can be extended to functions depending on more than one variable by introducing suitable finite increments. We denote by $\mathscr{C}_k$ the space of continuous functions from $[0,T]^k$ to V such that $f_{t_1,\ldots,t_k} = 0$ whenever $t_j = t_{j+1}$ for some $j = 1,\ldots,k-1$. For each $k \geq 1$ define an operator $\delta_k : \mathscr{C}_k \to \mathscr{C}_{k+1}$ by

$$\delta_k f_{t_1,\ldots,t_{k+1}} := \sum_{j=1}^{k+1} (-1)^{k+j} f_{t_1,\ldots,\hat{t}_j,\ldots,t_{k+1}} \tag{3.2}$$

where $\hat{t}_j$ means the argument is omited. In particular we have that $\delta_1 f_{st} = f_t - f_s$ and

$$\delta_2 f_{sut} = f_{st} - f_{su} - f_{ut}.$$

It can be shown [50] that

$$0 \to \mathbb{R} \to \mathscr{C}_1 \xrightarrow{\delta_1} \mathscr{C}_2 \xrightarrow{\delta_2} \mathscr{C}_3 \xrightarrow{\delta_3} \cdots$$

is an *acyclic cochain complex*. This means that for all $k \geq 1$ one has $\delta_{k+1} \circ \delta_k = 0$ and $\ker \delta_{k+1} = \delta_k \mathscr{C}_k$. Concretely, each time $\delta_{k+1} f = 0$ we can find a function $g \in \mathscr{C}_k$ such that $f = \delta_k g$. From now on, we set

$$\mathscr{C}_* = \bigoplus_{k \geq 1} \mathscr{C}_k$$

and by a slight abuse of notation we obtain an operator $\delta : \mathscr{C}_* \to \mathscr{C}_*$ such that $\delta^2 = 0$.

Given $\gamma > 0$, and F in $\mathscr{C}_2$ define

$$\|F\|_\gamma = \sup_{(s,t) \in [0,T]^2} \frac{|F_{st}|}{|t-s|^\gamma}.$$

Let $\mathscr{C}_2^\gamma \subset \mathscr{C}_2$ denote the space of functions such that $\|F\|_\gamma < \infty$. Also, let

$$\mathscr{C}_2^{1+} = \bigcup_{\gamma > 1} \mathscr{C}_2^\gamma.$$

Given $\gamma, \rho > 0$ and $F \in \mathscr{C}_3$ define

$$\|F\|_{\gamma,\rho} = \sup_{(s,u,t) \in [0,T]^3} \frac{|F_{sut}|}{|u-s|^\gamma |t-u|^\rho}$$

and

$$\|F\|_\gamma = \inf \Big\{ \sum \|G_i\|_{\rho_i, \gamma - \rho_i} : F = \sum G_i, 0 < \rho_i < \gamma \Big\}.$$

Let $\mathscr{C}_3^\gamma \subset \mathscr{C}_3$ denote the space of functions such that $\|F\|_\gamma < \infty$ and set

$$\mathscr{C}_3^{1+} = \bigcup_{\gamma > 1} \mathscr{C}_3^\gamma.$$

We will also need to consider the spaces $Z\mathscr{C}_3^\gamma = \mathscr{C}_3^\gamma \cap \ker \delta$ and $Z\mathscr{C}_3^{1+} = \mathscr{C}_3^{1+} \cap \ker \delta$.

The following result by M. Gubinelli is of central importance to the theory of branched rough paths.

Theorem 3.1.3 (Sewing Lemma, [51])**.** *There exists a unique linear map* $\Lambda : Z\mathscr{C}_3^{1+} \to \mathscr{C}_2^{1+}$ *such that* $\Lambda\delta = \mathrm{id}$. *Moreover, its restriction to* $Z\mathscr{C}_3^\gamma$ *satisfies the bound*

$$\|\Lambda F\|_\gamma \leq \frac{1}{1 - 2^{-\gamma}} \|F\|_\gamma.$$

In the case V is also an algebra, given a continuous function f and $F, G \in \mathscr{C}_2$ we define new functions $fF \in \mathscr{C}_2$, $FG \in \mathscr{C}_3$ and $F \cdot G \in \mathscr{C}_2$ by

$$(fF)_{st} = f_s F_{st}, \quad (FG)_{sut} = F_{su} G_{ut}, \quad (F \cdot G)_{st} = F_{st} G_{st}.$$

It is easy to see that the following relations

$$\delta(fF) = f\delta F - (\delta f)F \tag{3.3}$$

and

$$\delta(F \cdot G) = (\delta F) \cdot (\delta G) + (Fe + eF) \cdot \delta G + \delta F \cdot (Ge + eG) + FG + GF \tag{3.4}$$

hold, where $e \in \mathscr{C}_2$ is the function given by $e_{st} = 1$ for all $(s,t) \in [0,T]^2$. The space $\mathscr{C}_2$ is an algebra under the $\cdot$ product with unit e.

3.2 Wavelets

Let L_c^2 denote the space of functions $f : \mathbb{R} \to \mathbb{C}$ such that $t \mapsto |f(x)|^2$ is integrable with respect to $\mathrm{d}x$. The standard Fourier transform $\mathscr{F} : L_c^2 \to L_c^2$ given by

$$(\mathscr{F} f)(\omega) := \int_{-\infty}^{\infty} f(x) \mathrm{e}^{-2\pi \mathrm{i} x\omega} \, \mathrm{d}x \tag{3.5}$$

describes the global "frequency content" of the function $f \in L_c^2$. This information can sometimes be used to describe the global oscillatory behaviour of a given function but it is too sensitive to discontinuities as the "filtering function" $\mathrm{e}^{-2\pi \mathrm{i} x\omega}$ is spread out over the whole real line. In other words, a change in local behaviour of f –think of a jump discontinuity at a given point, for example– translates into a modification of the whole spectrum content represented by its Fourier transform $\mathscr{F} f$; moreover, given this frequency representation it is in general hard to pinpoint where such events occur as the exponential kernel appearing in eq. (3.5) is not well localised in space. One way to attempt to solve this problem is to introduce a *windowed* Fourier transform. A window function $w : \mathbb{R} \to \mathbb{R}$ is, roughly, a compactly supported and smooth function. Spatial localisation can then be achieved by first windowing the function f and then taking the standard Fourier transform, i.e.

$$(\mathscr{F}_w f)(\omega, x) := \int_{-\infty}^{\infty} f(y) w(y - x) \mathrm{e}^{-2\pi \mathrm{i} y\omega} \, \mathrm{d}y. \tag{3.6}$$

Observe that this *windowed Fourier transform* depends on two parameters, where the extra parameter $x \in \mathbb{R}$ now describes the centre of the spatial localisation. We remark that here the filtering function $y \mapsto w(y-x)\mathrm{e}^{-2\pi \mathrm{i} y\omega}$ is now well localised in space and in frequency. Under suitable assumptions on the window function, it can be shown that the coefficients eq. (3.6) are enough to characterise (and reconstruct) the original function f by means of the formula

$$f(x) = \frac{1}{w(0)} \int_{-\infty}^{\infty} (\mathscr{F}_w f)(\omega, x)\mathrm{e}^{2\pi \mathrm{i} x\omega}\,\mathrm{d}\omega.$$

In fact, only a countable number of such coefficients are actually needed; these coefficients correspond to discrete translate of the filter by some fixed $x_0 > 0$ and frequency ω_0:

$$c_{n,k}(f) \coloneqq (\mathscr{F}_w f)(nx_0, k\omega_0) = \int_{-\infty}^{\infty} f(y)g(y-nx_0)\mathrm{e}^{-2k\pi \mathrm{i} y\omega_0}\,\mathrm{d}y.$$

It turns out that there is another, related, type of transform, called the *wavelet transform*. Given a function $\psi : \mathbb{R} \to \mathbb{R}$ called the *mother wavelet* –whose precise properties will be detailed down below– and $a, b \in \mathbb{R}$ with $a \neq 0$ let

$$(\mathscr{W} f)(a,b) \coloneqq \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(x)\psi\left(\frac{x-b}{a}\right)\mathrm{d}x \tag{3.7}$$

be the *continuous wavelet transform* of $f \in L^2_c$. Note the resemblance with the windowed Fourier transform eq. (3.6). Loosely speaking, we require that the functions $\psi^{a,b}(x) = |a|^{-1/2}\psi(\frac{x-b}{a})$ be well localised in time and frequency, as for the Fourier transform. Then, eq. (3.7) provides another time-frequency description of a function $f \in L^2_c$. The first parameter is related to "scale", i.e. frequency localisation and b is simply is a time shift in order to change localisation in time. An example of an admissible wavelet function is the *Mexican hat function* $\psi(t) = (1-x^2)e^{-x^2/2}$; other, more interesting examples will be examined in what follows.

One of the main differences between eq. (3.6) and eq. (3.7) is that the window size in the windowed Fourier transform is the same regardless of its frequency localisation whereas the wavelets $\psi^{a,b}$ have their "window sizes" adapted to the frequency localisation. This allows for wavelets to better describe the behaviour of f at small scales; we will see that this translates into a better ability to describe function and distributions spaces than what can be done with the Fourier transform.

In this context there is also an inversion formula,

$$f = C_\psi^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{a^2}(\mathscr{W} f)(a,b)\psi^{a,b}\,\mathrm{d}a\mathrm{d}b. \tag{3.8}$$

where the normalising constant

$$C_\psi = \int_{-\infty}^{\infty} \frac{|(\mathscr{F}\psi)(\omega)|}{|\omega|}\,\mathrm{d}\omega$$

is assumed to be finite. If the mother wavelet ψ is integrable, then C_ψ can be finite only if $(\mathscr{F}\psi)(0) = 0$, i.e. if $\int \psi = 0$. One can also set up a discrete wavelet transform obtained by considering translates and scaling by fixed $a_0 > 1$ and $b_0 > 1$, i.e.

$$\psi_{n,k}(x) \coloneqq a_0^{n/2}\psi(a_0^n x - kb_0). \tag{3.9}$$

for $n, k \in \mathbb{Z}$. Unfortunately, since there is some redundancy in the wavelet coefficients

$$\langle f, \psi_{n,k} \rangle \coloneqq \int_{-\infty}^{\infty} f(x)\psi_{n,k}(x)\,\mathrm{d}x \tag{3.10}$$

in general there is no "inversion formula" just like eq. (3.8). This inversion problem is related to the fact that the $\psi_{n,k}$ should form an orthonormal basis of L_c^2, which may not be true for an arbitrary choice of ψ and a_0, b_0. The choice of mother wavelet ψ is only constrained by the fact that the normalising constant C_ψ appearing in eq. (3.8) should be finite. For reasons made clear in the above paragraphs, however, we also require that ψ is well concentrated in time and in frequency. It turns out that for some special choices of ψ, a_0 and b_0, the collection $(\psi_{n,k} : n, k \in \mathbb{Z})$ forms an orthonormal basis of L_c^2.

3.2.1 Multiresolution analysis

One of the principal examples of such a special choice is the Haar wavelet basis, defined in 1910 by A. Haar [53]. The mother wavelet is the step function

$$\psi(x) = \begin{cases} 1 & 0 \le x < \frac{1}{2} \\ -1 & \frac{1}{2} \le x < 1 \\ 0 & \text{otherwise} \end{cases}. \tag{3.11}$$

Even though this function is not very well localised, its translates and scalings do form an orthonormal basis of L_c^2, see [32] for a proof. In the proof, one builds the sequence of subspaces V_n consisting of functions in L_c^2 which are constant in each interval $[2^n k, 2^n(k+1))$. This nested sequence $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots \subset L_c^2$ is such that

$$\overline{\bigcup_{n\in\mathbb{Z}} V_n} = L_c^2$$

and

$$\bigcap_{n\in\mathbb{Z}} V_n = \{0\}.$$

Moreover, they satisfy $f \in V_n$ if and only if $f(2^n\cdot) \in V_0$ and $f \in V_0$ if and only if $f(\cdot - k) \in V_0$ for all $k, n \in \mathbb{Z}$. Also, the function φ such that $\varphi(x) = 1$ if $0 \le x < 1$ and 0 else is such that

$(\varphi(\cdot - k) : k \in \mathbb{Z})$ is a Schauder basis for V_0.

Definition 3.2.1. *A sequence of subspaces $(V_n : n \in \mathbb{Z})$ of L^2_c such that*

1. $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$,

2. $$\overline{\bigcup_{n\in\mathbb{Z}} V_n} = L^2_c,$$

3. $$\bigcap_{n\in\mathbb{Z}} V_n = \{0\},$$

4. $f \in V_n$ *if and only if* $f(2^n\cdot) \in V_0$ *for all* $n \in \mathbb{Z}$,

5. $f \in V_0$ *if and only if* $f(\cdot - k) \in V_0$ *for all* $k \in \mathbb{Z}$, *and*

6. *there is* $\varphi \in V_0$ *such that* $(\varphi(\cdot - k) : k \in \mathbb{Z})$ *is a Schauder basis for* V_0

is called a multiresolution analysis. *The function* φ *appearing in 6. is called the* scaling function *of the multiresolution analysis.*

We remark right away that the scaling condition 4. together with 6. implies that for each $n \in \mathbb{Z}$ the functions $\varphi_{n,k}(t) = 2^{n/2}\varphi(2^n t - k)$, $k \in \mathbb{Z}$ for an orthonormal (Schauder) basis for V_n. In particular, observe that the functions $\varphi_{1,k}(x) = \sqrt{2}\varphi(2x - k)$, $k \in \mathbb{Z}$ span V_1 by 4. and 6. above, and since $V_0 \subset V_1$ by 1. then there are coefficients $(\alpha_k : k \in \mathbb{Z})$ such that

$$\varphi(x) = \sum_{k\in\mathbb{Z}} \alpha_k \varphi(2x - k).$$

It turns out that setting

$$\psi(x) \coloneqq \sum_{k\in\mathbb{Z}} (-1)^k \alpha_{1-k} \varphi(2x - k)$$

and $\psi_{n,k}(x) \coloneqq 2^{-n/2}\psi(2^{-n}x - k)$ then $(\psi_{n,k} : k \in \mathbb{Z})$ spans the orthogonal complement W_n of V_n in V_{n-1}.

Proposition 3.2.2. *Let $(V_n :\in \mathbb{Z})$ be a multiresolution analysis. Then there exists an associated orthonormal wavelet basis $(\psi_{n,k} : n, k \in \mathbb{Z})$ of L^2_c such that, if P_n denotes the projection onto V_n then*

$$P_{n+1} = P_n + \sum_{k\in\mathbb{Z}} \langle \cdot, \psi_{n,k} \rangle \psi_{n,k}. \tag{3.12}$$

A simple consequence of 2. in definition 3.2.1 is that the projections $P_n f$ converge to f in L^2_c as $n \to \infty$. Then, eq. (3.12) implies that the decomposition

$$f = \sum_{k \in \mathbb{Z}} \langle f, \varphi_{0,k} \rangle \varphi_{0,k} + \sum_{n=0}^{\infty} \sum_{k \in \mathbb{Z}} \langle f, \psi_{n,k} \rangle \psi_{n,k} \tag{3.13}$$

also holds in L^2_c. Depending on the regularity properties of the mother wavelet, a multiresolution analysis can be used to describe and characterise functional and distributional classes other than L^2_c.

3.2.2 Function spaces

One of the main tools of the preceding work has been the Fourier transform. We now present the Schwartz space of rapidly decreasing functions, which is specially adapted to this kind of transformation. Given $n, k \in \mathbb{N}$ and a smooth function $f : \mathbb{R} \to \mathbb{C}$ let

$$\|f\|_{n,k} \coloneqq \sup_{t \in \mathbb{R}} |x^k f^{(n)}(x)| \tag{3.14}$$

The linear space $\mathcal{S}$ composed of all smooth functions such that $\|f\|_{n,k} < \infty$ for all $n, k \in \mathbb{N}$ is known as the *Schwartz space* in honour of L. Schwartz who defined them for the first time [108]. We observe that the classical space C_c^∞ of smooth functions with compact support –also known as test functions– is included in $\mathcal{S}$, and if $\phi \in \mathcal{S}$ then $(1 + x^2)\phi(x)$ is bounded and so

$$\int_{-\infty}^{\infty} |\phi(x)|^p \, \mathrm{d}x \le \sup_{x \in \mathbb{R}} |(1 + x^2)\phi(x)|^p \int_{-\infty}^{\infty} \frac{1}{1 + x^2} \, \mathrm{d}x < \infty$$

for all $1 \le p \le \infty$, hence we have the inclusions $C_c^\infty \subset \mathcal{S} \subset L^p$. A standard result then implies that the Schwartz space is dense in L^p in the L^p norm, for all $p \in [1, \infty]$.

The standard topology on $\mathcal{S}$ is given by the family of norms defined in eq. (3.14). This topology is induced by the metric

$$(\phi, \psi) \mapsto \sum_{n,k \ge 0} 2^{-(n+k)} \frac{\|\phi - \psi\|_{n,k}}{1 + \|\phi - \psi\|_{n,k}}$$

and $\mathcal{S}$ becomes a complete metric space. Its topological dual $\mathcal{S}'$, the space of *tempered distributions*, consists of all linear functionals $\xi : \mathcal{S} \to \mathbb{C}$ continuous with respect to the above topology. Given $\phi \in \mathcal{S}$, the linear functional $T_\phi : \mathcal{S} \to \mathbb{C}$ given by

$$\langle T_\phi, \psi \rangle \coloneqq \int_{-\infty}^{\infty} \phi(x)\psi(x) \, \mathrm{d}x$$

is continuous in the topology of $\mathcal{S}$ by a standard argument. Moreover, it is an injection since if $T_\phi = 0$ then $\langle T_\phi, \phi\rangle = \|\phi\|_{L^2}^2 = 0$ whence $\phi = 0$. We therefore have a continuous injection $\mathcal{S} \hookrightarrow \mathcal{S}'$. In fact, since the inclusion $\mathcal{S} \subset L^p$ is also continuous for all $p \in [1, \infty]$ and $\mathcal{S}$ is dense in L^p, in a similar way we have that $\mathcal{S} \hookrightarrow L^p \hookrightarrow \mathcal{S}'$ for all $1 \leq p \leq \infty$.

The main feature of the Schwartz space is that it bears a close relation to the Fourier transform.

Theorem 3.2.3. *The Fourier transform is a continuous automorphism of* $\mathcal{S}$.

Proof. The result is standard but we provide a (partial) proof nonetheless for the sake of completeness, and because it introduces two useful properties of the Fourier transform. Let $\phi \in \mathcal{S}$. We first observe from the definition that all the derivatives of ϕ, as well as its product by any polynomial still belong to $\mathcal{S}$. Next, let $\omega \in \mathbb{R}$; for any $h \in \mathbb{R}$ we then have

$$\begin{aligned} \mathcal{F}\phi(\omega + h) - \mathcal{F}\phi(\omega) &= \int_{-\infty}^{\infty} \phi(x) \mathrm{e}^{-2\pi \mathrm{i} x\omega} \left(\mathrm{e}^{-2\pi \mathrm{i} x h} - 1\right) \mathrm{d}x \\ &= -2\pi \mathrm{i} h \int_{-\infty}^{\infty} x\phi(x) \mathrm{e}^{-2\pi \mathrm{i} x\omega} \,\mathrm{d}x + \mathrm{o}(h). \end{aligned}$$

since $x \mapsto x^n \phi(x)$ is a Schwartz function –in particular it is integrable– for all $n \in \mathbb{N}$. Therefore, we obtain the formula

$$\frac{\mathrm{d}}{\mathrm{d}\omega} \mathcal{F}\phi = -2\pi \mathrm{i} \mathcal{F}(x\phi), \tag{3.15}$$

valid for all $\phi \in \mathcal{S}$. Iterating eq. (3.15) we obtain the formula

$$\frac{\mathrm{d}^n}{\mathrm{d}\omega^n} \mathcal{F}\phi = (-2\pi \mathrm{i}) \mathcal{F}(x^n \phi) \tag{3.16}$$

for all $n \in \mathbb{N}$. In particular, $\mathcal{F}\phi \in C^\infty$ for any $\phi \in \mathcal{S}$.

Now we have to evaluate the behaviour of the Fourier transform at infinity. For this, observe that

$$\begin{aligned} \omega \mathcal{F}\phi(\omega) &= \int_{-\infty}^{\infty} \omega \mathrm{e}^{-2\pi \mathrm{i} x\omega} \phi(x) \,\mathrm{d}x \\ &= -\frac{1}{2\pi \mathrm{i}} \int_{-\infty}^{\infty} \phi(x) \frac{\mathrm{d}}{\mathrm{d}x} e^{-2\pi \mathrm{i} x\omega} \,\mathrm{d}x \\ &= \frac{1}{2\pi \mathrm{i}} \mathcal{F}(\phi')(\omega) \end{aligned}$$

after integrating by parts and using the behaviour of ϕ at infinity. We have shown that

$$\mathcal{F}(\phi')(\omega) = 2\pi \mathrm{i} \omega \mathcal{F}\phi(\omega) \tag{3.17}$$

and therefore

$$\mathcal{F}(\phi^{(n)})(\omega) = (2\pi \mathrm{i} \omega)^n \mathcal{F}\phi(\omega) \tag{3.18}$$

is valid for all $n \in \mathbb{N}$.

Next, if ψ is any Schwartz function then

$$|\mathscr{F}\psi(\omega)| \leq \int_{-\infty}^{\infty} |\psi(x)| \,\mathrm{d}x < \infty$$

so in particular $\mathscr{F}\psi$ is bounded. This fact, together with formulas (3.16) and (3.18) imply that $\omega^k \frac{\mathrm{d}^n}{\mathrm{d}\omega^n}\mathscr{F}\phi$ is bounded for all $k, n \in \mathbb{N}$, that is, $\mathscr{F}\phi \in \mathcal{S}$ whenever $\phi \in \mathcal{S}$.

Continuity is implied by the fact that if $\phi_n \to \phi$ in $\mathcal{S}$, in particular $\phi_n \to \phi$ uniformly. A proof of the invertibility of $\mathscr{F}$ can be found in [108]. □

The continuity of this transformations allows to extend $\mathscr{F}$ to $\mathcal{S}'$. Indeed, if $\xi \in \mathcal{S}'$ is a tempered distribution, then $\langle \mathscr{F}\xi, \psi\rangle = \langle \xi, \mathscr{F}\psi\rangle$ uniquely defines the tempered distribution $\mathscr{F}\xi \in \mathcal{S}'$.

Other classes of interesting subspaces of $\mathcal{S}'$ are the *Besov spaces* $B^s_{p,q}$ for $s \in \mathbb{R}$, $p, q \in [0, \infty]$ (although we will only be interested in the cases where $p, q \in (1, \infty)$). These spaces admit various –equivalent, of course– definitions, but essentially they capture the local behaviour (regularity, integrability) of its members, see for example [117] for a precise definition in these terms. The first parameter $s \in \mathbb{R}$ is closely related to smoothness in the fashion of the Hölder spaces C^α or the smoothness parameter in the Sobolev spaces W^s_p. The second parameter $0 \leq p \leq \infty$ describes integrability, and plays basically the same role as in the Sobolev spaces $W^{s,p}$. Finally, the third parameter $0 \leq q \leq \infty$ corresponds to approximation and, in essence, describes the growth of the coefficients of an approximating function series; c.f. the coefficients of the Fourier series of an L^2 function are in ℓ^2. In fact, the Besov spaces contain both C^α and W^s_p as special cases, hence they provide a finer way to analyse functions on the real line.

For our purposes it will enough to characterise $B^s_{p,q}$ in terms of discrete wavelet expansions *à la* eq. (3.13), as is done in [90]. See also [61]. Let (ϕ, ψ) be the scaling function and the mother wavelet associated to the Daubechies orthonormal r-regular multresolution analysis, i.e. ϕ is of class C^r with compact support. Recall that we denote $\phi_k = \phi(\cdot - k)$ and $\psi_{n,k} = 2^{n/2}\psi(2^n \cdot -k)$ for $k, n \in \mathbb{Z}$.

Definition 3.2.4. *Let $1 \leq p, q \leq \infty$ and $|s| < r$. The Besov space $B^s_{p,q}$ is the space of distributions $\xi \in \mathcal{S}'$ such that*

$$\|\xi\|_{B^s_{p,q}} := \left(\sum_{k\in\mathbb{Z}} |\langle \xi, \phi_k\rangle|^p\right)^{1/p} + \left[\sum_{n=0}^{\infty} 2^{j(qs+q/2-q/p)}\left(\sum_{k\in\mathbb{Z}} |\langle \xi, \psi_{n,k}\rangle|^p\right)^{q/p}\right]^{1/q} < \infty \tag{3.19}$$

with the usual modifications when $p, q = \infty$.

It can be shown that each $B^s_{p,q}$ is a Banach space, for $s \in \mathbb{R}$ and $1 \leq p, q \leq \infty$. Moreover, the following equalities hold: $B^s_{\infty,\infty} = C^s$ for $0 < s < 1$ and $B^s_{p,p} = W^{s,p}$ for $s > 0$ not an integer.

Part II

Results

Chapter 4

Hopf-algebraic deformations of products

4.1 Introduction

Chaos expansions and Wick products have notoriously been thought of as key steps in the renormalisation process in perturbative quantum field theory (QFT). The technical reason for this is that they allow to remove contributions to amplitudes (say, probability transitions between two physical states) that come from so-called diagonal terms – from which divergences in the calculation of those amplitudes may originate. Rota and Wallstrom [107] addressed these issues from a strictly combinatorial point of view using, in particular, the structure of the lattice of set partitions. These are the same techniques that are currently used intensively in the approach by Peccati and Taqqu in the context of Wiener chaos and related phenomena. We refer to their book [100] for a detailed study and the classical results on the subject, as well as for a comprehensive bibliography and historical survey.

Recently, the interest in the fine structure of cumulants and Wick products for non-Gaussian variables has been revived, since they both play important roles in M. Hairer's theory of regularity structures [56]. See, for instance, references [20, 59]. The progress in these works relies essentially on describing the underlying algebraic structures in a transparent way. Indeed, the combinatorial complexity of the corresponding renormalisation process requires the introduction of group-theoretical methods such as, for instance, renormalisation group actions and comodule Hopf algebra structures [15]. Another reference of interest on generalised Wick polynomials in view of the forthcoming developments is the recent paper [82].

Starting from these remarks, in this chapter we shall discuss algebraic constructions related to moment–cumulant relations as well as Wick products, using Hopf algebra techniques. A key observation, that seems to be new in spite of being elementary and powerful, relates to the interpretation of multivariate moments of a family of random variables as a linear form on a suitable Hopf algebra. It turns out that the operation of convolution with this linear form happens to encode much of the theory of Wick products and polynomials. On the one hand, this approach enlightens the classical theory, as various structure theorems in the theory of chaos expansions follow immediately from elementary Hopf algebraic constructions, and therefore are given by the latter a group-theoretical meaning. Our methods should be compared with the combinatorial approach in [100]. On the other hand, we show a natural relation with the results and techniques that have been developed in the theory of regularity structures.

Our approach has been partially motivated by similarities with methods that have been developed for bosonic and fermionic Fock spaces by C. Brouder et al. [11, 12] to deal with interacting fields and non-trivial vacua in perturbative QFT. This is not surprising since, whereas the combinatorics of Gaussian families is reflected in the computation of averages of creation and annihilation operators over the vacuum in QFT, combinatorial properties of non-Gaussian families correspond instead to averages over non-trivial vacua.

The main idea of this chapter is that the coproduct of a bialgebra allows to deform the product and that this permits to encode interesting constructions such as generalised Wick polynomials. In the last sections of this paper, we show how the above ideas can be used in more general contexts, which include regularity structures. Regarding the latter, we mention that these ideas have been used and greatly expanded in a series of recent papers [15, 20, 56] on renormalisation of regularity structures. These papers handle products of *random* distributions which can be ill-defined and need to be *renormalised*. The procedure is rather delicate since the renormalisation, which we rather call *deformation* in this paper, must preserve other algebraic and analytical structures. Without explaining in detail the rather complex constructions appearing in [15, 20, 56], we describe how one can formalise this deformed (renormalised) product of distributions by means of a comodule structure.

4.1.1 Generalised Wick polynomials

The main results of the first part of this chapter (Theorems 4.5.1 and 4.5.4) are multivariate generalisations of the following statements for single real-valued random variable X with finite moments of all orders.

We denote by $H := \mathbb{R}[x]$ the algebra of polynomials in the variable x, endowed with the standard product

$$x^n \cdot x^m := x^{n+m}, \tag{4.1}$$

for $n, m \geq 0$. We equip H with the cocommutative coproduct $\Delta : H \to H \otimes H$ defined by

$$\Delta x^n := \sum_{k=0}^{n} \binom{n}{k} x^{n-k} \otimes x^k. \tag{4.2}$$

Product eq. (4.1) and coproduct eq. (4.2) together define a connected graded commutative bialgebra, and therefore a Hopf algebra on H. On the dual space H^* a dual product $\alpha \star \beta \in H^*$ can be defined in terms of eq. (4.2)

$$(\alpha \star \beta)(x^n) := (\alpha \otimes \beta)\Delta x^n,$$

for $\alpha, \beta \in H^*$. This product is commutative and associative, and the space $\mathscr{G}(H) := \{\lambda \in H^* : \lambda(1) = 1\}$ forms a group for this multiplication law.

Given a real-valued random variable X with all moments finite, we define the functional $\mu \in H^*$ given by $\mu(x^n) := \mu_n = \mathbb{E}(X^n)$. Then $\mu \in \mathscr{G}(H)$ and therefore its inverse μ^{-1} in $\mathscr{G}(H)$ is well defined.

Theorem 4.1.1 (Wick polynomials)**.** *We define $W := \mu^{-1} \star \mathrm{id} : H \to H$, i.e., the linear operator such that*

$$W(x^n) = (\mu^{-1} \otimes \mathrm{id})\Delta x^n = \sum_{k=0}^{n} \binom{n}{k} \mu^{-1}(x^{n-k})\, x^k. \tag{4.3}$$

Then

- *$W : H \to H$ is the only linear operator such that*

$$W(1) = 1, \qquad \frac{\mathrm{d}}{\mathrm{d}x} \circ W = W \circ \frac{\mathrm{d}}{\mathrm{d}x}, \qquad \mu(W(x^n)) = 0, \tag{4.4}$$

  for all $n \geq 1$.

- *$W : H \to H$ is the only linear operator such that for all $n \geq 0$*

$$x^n = (\mu \otimes W)\Delta x^n = \sum_{k=0}^{n} \binom{n}{k} \mu(x^{n-k})\, W(x^k).$$

We call $W(x^n) \in H$ the Wick polynomial of degree n associated to the law of X. If X is a standard Gaussian random variable then the recurrence eq. (4.4) shows that $W(x^n)$ is the Hermite polynomial H_n. Therefore eq. (4.3) gives an explicit formula for such generalised Wick polynomials in terms of the inverse μ^{-1} of the linear functional μ in the group $\mathscr{G}(H)$.

The Wick polynomial W permits to define a *deformation* of the Hopf algebra H.

Theorem 4.1.2. *The linear operator $W : H \to H$ has an inverse $W^{-1} : H \to H$ given by $W^{-1} = \mu \star \mathrm{id}$. If we define for $n, m \geq 0$ the product*

$$x^n \cdot_\mu x^m := W(W^{-1}(x^n) \cdot W^{-1}(x^m)),$$

and define similarly a twisted coproduct Δ_μ, then H endowed with $\cdot_\mu$, Δ_μ and $\varepsilon_\mu := \mu$ is a bicommutative Hopf algebra. The map W becomes an isomorphism of Hopf algebras. In *particular*

$$W(x^{n_1+\cdots+n_k}) = W(x^{n_1}) \cdot_\mu W(x^{n_2}) \cdot_\mu \cdots \cdot_\mu W(x^{n_k}),$$

for all $n_1, \ldots, n_k \in \mathbb{N}$.

Furthermore, we present Hopf-algebraic versions of classical multivariate formulae on relations between moments and cumulants.

We recall that in the case of a single random variable X with finite moments of all orders, the sequence $(\kappa_n)_{n\geq 0}$ of cumulants of X is defined by the following formal power series relation between exponential generating functions

$$\exp\left(\sum_{n\geq 0} \frac{t^n}{n!}\kappa_n\right) = \sum_{n\geq 0} \frac{t^n}{n!}\mu_n, \tag{4.5}$$

where t is a formal variable and $\mu_n = \mathbb{E}(X^n)$ is the nth-order moment of X. Note that $\mu_0 = 1$ and $\kappa_0 = 0$. Equation eq. (4.5) is equivalent to the classical recursion

$$\mu_n = \sum_{m=1}^{n} \binom{n-1}{m-1} \kappa_m \mu_{n-m}. \tag{4.6}$$

In fact, equation eq. (4.5) together with eq. (4.6) provide the definition of the classical Bell polynomials, which, in turn, are closely related to the Faà di Bruno formula [105].

Then we show multivariate generalisation of the following formulae

Theorem 4.1.3. *Setting* $\mu, \kappa \in H^*$, $\mu(x^n) := \mu_n$ *and* $\kappa(x^n) := \kappa_n$, $n \geq 0$, *we have the relations*

$$\mu = \exp^\star(\kappa) = \varepsilon + \sum_{n\geq 1} \frac{1}{n!}\kappa^{\star n}, \tag{4.7}$$

$$\kappa = \log^\star(\mu) = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n}(\mu - \varepsilon)^{\star n}, \tag{4.8}$$

where $\varepsilon(x^k) := \mathbf{1}(k = 0)$.

The above formulae eq. (4.7) and eq. (4.8) are Hopf-algebraic interpretations of the classical *Leonov–Shiryaev relations* [78], see eq. (4.10) and eq. (4.11) below.

4.1.2 Deformation of products

Theorem 4.1.2 above introduces the idea of a deformed product $\cdot_\mu$ in a polynomial algebra. This idea is used in a very important way in the recent theory of regularity structures [15, 20, 56], which is based on *products of random distributions*, i.e. of generalised functions on $\mathbb{R}^d$. Such products are in fact ill-defined and need to be *renormalised*; this operation corresponds algebraically to a deformation of the standard pointwise product, and is achieved through a comodule structure which extends the coproduct eq. (4.2) to a much larger class of generalised monomials.

In the last sections the notion of a deformed product is extended to more general comodules and we discuss one important and instructive example, the space of decorated rooted trees endowed with the extraction-contraction operator. This setting is relevant for branched rough paths [51], and constitutes a first step towards the more complex framework of regularity structures [15].

4.1.3 Overview

In Section 4.2 a brief review classical multivariate moment–cumulants relations is given. Section 4.3 provides an interpretation of these relations in a Hopf-algebraic context. In Section 4.4 the previous approach to generalised Wick polynomials is extended. Section 4.5 is devoted to Hopf algebra deformations, which are applied to Wick polynomials in Section 4.6. In Section 4.7 still another interpretation of Wick polynomials in terms of a suitable comodule structure is introduced. Section 4.8 explains the deformation of the pointwise product on functions. Section 4.9 addresses the problem of extending these results to Hopf algebras of non-planar decorated rooted trees replacing monomials.

For convenience and in view of applications to scalar real-valued random variables, we fix the field of real numbers $\mathbb{R}$ as ground field. Notice however that algebraic results and constructions in this chapter depend only on the ground field being of characteristic zero.

This chapter is based on the preprint [39].

4.2 Joint cumulants and moments

We start by briefly reviewing classical multivariate moment–cumulant relations. Given an index set A, we denote by $\mathscr{M}(A)$ the set of all finitely supported functions $\alpha \colon A \to \mathbb{N}$. For $\alpha \in \mathscr{M}(A)$ we denote by $\operatorname{supp}\alpha = \{a \in A : \alpha_a \neq 0\}$ its support. Observe that if A is itself finite, then $\mathscr{M}(A)$ coincides with the set $\mathbb{N}^A$ of all $\mathbb{N}$-valued functions on A.

The set $\mathscr{M}(A)$ is a poset under pointwise majoration, i.e. we say that $\alpha \leq \beta$ if and only if $\alpha_a \leq \beta_a$ for all $a \in A$. Moreover, it is an abelian semigroup under pointwise adition $(\alpha + \beta)_a = \alpha_a + \beta_a$. In fact, it is the free commutative semigroup generated by the indicator functions of the one-element sets $\{a\}$ for $a \in A$.

Suppose we are given a collection $\mathscr{X} = (x_a : a \in A)$ of commuting variables. We define the powers $x^0 := 1$ and

$$x^\alpha := \prod_{a \in A} x_a^{\alpha_a}.$$

We also define, for $\alpha, \beta \in \mathscr{M}(A)$ with $\beta \leq \alpha$,

$$\alpha! := \prod_{a \in A} \alpha_a!, \quad \binom{\alpha}{\beta} := \prod_{a \in A} \binom{\alpha_a}{\beta_a},$$

and more in general we let

$$\binom{\alpha}{\beta^1, \ldots, \beta^p} := \prod_{a \in A} \frac{\alpha_a!}{\beta_a^1! \cdots \beta_a^p!}$$

if $\beta^1, \ldots, \beta^p \in \mathscr{M}(A)$ are such that $\alpha = \beta^1 + \cdots + \beta^p$, and zero otherwise. Note that for a given $\alpha \in \mathscr{M}(A)$ there is only a finite number of elements in $\mathscr{M}(A)$ satisfying this condition. Finally, for $\alpha \in \mathscr{M}(A)$ we let

$$|\alpha| := \sum_{a \in A} \alpha_a < \infty.$$

4.2.1 Cumulants

If we have a finite family of random variables $(X_a \colon a \in A)$ such that X_a has finite moments of all orders for every $a \in A$, then the analogue of the exponential formula eq. (4.5) holds

$$\exp\left(\sum_{\alpha \in \mathcal{M}(A)} \frac{t^\alpha}{\alpha!} \kappa_\alpha \right) = \sum_{\alpha \in \mathcal{M}(A)} \frac{t^\alpha}{\alpha!} \mu_\alpha, \tag{4.9}$$

This defines in a unique way the family $(\kappa_\alpha, \alpha \in \mathcal{M}(A))$ of joint cumulants of $(X_a, a \in A)$ once the family of corresponding joint moments $(\mu_\alpha, \in \mathcal{M}(A))$ is given. When it is necessary to specify the dependence of κ_α on $(X_a \colon a \in A)$ we shall write $\kappa_\alpha(X)$, and similarly for μ_α.

Identifying a subset $B \subseteq A$ with its indicator function $\mathbf{1}_B \in \{0,1\}^A \subset \mathscr{M}(A)$, we can use the notation κ_B and μ_B for the corresponding joint cumulants and moments. The families $(\kappa_B, B \subseteq A)$ and $(\mu_B, B \subseteq A)$ satisfy the so-called *Leonov–Shiryaev relations* [78, 110]

$$\mu_B = \sum_{\pi \in \mathscr{P}(B)} \prod_{C \in \pi} \kappa_C \tag{4.10}$$

$$\kappa_B = \sum_{\pi \in \mathscr{P}(B)} (|\pi| - 1)! \, (-1)^{|\pi|-1} \prod_{C \in \pi} \mu_C, \tag{4.11}$$

where we write $\mathscr{P}(B)$ for the set of all set partitions of B, namely, all collections π of subsets (blocks) of B such that $\cup_{C \in \pi} C = B$ and elements of π are pairwise disjoint; moreover $|\pi|$ denotes the number of blocks of π, which is finite since B is finite. Formulae (4.10) and (4.11) have been intensively studied from a combinatorial perspective, see, e.g., [100, Chapter 2]. Regarding the properties of cumulants we refer the reader to [110].

Formula (4.10) has in fact been adopted, for instance, in [59] as a recursive definition for $(\kappa_B, B \subseteq A)$. This approach does indeed determine the cumulants uniquely by induction over the cardinality $|B|$ of the finite set B. This follows from the right-hand side containing κ_B, which is what we want to define, as well as κ_C for some C with $|C| < |B|$, which have been already defined in lower order.

Although this recursive approach seems less general than the one via exponential generating functions as in eq. (4.9), since it forces to consider only $\alpha \in \{0,1\}^A$, it turns out that they are equivalent. Indeed, replacing $(X_a : a \in A)$ with $(Y_b : b \in A \times \mathbb{N})$, where $Y_b := X_a$ for $b = (a,k) \in A \times \mathbb{N}$, then for $\alpha \in \mathscr{M}(A)$ we have

$$\kappa_\alpha(X) = \kappa_B(Y), \qquad B = \{(a,k) : a \in A,\ 1 \le k \le \alpha_a\}.$$

Now we show that the Leonov–Shiryaev relations eqs. (4.10) and (4.11) have an elegant Hopf-algebraic interpretation which also extends to Wick polynomials. Notice that a different algebraic interpretation of eqs. (4.10) and (4.11) has been given in terms of Möbius calculus [100, 110]. Moreover, the idea of writing moment–cumulant relations in terms of convolution products is closely related to Rota's Umbral calculus [71, 106].

4.3 From cumulants to Hopf algebras

In this section we explain how classical moment–cumulant relations can be encoded using Hopf algebra techniques. These results may be folklore among followers of Rota's combinatorial approach to probability, and, as we already alluded at, there exist actually in the literature already various other algebraic descriptions of moment–cumulant relations (via generating series as well as more sophisticated approaches in terms of umbral calculus, tensor algebras and set partitions). Our approach is most suited regarding our later applications, i.e., the Hopf algebraic study of Wick products. Since these ideas do not seem to be well-known to probabilists, we believe that they deserve a detailed presentation.

4.3.1 Moment–cumulant relations via multisets

Throughout the chapter we consider a fixed collection of real-valued random variables $(X_a : a \in A)$ defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for an index set A. We suppose that X_a has finite moments of all orders for every $a \in A$. Consider a collection $\mathcal{X} = (x_a : a \in A)$ of variables and let $H := \mathbb{R}[\mathcal{X}]$ be the commutative polynomial Hopf algebra described in Section 2.2

Definition 4.3.1. *For every $\alpha \in \mathcal{M}(A)$, we define the cumulant $\mathbf{E}_c(X_\alpha)$ inductively over $|\alpha|$ by $\mathbf{E}_c(X_0) = 0$ and else*

$$\mathbb{E}(X^\alpha) = \sum_{n=1}^{|\alpha|} \frac{1}{n!} \sum_{\beta^1,\ldots,\beta^n \in \mathcal{M}(A)} \binom{\alpha}{\beta^1,\ldots,\beta^n} \prod_{i=1}^{n} \mathbf{E}_c(X_{\beta^i}). \tag{4.12}$$

Remark 4.3.2. If $\alpha \in \mathcal{M}(A) \cap \{0,1\}^A$, then eq. (4.12) reduces to the first Leonov–Shiryaev relation eq. (4.10), since on the right-hand side of eq. (4.12) $\beta^1,\ldots,\beta^n \in \mathcal{M}(A)$ are also in $\{0,1\}^A$ and in particular the binomial coefficient (when non-zero) is equal to 1. △

As we will show in eq. (4.14) below, expression (4.12) is equivalent to the usual formal power series definition of cumulants (whose exponential generating series is the logarithm of the exponential generating series of moments). As for eq. (4.10), expression (4.12) does indeed determine the cumulants uniquely by induction over $|\alpha|$. This is because the right-hand side only involves $\mathbf{E}_c(X_\alpha)$, which is what we want to define, as well as $\mathbf{E}_c(X_{\bar\alpha})$ for some $\bar\alpha$ with $|\bar\alpha| < |\alpha|$, which is already defined by the inductive hypothesis. Define two linear functionals on H:

$$\begin{aligned} \mu :\ & H \to \mathbb{R} & \qquad \kappa :\ & H \to \mathbb{R} \\ & x^\alpha \mapsto \mu(x^a) := \mathbb{E}(X^\alpha) & & x^\alpha \mapsto \kappa(x^\alpha) := \mathbf{E}_c(X_\alpha), \end{aligned} \tag{4.13}$$

with $\mu(1) := 1$ and $\kappa(1) := 0$.

4.3.2 Exponential generating functions

Let us fix a finite subset $S = \{a_1,\ldots,a_p\} \subset A$. For $\beta \in \mathcal{M}(S)$ we set

$$t^\beta := \prod_{i=1}^{p} (t_i)^{\beta_{a_i}},$$

where $t = (t_i : i = 1,\ldots,p)$ are commuting variables. Then we define the exponential generating function of $\lambda \in H^*$ as the formal series

$$\phi_\lambda(t,S) := \sum_{\beta \in \mathcal{M}(S)} \frac{t^\beta}{\beta!} \lambda(\beta) \in \mathbb{R}[[t]].$$

Then from definition 4.3.1 we get the usual exponential relation between the exponential moment and cumulant generating functions of μ and κ, analogous to eqs. (4.5) and (4.9):

$$\begin{aligned}\phi_\mu(t,S) &= \sum_{\beta\in\mathcal{M}(S)} \frac{t^\beta}{\beta!}\mu(\beta)\\ &= \sum_{n\geq 0}\frac{1}{n!}\sum_{\beta\in\mathcal{M}(S)}\sum_{\beta^1,\dots,\beta^n\in\mathcal{M}(S)}\frac{1}{\beta!}\binom{\beta}{\beta^1,\dots,\beta^n}\prod_{i=1}^n\left(t^{\beta^i}\kappa(\beta^i)\right)\\ &= \sum_{n\geq 0}\frac{1}{n!}\sum_{\beta^1,\dots,\beta^n\in\mathcal{M}(S)}\prod_{i=1}^n\left(\frac{t^{\beta^i}}{\beta^i!}\kappa(\beta^i)\right)\\ &= \sum_{n\geq 0}\frac{1}{n!}\left(\sum_{\beta\in\mathcal{M}(S)}\frac{t^\beta}{\beta!}\kappa(\beta)\right)^n = \exp(\phi_\kappa(t,S)).\end{aligned} \tag{4.14}$$

From eq. (4.14) we obtain another recursive relation between moments and cumulants. We have

$$\mu(x^\alpha\cdot x_a) = \sum_{\beta^1,\beta^2\in\mathcal{M}(A)}\binom{\alpha}{\beta^1,\beta^2}\kappa(x^{\beta^1}\cdot x_a)\mu(\beta^2). \tag{4.15}$$

This recursion is the multivariate analogue of the one in eq. (4.6).

4.3.3 Moment–cumulant relations and Hopf algebras

Recall that H is a graded connected Hopf algebra with product induced by the semigroup structure on $\mathcal{M}(A)$ by $x^\alpha\cdot x^\beta = x^{\alpha+\beta}$. The coproduct $\Delta : H\to H\otimes H$ can be equally described as

$$\Delta x^\alpha = \sum_{\beta+\beta'=\alpha}\binom{\alpha}{\beta,\beta'}x^\beta\otimes x^{\beta'}. \tag{4.16}$$

It is the unique algebra map such that the variables x_a are primitive. The counit is defined by $\langle\varepsilon, x^\alpha\rangle = 1_{\alpha=0}$.

We denote $\Delta^0 := \mathrm{id} : H\to H$, $\Delta^1 := \Delta : H\to H\otimes H$, and for $n\geq 2$:

$$\Delta^n := (\Delta\otimes\mathrm{id})\Delta^{n-1} : H\to H^{\otimes(n+1)}.$$

Proposition 4.3.3. *We have*

$$\mu = \exp^\star(\kappa) = \varepsilon + \sum_{n\geq 1}\frac{1}{n!}\kappa^{\star n}. \tag{4.17}$$

Proof. By the definitions of Δ and Δ^n we find that

$$\kappa^{\otimes n}\Delta^{n-1}A = \sum_{B_1,\ldots,B_n\in\mathcal{M}(\mathcal{A})} \binom{A}{\beta_1\ldots\beta_n} \prod_{i=1}^{n} \kappa(\beta_i),$$

and, by eq. (4.12), this yields the result. □

Remark 4.3.4. Formula eq. (4.17) is the Hopf-algebraic analogue of the first Leonov–Shiryaev relation eq. (4.10). △

Since $\kappa(\emptyset) = 0$, $\kappa^{\otimes n}\Delta^{n-1}(A)$ vanishes whenever $|A| > n$. Similarly, under the same assumption, $(\mu-\varepsilon)^{\star n}(A) = 0$. It follows that one can handle formal series identities such as $\log^{\star}(\exp^{\star}(\kappa)) = \kappa$ or $\exp^{\star}(\log^{\star}(\mu)) = \mu$ without facing convergence issues. In particular

Proposition 4.3.5. *We have*

$$\kappa = \log^{\star}(\mu) = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} (\mu-\varepsilon)^{\star n}. \tag{4.18}$$

From Proposition 4.3.5 we obtain the formula

$$\mathbf{E}_c(X_\alpha) = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} \sum_{\beta_1,\ldots,\beta_n\in\mathcal{M}(A)} \binom{\alpha}{\beta_1\ldots\beta_n} \prod_{i=1}^{n} \mathbb{E}(X^{\beta_i}) \tag{4.19}$$

which may be considered the inverse to eq. (4.12).

Remark 4.3.6. The formula (4.18) is the Hopf-algebraic analogue of the second Leonov–Shiryaev relation eq. (4.11). Moreover, for $\alpha \in \mathcal{M}(A)\cap\{0,1\}^A$, then eq. (4.19) also reduces to the second Leonov–Shiryaev relation eq. (4.11), since on the right-hand side of eq. (4.19) $\beta_1,\ldots,\beta_n \in \mathcal{M}(A)$ are also in $\{0,1\}^A$. △

4.3.4 A sub-coalgebra

If one prefers to work in the combinatorial framework of the Leonov–Shiryaev formulae eq. (4.10)-eq. (4.11) rather than with eq. (4.12)-eq. (4.19), then one may consider the linear span J of the set $\{x^\alpha : \alpha \in \mathcal{M}(A)\cap\{0,1\}^A\}$, namely all monomials where the partial degree of each variable is at most one.

Then J is a linear subspace of H, which is *not* a sub-algebra of H for the product, since for example any single variable $x_a \in J$ but $x_a \cdot x_a = x_a^2 \notin J$. The coproduct Δ defined in eq. (4.16) coacts however nicely on J since if $x^\alpha \in J$ then

$$\Delta x^\alpha = \sum_{\beta_1+\beta_2=\alpha} \beta_1\otimes\beta_2 \in J\otimes J.$$

Moreover the restriction of ε to J defines a counit for (J, Δ). Therefore J is a sub-coalgebra of H. With a slight abuse of notation we still write $\star$ for the dual product on J^*

$$\langle f \star g, x^\alpha \rangle := \langle f \otimes g, \Delta x^\alpha \rangle,$$

for $x^\alpha \in J$ and $f, g \in J^*$. If we denote as before

$$\begin{aligned} \mu :\ & J \longrightarrow \mathbb{R} \\ & x^\alpha \mapsto \mu(x^\alpha) := \mathbb{E}(X^\alpha) \end{aligned} \qquad \begin{aligned} \kappa :\ & J \longrightarrow \mathbb{R} \\ & x^\alpha \mapsto \kappa(x^\alpha) := \mathbf{E}_c(X_\alpha), \end{aligned}$$

with $\mu(\emptyset) := 1$ and $\kappa(\emptyset) := 0$, then the Leonov–Shiryaev relations eq. (4.10)-eq. (4.11) can be rewritten in J^* as, respectively,

$$\mu = \exp^\star(\kappa) = \varepsilon + \sum_{n \geq 1} \frac{1}{n!} \kappa^{\star n}$$

and

$$\kappa = \log^\star(\mu) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} (\mu - \varepsilon)^{\star n}.$$

4.4 Wick Products

The theory of Wick products, as well as the related notion of chaos decomposition, play an important role in various fields of applied probability. Both have deep structural features in relation to the fine structure of the algebra of square integrable functions associated to one or several random variables. The aim of this section and the following ones is to revisit the theory on Hopf algebraic grounds. The basic observation is that the formula for the Wick product is closely related to the recursive definition of antipode in a connected graded Hopf algebra. This approach seems to be new, also from the point of view of concurring approaches such as umbral calculus [71, 106] or set partition combinatorics *à la* Rota–Wallstrom [107].

4.4.1 Wick polynomials

We are going to use extensively the notion of Wick polynomials for a collection of (not necessarily Gaussian) random variables which is defined as follows.

Definition 4.4.1. *Given a collection $(X_a : a \in A)$ of random variables with finite moments of all orders, for any $\alpha \in \mathcal{M}(A)$ the Wick polynomial $:X_\alpha:$ is a random variable defined recursively* *by setting* $:X_0: \; = 1$ *and postulating that*

$$X^\alpha = \sum_{\beta_1,\beta_2 \in \mathcal{M}(A)} \binom{\alpha}{\beta_1\,\beta_2} \mathbb{E}(X^{\beta_1}) :X_{\beta_2}: . \tag{4.20}$$

As for cumulants, eq. (4.20) is sufficient to define $:X_\alpha:$ by recursion over $|\alpha|$. Indeed, the term with $\beta_2 = \alpha$ is precisely the quantity we want to define, and all other terms only involve Wick polynomials $:X_\beta:$, for $\beta \in \mathcal{M}(A)$ with $|\beta| < |\alpha|$.

It is now clear that formula eq. (4.20) can be lifted to H as

$$x^\alpha = \sum_{\beta_1,\beta_2 \in \mathcal{M}(\mathcal{A})} \binom{\alpha}{\beta_1\,\beta_2} \mathbb{E}(X^{\beta_1}) :x_{\beta_2}: , \tag{4.21}$$

and written in Hopf algebraic terms as follows

$$x^\alpha = (\mu \star W)(x^\alpha) = (\mu \otimes W)\Delta x^\alpha, \tag{4.22}$$

for $A \in \mathcal{M}(\mathcal{A})$. We have set $W : H \to H$, $W(x^\alpha) := \; :x_\alpha:$ and call W the Wick product map (see Theorem 4.5.4 for a justification of the terminology). Notice that it depends on the joint distribution of the X_as. Formula 4.22 is the Hopf algebraic analogue of the definition of the Wick polynomial $:X_B:$ used in references [59, 82]. Moreover, introducing the algebra map $\mathrm{ev} : x^\alpha \longmapsto X^\alpha$ from H to the algebra of random variables generated by $(X_a : a \in A)$, one gets by a recursion over $|\alpha|$ that $\mathrm{ev}(:x^\alpha:) = \; :X_\alpha:$ (for that reason, from now on we will call slightly abusively both $:x^\alpha:$ and the random variable $:X_\alpha:$ the Wick polynomial associated to x^α).

4.4.2 A Hopf algebraic construction

We want to present now a closed Hopf algebraic formula for the Wick polynomials introduced in Definition 4.4.1. We define the set $\mathcal{G}(H) := \{\lambda \in H^* : \lambda(1) = 1\}$. Then it is well known that $\mathcal{G}(H)$ is a group for the $\star$-product. Indeed, any $\lambda \in \mathcal{G}(H)$ has an inverse λ^{-1} in $\mathcal{G}(H)$ given by

$$\lambda^{-1} = \sum_{n \geq 0} (\varepsilon - \lambda)^{\star n}. \tag{4.23}$$

As usual, this infinite sum defines an element of H^* since, evaluated on any monomial $x^\alpha \in H$, it reduces to a finite number of terms.

Theorem 4.4.2. *Let* $\mu \in \mathcal{G}(H)$ *be given by* $\mu(x^\alpha) = \mathbb{E}(X^\alpha)$ *and* $W(x^\alpha) = \; :x^\alpha:$ *, then for all* $\alpha \in \mathcal{M}(A)$

$$:x^\alpha: \; = (\mu^{-1} \star \mathrm{id})(x^\alpha) = (\mu^{-1} \otimes \mathrm{id})\Delta x^\alpha. \tag{4.24}$$

Proof. The identity follows from eq. (4.22) and from the associativity of the $\star$ product. $\square$

From eq. (4.23) and eq. (4.24) we obtain

Proposition 4.4.3. *Wick polynomials have the explicit expansion*

$$:x^\alpha: \; = x^\alpha + \sum_{n\geq 1}(-1)^n \sum_{\beta\in\mathcal{M}(A)} \sum_{\substack{\beta_1,\dots,\beta_n\in\mathcal{M}(A)\\ \beta_i\neq 0}} \binom{\alpha}{\beta_1,\dots,\beta_n,\beta} \mu(x^{\beta_1})\cdots\mu(x^{\beta_n})x^\beta.$$

4.5 Hopf algebra deformations

The group $\mathcal{G}(H) = \{\lambda \in H^* : \lambda(\mathbf{1}) = 1\}$ equipped with the $\star$ product acts canonically on H by means of the map $\varphi_\lambda : H \to H$

$$\varphi_\lambda(x^\alpha) := (\lambda \otimes \mathrm{id})\Delta x^\alpha, \tag{4.25}$$

for $\lambda \in \mathcal{G}(H)$ and a monomial $x^\alpha \in H$. In other words, $\varphi_\lambda = \lambda \star \mathrm{id} = \mathrm{id} \star \lambda$, since Δ is cocommutative. This is a group action since one checks easily using the coassociativity of Δ that

$$\varphi_{\lambda_1\star\lambda_2} = \varphi_{\lambda_1} \circ \varphi_{\lambda_2},$$

so that in particular

$$(\varphi_\lambda)^{-1} = \varphi_{\lambda^{-1}}.$$

These maps φ_λ, being invertible, allow to define *deformations* of the standard product $\cdot$ in H, as well as of the coproduct Δ defined in eq. (4.16) and of the counit. Namely we define $\cdot_\lambda : H \otimes H \to H$, $\Delta_\lambda : H \to H \otimes H$ and ε_λ by

$$\begin{aligned} x^\alpha \cdot_\lambda x^\beta &:= \varphi_\lambda^{-1}(\varphi_\lambda(x^\alpha) \cdot \varphi_\lambda(x^\beta)), \\ \Delta_\lambda x^\alpha &:= (\varphi_\lambda^{-1} \otimes \varphi_\lambda^{-1})\Delta\varphi_\lambda x^\alpha, \\ \langle \varepsilon_\lambda, x^\alpha\rangle &:= \langle \varepsilon, \varphi_\lambda(x^\alpha)\rangle = \langle \lambda, x^\alpha\rangle. \end{aligned} \tag{4.26}$$

Although $\varepsilon_\lambda = \lambda$, we find it useful to introduce the notation ε_λ to feature the new role of λ as a counit.

Notice that, as $\langle \lambda, \mathbf{1}\rangle = 1$, we have $\varphi_\lambda(\mathbf{1}) = \mathbf{1}$ and $x^\alpha \cdot_\lambda \mathbf{1} = x^\alpha$. Dually,

$$(\varepsilon_\lambda \otimes \mathrm{id}) \circ \Delta_\lambda(x^\alpha) = (\varepsilon \otimes \varphi_{\lambda^{-1}})\Delta\varphi_\lambda(x^\alpha) = x^\alpha.$$

Then we have

Theorem 4.5.1. *For any $\lambda \in \mathscr{G}(H)$, the quintuple $(H, \cdot_\lambda, \mathbf{1}, \Delta_\lambda, \varepsilon_\lambda)$ defines a Hopf algebra. The map*

$$\varphi_\lambda^{-1} : (H, \cdot, \mathbf{1}, \Delta, \varepsilon) \to (H, \cdot_\lambda, \mathbf{1}, \Delta_\lambda, \varepsilon_\lambda)$$

is an isomorphism of Hopf algebras.

Proof. Although the Theorem follows directly from the properties of conjugacy, we detail the proof. Associativity of $\cdot_\lambda$ and coassociativity of Δ_λ follow directly. First,

$$(x^\alpha \cdot_\lambda x^\beta) \cdot_\lambda x^\gamma = \varphi_\lambda^{-1}(\varphi_\lambda(x^\alpha) \cdot \varphi_\lambda(x^\beta) \cdot \varphi_\lambda(x^\gamma)) = x^\alpha \cdot_\lambda (x^\beta \cdot_\lambda x^\gamma),$$

which shows associativity. Coassociativity is simple to see as well

$$\begin{aligned}(\Delta_\lambda \otimes \mathrm{id})\Delta_\lambda x^\alpha &= (\varphi_\lambda^{-1} \otimes \varphi_\lambda^{-1} \otimes \varphi_\lambda^{-1})(\Delta \otimes \mathrm{id})\Delta\varphi_\lambda x^\alpha \\ &= (\varphi_\lambda^{-1} \otimes \varphi_\lambda^{-1} \otimes \varphi_\lambda^{-1})(\mathrm{id} \otimes \Delta)\Delta\varphi_\lambda x^\alpha \\ &= (\mathrm{id} \otimes \Delta_\lambda)\Delta_\lambda x^\alpha.\end{aligned}$$

We check now the compatibility relation between $\cdot_\lambda$ and Δ_λ:

$$\begin{aligned}(\Delta_\lambda x^\alpha) \cdot_\lambda (\Delta_\lambda \beta) &= (\varphi_\lambda^{-1} \otimes \varphi_\lambda^{-1})\Big(\big((\varphi_\lambda \otimes \varphi_\lambda)\Delta_\lambda x^\alpha\big) \cdot \big((\varphi_\lambda \otimes \varphi_\lambda)\Delta_\lambda x^\beta\big)\Big) \\ &= (\varphi_\lambda^{-1} \otimes \varphi_\lambda^{-1})\big((\Delta\varphi_\lambda x^\alpha) \cdot (\Delta\varphi_\lambda x^\beta)\big) \\ &= (\varphi_\lambda^{-1} \otimes \varphi_\lambda^{-1})\Delta\Big(\varphi_\lambda x^\alpha \cdot \varphi_\lambda x^\beta\Big) \\ &= \Delta_\lambda(x^\alpha \cdot_\lambda x^\beta).\end{aligned}$$

Finally, we check that $\varphi_\lambda^{-1} : (H, \cdot, \mathbf{1}, \Delta, \varepsilon) \to (H, \cdot_\lambda, \mathbf{1}, \Delta_\lambda, \varepsilon_\lambda)$ is a bialgebra morphism:

$$\varphi_\lambda^{-1}(x^\alpha \cdot x^\beta) = \varphi_\lambda^{-1}(x^\alpha) \cdot_\lambda \varphi_\lambda^{-1}(x^\beta),$$

$$(\varphi_\lambda^{-1} \otimes \varphi_\lambda^{-1})\Delta x^\alpha = \Delta_\lambda \varphi_\lambda^{-1} x^\alpha.$$

We have proved until now that φ_λ^{-1} is a isomorphism of bialgebras. Since $(H, \cdot_\lambda, \mathbf{1}, \Delta_\lambda, \varepsilon_\lambda)$ is a graded connected bialgebra, it has an antipode. Since moreover the antipode of a Hopf algebra is unique, we obtain that φ_λ^{-1} preserves the antipode as well. □

Remark 4.5.2. The construction of eq. (4.26) and Theorem 4.5.1 works also if we replace φ_λ with any linear invertible map $\varphi : H \to H$ such that $\varphi(\mathbf{1}) = \mathbf{1}$. Indeed, in the above considerations we have never used the formula eq. (4.25) which defines φ_λ. △

Remark 4.5.3. There is nothing particular about the structure of H used in the proof of Theorem 4.5.1, so the same proof works more generally for any Hopf algebra H and linear invertible map $\varphi : H \to H$ preserving the unit. △

In the particular case of $\lambda = \mu$, where μ is the moment functional defined in eq. (4.13), we obtain by Theorem 4.4.2 and Theorem 4.5.1:

Theorem 4.5.4. *The Wick product map* $W(x^\alpha) = \,:x^\alpha:$ *is equal to* $\varphi_{\mu^{-1}}$. *Therefore* $W : (H, \cdot, \mathbf{1}, \Delta, \varepsilon) \to (H, \cdot_\mu, \mathbf{1}, \Delta_\mu, \varepsilon_\mu)$ *is a Hopf algebra isomorphism, in particular*

$$:x^\alpha \cdot x^\beta: \; = \; :x^\alpha: \;\cdot_\mu\; :x^\beta: \,,$$

for all monomials $x^\alpha, x^\beta \in H$.

More generally, we obtain for any $\alpha_1, \ldots, \alpha_n \in \mathcal{M}(A)$ that

$$:x^{\alpha_1} \cdots x^{\alpha_n}: \; = \; :x^{\alpha_1}: \;\cdot_\mu \cdots \cdot_\mu\; :x^{\alpha_n}: \,. \tag{4.27}$$

We notice at last an interesting additional result expressing abstractly compatibility relations between the two Hopf algebra structures on H (see also Proposition 4.7.3 below). We recall that a linear space M is a left comodule over the coalgebra (H, Δ, ε) if there is linear map $\rho : M \to H \otimes M$ such that

$$(\Delta \otimes \mathrm{id}_M)\rho = (\mathrm{id}_H \otimes \rho)\rho, \qquad (\varepsilon \otimes \mathrm{id}_M)\rho = \mathrm{id}_M \,. \tag{4.28}$$

A left comodule endomorphism of M is then a linear map $f : M \to M$ such that

$$\rho \circ f = (\mathrm{id}_H \otimes f)\rho.$$

In particular the coalgebra (H, Δ, ε) is a left comodule over itself, with $\rho = \Delta$.

Proposition 4.5.5. *If we consider* H *as a left comodule over itself, then* φ_λ *is a left comodule morphism for all linear* $\lambda : H \to \mathbb{R}$, *namely*

$$\Delta\varphi_\lambda = (\mathrm{id} \otimes \varphi_\lambda)\Delta. \tag{4.29}$$

In particular the Wick product map W *is a left comodule endomorphism of* (H, Δ, ε).

Proof. We have

$$\begin{aligned}
\Delta\varphi_\lambda &= (\lambda \otimes \mathrm{id} \otimes \mathrm{id})(\mathrm{id} \otimes \Delta)\Delta \\
&= (\lambda \otimes \mathrm{id} \otimes \mathrm{id})(\Delta \otimes \mathrm{id})\Delta \\
&= (\mathrm{id} \otimes \lambda \otimes \mathrm{id})(\Delta \otimes \mathrm{id})\Delta \\
&= (\mathrm{id} \otimes \lambda \otimes \mathrm{id})(\mathrm{id} \otimes \Delta)\Delta \\
&= (\mathrm{id} \otimes \varphi_\lambda)\Delta,
\end{aligned}$$

where we have used, in this order, coassociativity, cocommutativity and then coassociativity again. $\square$

4.6 Wick products as Hopf algebra deformations

Let $a \in A$. We define now the functional $\zeta_a : H \to \mathbb{R}$ given by $\langle \zeta_a, x^\alpha \rangle = 1$ if and only if $x^\alpha = x_a$. Then we define the operator $\partial_a : H \to H$ as $\partial_a := \zeta_a \star \mathrm{id} = \varphi_{\zeta_a}$ in the notation eq. (4.25), namely

$$\partial_a x^\alpha = (\zeta_a \otimes \mathrm{id})\Delta x^\alpha.$$

It is simple to see that ∂_a acts as a formal partial derivation with respect to x_a, namely it satisfies for $\alpha, \beta \in \mathcal{M}(A)$ and $a, b \in A$

$$\partial_a x_b = 1_{(a=b)}\mathbf{1}, \qquad \partial_a(x^\alpha \cdot x^\beta) = \partial_a(x^\alpha) \cdot x^\beta + x^\alpha \cdot \partial_a(x^\beta),$$

since ζ_a satisfies $\zeta_a(\mathbf{1}) = 0$ and $\zeta_a(x^\alpha \cdot x^\beta) = \zeta_a(x^\alpha)\langle \varepsilon, x^\beta \rangle + \langle \varepsilon, x^\alpha \rangle \zeta_a(x^\beta)$, namely ζ_a is an infinitesimal character.

Then the following result is a reformulation in our setting of [82, Proposition 3.4].

Theorem 4.6.1. *The family of polynomials* $(:x^\alpha:, \alpha \in \mathcal{M}(A))$ *is the only collection such that* $:\mathbf{1}: = \mathbf{1}$ *and for all non-null* $\alpha \in \mathcal{M}(A)$ *and* $a \in \mathcal{A}$

$$\partial_a :x^\alpha: = :\partial_a x^\alpha: \quad \text{and} \quad \langle \mu, :x^\alpha: \rangle = 0. \tag{4.30}$$

Proof. Since $\mu \in \mathcal{G}(H)$ equation eq. (4.24) implies

$$\langle \mu, :x^\alpha: \rangle = \langle \mu^{-1} \star \mu, x^\alpha \rangle = \langle \varepsilon, x^\alpha \rangle.$$

Using eq. (4.29) for $\lambda = \zeta_a$ we obtain

$$\Delta \partial_a = (\mathrm{id} \otimes \partial_a)\Delta.$$

We conclude from eq. (4.24) that

$$\partial_a :x^\alpha: = (\zeta_a \star \mu^{-1} \star \mathrm{id})(x^\alpha) = (\mu^{-1} \star \zeta_a \star \mathrm{id})(x^\alpha) = :\partial_a x^\alpha:$$

by the associativity and commutativity of $\star$. Therefore $:x^\alpha:$ satisfies eq. (4.30). The converse follows from the fact that eq. (4.30) defines by recurrence a unique family. $\square$

4.6.1 Back to simple subsets

As in Subsection 4.3.4, we can restrict the whole discussion to Wick polynomials associated to finite sets $\alpha \in \mathcal{M}(A) \cap \{0,1\}^A$ and their linear span J. Indeed, if $\alpha \in J$ then $:x^\alpha: \; = W(x^\alpha)$ also belongs to J and is defined by the recursion

$$x^\alpha = \sum_{\beta_1+\beta_2=\alpha} \mathbb{E}(X^{\beta_1}) :X_{\beta_2}: \, .$$

As in Theorem 4.4.2, we have $W = \mu^{-1} \star \mathrm{id}$ and $\mathrm{id} = \mu \star W$, and, as in Proposition 4.4.3,

$$\begin{aligned} :x^\alpha: \; &= x^\alpha + \\ &+ \sum_{n\geq 1} (-1)^n \sum_{\beta \in \mathcal{M}(A)} \sum_{\beta_1,\dots,\beta_n \in \mathcal{M}(A)\setminus\{0\}} \mathbb{1}_{(\beta++\beta_1+\cdots+\beta_n=\alpha)} \mu(x^{\beta_1}) \cdots \mu(x^{\beta_n}) x^\beta \end{aligned}$$

for all $\alpha \in \mathcal{M}(A) \cap \{0,1\}^A$.

However, as we have sees in Section 4.5 above, it is more interesting to work on the bialgebra H than on the coalgebra J, see in particular Theorem 4.5.4.

4.7 On the inverse of unital functionals

As we have seen in Theorem 4.4.2, the element $\mu^{-1} \in \mathcal{G}(H)$ plays an important role in the Hopf algebraic representation eq. (4.24) of Wick products. From eq. (4.23) we obtain a general way to compute μ^{-1}. We discuss now another way to represent μ^{-1} by means of a comodule structure, which is directly inspired by [15, 20], see Section 4.10.

Let us consider now a linear functional $\lambda : H \to \mathbb{R}$ which is also an unital algebra morphism (or *character*), namely such that $\lambda(\varnothing) = 1$ and $\lambda(x^{\alpha+\beta}) = \lambda(x^\alpha)\lambda(x^\beta)$ for all $\alpha, \beta \in \mathcal{M}(A)$. Then we have a simpler way to compute its inverse: namely as $\lambda^{-1} = \lambda \circ S$, where $S : H \to H$ is the *antipode*, i.e. the only linear map such that

$$S \star \mathrm{id} = \mathrm{id} \star S = \mathbf{1}\varepsilon,$$

where $\mathbf{1}$ is the unit and ε the counit of H.

However, this does not seem to help in the computation of μ, since moments are notoriously not multiplicative in general. This problem is circumvented by extending μ to a character $\hat{\mu}$ defined on a larger Hopf algebra $\hat{H}$ which is constructed from H, such that the inverse $\hat{\mu}^{-1}$ will be computed via the antipode of $\hat{H}$.

Definition 4.7.1. *Let $\hat{H}$ be the free commutative unital algebra (the algebra of polynomials)* *generated by H. We denote by $\bullet$ the product in $\hat{H}$ and we define the coproduct $\hat{\Delta} : \hat{H} \to \hat{H} \otimes \hat{H}$ given by $\hat{\Delta}(\iota x^\alpha) = (\iota \otimes \iota)\Delta x^\alpha$ and*

$$\hat{\Delta}(x^{\alpha_1} \bullet x^{\alpha_2} \bullet \cdots \bullet x^{\alpha_n}) = (\hat{\Delta} x^{\alpha_1}) \bullet (\hat{\Delta} x^{\alpha_2}) \bullet \cdots \bullet (\hat{\Delta} x^{\alpha_n}),$$

where $\iota : H \to \hat{H}$ is the canonical injection (which we will omit whenever this does not cause confusion). The unit of $\hat{H}$ is $\mathbf{1}$ and the counit is defined by $\hat{\varepsilon}(x^{\alpha_1} \bullet x^{\alpha_2} \bullet \cdots \bullet x^{\alpha_n}) = \varepsilon(x^{\alpha_1}) \cdots \varepsilon(x^{\alpha_n})$.

Since $\hat{H}$ is a polynomial algebra, $\hat{\Delta}$ is well-defined by specifying its action on the elements of $\mathcal{M}(A)$, and requiring it to be multiplicative. It turns the space $\hat{H}$ into a connected graded Hopf algebra, where the grading is

$$|x^{\alpha_1} \bullet x^{\alpha_2} \bullet \cdots \bullet x^{\alpha_n}| := |\alpha_1| + |\alpha_2| + \cdots + |\alpha_n|.$$

The antipode $\hat{S} : \hat{H} \to \hat{H}$ of $\hat{H}$ can be computed by recurrence with the classical formula

$$\hat{S} x^\alpha = -x^\alpha - \sum_{\beta_1, \beta_2 \in \mathcal{M}(A) \setminus \{0\}} \binom{\alpha}{\beta_1\ \beta_2} \big[\hat{S} x^{\beta_1}\big] \bullet x^{\beta_2},$$

where we dropped the injection ι for notational convenience. A closed formula for $\hat{S}$ follows

$$\hat{S} x^\alpha = -x^\alpha + \sum_{n \geq 2} (-1)^n \sum_{\beta_1, \ldots, \beta_n \notin \{0\}} \binom{\alpha}{\beta_1, \ldots, \beta_n} \beta_1 \bullet \beta_2 \bullet \cdots \bullet \beta_n. \tag{4.31}$$

We denote by $\mathcal{C}(\hat{H})$ the set of characters on $\hat{H}$. This is a group for the $\hat{\star}$ convolution, dual to $\hat{\Delta}$.

Proposition 4.7.2. *The restriction map $R : \mathcal{C}(\hat{H}) \to \mathcal{G}(H)$, $R\hat{\lambda} := \hat{\lambda}|_H$ defines a group isomorphism.*

Proof. The map is clearly bijective, since a character on $\hat{H}$ is uniquely determined by its values on H, and every $\lambda \in \mathcal{G}(H)$ gives rise in this way to a $\hat{\lambda} \in \mathcal{C}(\hat{H})$ such that $R\hat{\lambda} = \lambda$.

It remains to show that R is a group morphism. This follows from

$$R(\hat{\phi} \,\hat{\star}\, \hat{\psi})(x^\alpha) = (\hat{\phi} \otimes \hat{\psi})\hat{\Delta} x^\alpha = (\hat{\phi}|_H \otimes \hat{\psi}|_H)\Delta x^\alpha = (R\hat{\phi}) \star (R\hat{\psi})(x^\alpha),$$

where $\hat{\phi}, \hat{\psi} \in \mathcal{C}(\hat{H})$ and $x^\alpha \in H$. □

For all $\lambda \in \mathcal{G}(H)$ we call $\hat{\lambda}$ the only character on $\hat{H}$ which is mapped to λ by the isomorphism R. By the previous proposition we obtain, in particular, that $(\hat{\lambda})^{-1}|_H = \lambda^{-1}$ for all $\lambda \in \mathcal{G}(H)$. Since $\hat{\lambda}$ is a character on $\hat{H}$, we have $(\hat{\lambda})^{-1} = \hat{\lambda} \circ \hat{S}$. Therefore

$$\lambda^{-1} = (\,\hat{\lambda} \circ \hat{S}\,)|_H. \tag{4.32}$$

This formula can be used specifically to compute the inverse μ^{-1} of the functional μ in eq. (4.24).

4.7.1 A comodule structure

The above considerations suggest that we can introduce the following additional structure: if we define $\delta : H \to \hat{H} \otimes H$, $\delta := (\iota \otimes \mathrm{id})\Delta$, where $\iota : H \to \hat{H}$ is the canonical injection of Definition 4.7.1, then H is turned into a *left comodule* over $\hat{H}$, namely we have

$$(\hat{\Delta} \otimes \mathrm{id}_H)\,\delta = (\mathrm{id}_{\hat{H}} \otimes \delta)\,\delta, \qquad \mathrm{id}_H = (\hat{\varepsilon} \otimes \mathrm{id}_H)\,\delta, \tag{4.33}$$

see eq. (4.28) above. Note that eq. (4.33) is in fact just the coassociativity of Δ on H in disguise.

Then we can rewrite the Hopf algebraic representation eq. (4.24) of Wick polynomials as follows:

$$:x^\alpha: \; = (\hat{\mu} \circ \hat{S} \otimes \mathrm{id})\,\delta x^\alpha, \tag{4.34}$$

for a monomial $x^\alpha \in H$, where $\hat{\mu}$ is the $\bullet$-multiplicative extension of μ from H to $\hat{H}$. Expanding this formula by means of the closed formula for $\hat{S}$, one recovers, by different means, Proposition 4.4.3.

From Proposition 4.5.5 above, we obtain

Proposition 4.7.3. *We define the action of $\mathscr{C}(\hat{H})$ on H by*

$$\psi_{\hat{\lambda}} : H \to H, \qquad \psi_{\hat{\lambda}}(x^\alpha) = (\hat{\lambda} \otimes \mathrm{id})\,\delta x^\alpha, \qquad \hat{\lambda} \in \mathscr{C}(\hat{H}),$$

for $x^\alpha \in H$. Then $\psi_{\hat{\lambda}}$ is comodule morphism for all $\hat{\lambda} \in \mathscr{C}(\hat{H})$, namely

$$\delta \circ \psi_{\hat{\lambda}} = (\mathrm{id}_{\hat{H}} \otimes \psi_{\hat{\lambda}})\,\delta.$$

4.8 Deformation of pointwise multiplication

We show now that the ideas of the previous sections can be generalised and used to define deformations of other products. The main example for us is the pointwise product on functions $f : \mathbb{R}^d \to \mathbb{R}$, and we explain in the next sections how these ideas appear in regularity structures.

Let us consider the vector space V freely generated by a family $T = (\tau_i,\, i \in I)$. We denote by $\mathfrak{T}$ the free commutative monoid on T, with neutral element $\varnothing \in \mathfrak{T} \setminus T$. We define also $(C, \odot, \varnothing)$ as the unital free commutative algebra generated by T; then C is the vector space freely generated by $\mathfrak{T}$ and we have a canonical embedding

$$V \hookrightarrow C.$$

We suppose that C is a left-comodule over a Hopf algebra $(\hat{C}, \bullet, \varnothing, \hat{\Delta}, \hat{\varepsilon})$, with coaction $\delta : C \to \hat{C} \otimes C$ satisfying the analogue of (4.33). We stress that the coaction δ is not supposed to be multiplicative with respect to the $\odot$ product in C.

We say that $\lambda : \hat{C} \to \mathbb{R}$ is *unital* if it is a linear functional such that $\lambda(\varnothing) = 1$. Then we define, as in the previous section,

$$\psi_\lambda : C \to C, \qquad \psi_\lambda := (\lambda \otimes \mathrm{id})\delta, \tag{4.35}$$

for every unital $\lambda : \hat{C} \to \mathbb{R}$. It is easy to see that

$$\psi_\lambda \circ \psi_{\lambda'} = \psi_{\lambda \,\hat{\star}\, \lambda'},$$

where $\hat{\star}$ is the convolution product with respect to the coproduct $\hat{\Delta} : \hat{C} \to \hat{C} \otimes \hat{C}$. We then define the product $\odot_\lambda$ on C as

$$\alpha \odot_\lambda \beta := \psi_\lambda^{-1}[(\psi_\lambda \alpha) \odot (\psi_\lambda \beta)], \qquad \alpha, \beta \in C, \tag{4.36}$$

where $\psi_\lambda^{-1} = \psi_{\lambda^{-1}}$ and $\lambda^{-1} : \hat{C} \to \mathbb{R}$ is the inverse of λ with respect to the $\hat{\star}$ convolution product. It is easy to see that the product $\odot_\lambda$ is associative and commutative, arguing as in the proof of Theorem 4.5.1. Since we have not supposed the coaction δ to be multiplicative with respect to the $\odot$ product in C, the product $\odot_\lambda$ is in general different from $\odot$.

Definition 4.8.1. *The map $\psi_\lambda^{-1} = \psi_{\lambda^{-1}}$ is called the generalized Wick λ-product map.*

We denote by $\mathscr{C} := \mathscr{C}(\mathbb{R}^d)$ the space of continuous functions $f : \mathbb{R}^d \to \mathbb{R}$, for a fixed $d \geq 1$; we endow $\mathscr{C}$ with the associative commutative product $\cdot$ given by the pointwise multiplication. We consider the spaces $\mathrm{Lin}(V, \mathscr{C})$ and $\mathrm{Lin}(C, \mathscr{C})$ of linear functionals from V, respectively C, to $\mathscr{C}$. One can think to $\mathrm{Lin}(V, \mathscr{C})$ as a space of T-indexed functions: this is typically what happens in perturbative expansions indexed by combinatorial objects (sequences, in usual Taylor expansions or in Lyons' classical theory of geometric rough paths, or more complex objects such as trees or forests, as in Gubinelli's theory of non-geometric rough paths or in the theory of regularity structures, for example).

For $\Pi \in \mathrm{Lin}(V, \mathscr{C})$ and $\Gamma \in \mathrm{Lin}(C, \mathscr{C})$ we use the notation

$$\Pi(\tau) = \langle \Pi, \tau \rangle \in \mathscr{C}, \quad \tau \in V,$$
$$\Gamma(\alpha) = \langle \Gamma, \alpha \rangle \in \mathscr{C}, \quad \alpha \in C.$$

Definition 4.8.2. *For every* $\Gamma \in \mathrm{Lin}(C, \mathscr{C})$ *we denote by* V_Γ *the vector space freely generated by* $\{\Gamma(\alpha), \alpha \in \mathfrak{T}\}$. *The canonical map from* V_Γ *to* $\mathscr{C}$ *is called the evaluation map and is written* ev.

Notice that V_Γ is not in general embedded in $\mathscr{C}$ (the evaluation map is not injective): we view here the $\Gamma(\alpha)$ as forming a basis of V_Γ although, as elements of $\mathscr{C}$ they could possibly satisfy linear relations. One should think of the $\Gamma(\alpha)$'s as $\mathfrak{T}$-indexed functions: the constructions we will perform on them will depend also on their indexing. To avoid ambiguities, we reserve from now on the notation $\Gamma(\alpha)$ to denote elements of V_Γ and the notation $\langle \Gamma, \alpha \rangle$ to denote elements in $\mathscr{C}$.

Definition 4.8.3. *We call every* $\Gamma \in \mathrm{Lin}(C, \mathscr{C})$ *a* generalised product *on* T. *If* $\Gamma \in \mathrm{Lin}(C, \mathscr{C})$ *is an algebra unital morphism from* $(C, \odot)$ *to* $(\mathscr{C}, \cdot)$*, namely if*

$$\langle \Gamma, \varnothing \rangle = 1, \qquad \langle \Gamma, \tau_1 \odot \cdots \odot \tau_n \rangle = \langle \Gamma, \tau_1 \rangle \cdots \langle \Gamma, \tau_n \rangle, \tag{4.37}$$

for $\tau_1, \ldots, \tau_n \in T$*,* $n \geq 1$*, then* Γ *is called a* canonical product*, where* $\cdot$ *is the pointwise multiplication on* $\mathscr{C}$*.*

A generalised product is therefore a very general concept since we assume a priori no relation between $\langle \Gamma, \alpha \rangle$ and $\langle \Gamma, \alpha' \rangle$ for $\alpha, \alpha' \in \mathfrak{T}$ and $\alpha \neq \alpha'$. However we are mainly interested in deformations of canonical products, see below.

Lemma 4.8.4. *For every* $\Pi \in \mathrm{Lin}(V, \mathscr{C})$ *there exists a unique canonical product* $R\Pi \in \mathrm{Lin}(C, \mathscr{C})$ *such that*

$$\langle \Pi, \tau \rangle = \langle R\Pi, \tau \rangle, \qquad \forall\, \tau \in T.$$

Proof. The property (4.37) allows to construct $R\Pi$ uniquely by recursion. □

Now we show that a generalised product Γ allows to define a genuine commutative and associative product on the space V_Γ.

Definition 4.8.5. *For every* $\Gamma \in \mathrm{Lin}(C, \mathscr{C})$ *we define the commutative and associative product* $\mathscr{M}_\Gamma$ *on* V_Γ

$$\mathscr{M}_\Gamma(\Gamma(\alpha), \Gamma(\beta)) := \Gamma(\alpha \odot \beta), \qquad \forall\, \alpha, \beta \in C.$$

Then $(V_\Gamma, \mathscr{M}_\Gamma, \langle \Gamma, \varnothing \rangle)$ is a commutative unital algebra and, by the very definition, $\Gamma : (C, \odot) \to (V_\Gamma, \mathscr{M}_\Gamma)$ is an algebra isomorphism. In particular, for every $\Pi \in \mathrm{Lin}(V, \mathscr{C})$, $R\Pi : (C, \odot) \to (V_{R\Pi}, \mathscr{M}_{R\Pi})$ is an algebra morphism and $\mathscr{M}_{R\Pi}$ is mapped by the evaluation map to the canonical pointwise product on $\mathscr{C}$.

We want now to define *deformations* of the products of Definition 4.8.5, taking inspiration from Section 4.5.

Definition 4.8.6. *For every $\Gamma \in \mathrm{Lin}(C, \mathscr{C})$ and every unital $\lambda : \hat{C} \to \mathbb{R}$ we can define a product $\mathcal{M}_\Gamma^\lambda$ on V_Γ by*

$$\mathcal{M}_\Gamma^\lambda(\Gamma(\alpha), \Gamma(\beta)) \coloneqq \Gamma(\alpha \odot_\lambda \beta), \qquad \forall\, \alpha, \beta \in C, \tag{4.38}$$

such that $\Gamma : (C, \odot_\lambda) \to (V_\Gamma, \mathcal{M}_\Gamma^\lambda)$ is an algebra isomorphism.

We say that $\mathcal{M}_\Gamma^\lambda$ is a λ-deformation of $\mathcal{M}_\Gamma$, since in the case λ is the counit $\hat{\varepsilon}$ of $\hat{C}$, the coaction property (4.33) of δ implies that $M_{\hat{\varepsilon}}$ is the identity map on C, hence $\mathcal{M}_\Gamma^{\hat{\varepsilon}}$ coincides with $\mathcal{M}_\Gamma$. In particular we have

Definition 4.8.7. *For every $\Pi \in \mathrm{Lin}(V, \mathscr{C})$ we can define a λ-deformation $\cdot_\lambda \coloneqq \mathcal{M}_{R\Pi}^\lambda$ of the canonical product $\mathcal{M}_{R\Pi}$ on $V_{R\Pi}$, such that*

$$\Pi(\tau_1) \cdot_\lambda \cdots \cdot_\lambda \Pi(\tau_n) \coloneqq R\Pi(\tau_1 \odot_\lambda \cdots \odot_\lambda \tau_n), \tag{4.39}$$

and $\tau_1, \ldots, \tau_n \in T$.

We stress again that, unlike the pointwise multiplication $\cdot$, the λ-deformation $\cdot_\lambda$ is not defined on $\mathscr{C}$ but rather, for every fixed $\Pi \in \mathrm{Lin}(V, \mathscr{C})$, on $V_{R\Pi}$. As stated below Definition 4.8.6, the deformation $\mathcal{M}_{R\Pi}^{\hat{\varepsilon}}$ coincides with the canonical pointwise product on $V_{R\Pi}$.

Lemma 4.8.8. *For every unital $\lambda : \hat{C} \to \mathbb{R}$ and $\Gamma \in \mathrm{Lin}(C, \mathscr{C})$ the map*

$$\Gamma^\lambda \coloneqq \Gamma \circ \psi_\lambda^{-1} \in \mathrm{Lin}(C, \mathscr{C}) \tag{4.40}$$

defines an algebra isomorphism from $(C, \odot)$ to $(V_{\Gamma^\lambda}, \mathcal{M}_\Gamma^\lambda)$.

Proof. By eq. (4.38)

$$\mathcal{M}_\Gamma^\lambda(\Gamma^\lambda(\alpha), \Gamma^\lambda(\beta)) = \Gamma^\lambda(\alpha \odot \beta), \qquad \forall\, \alpha, \beta \in C,$$

and the claim follows. □

In particular for $\Gamma = R\Pi$ we obtain by eq. (4.39)

$$\Gamma^\lambda(\tau_1) \cdot_\lambda \cdots \cdot_\lambda \Gamma^\lambda(\tau_n) \coloneqq \Gamma^\lambda(\tau_1 \odot \cdots \odot \tau_n), \tag{4.41}$$

which is reminiscent of eq. (4.27).

Example 4.8.9. In the setting of the previous sections, we can consider $T = \mathcal{A}$, so that in this case $V = \mathbb{R}^{\mathcal{A}}$, $C = H$ and $\hat{C} = \hat{H}$. Then the most natural choice of $\Pi \in \mathrm{Lin}(V, \mathscr{C}(\mathbb{R}^{\mathcal{A}}))$ is given by $\Pi X_a \coloneqq t_a$, where $t_a : \mathbb{R}^{\mathcal{A}} \to \mathbb{R}$ is the evaluation of the a-component, and $R\Pi : H \to \mathscr{C}$ is

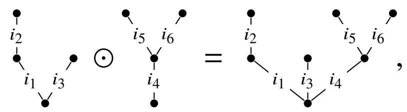
$$\langle R\Pi, X_{a_1} \ldots X_{a_n} \rangle \coloneqq t_{a_1} \cdots t_{a_n}, \qquad a_1, \ldots, a_n \in \mathcal{A},$$

Then eq. (4.41) is the analogue of eq. (4.27) in this context, while eq. (4.39) defines a deformation $\cdot_\lambda$ of the pointwise product of $t_{a_1}, \ldots, t_{a_n}$. This point of view will be generalised in the next section. $\diamond$

4.9 Wick products of trees

We now discuss the main example we have in mind of the general construction in Section 4.8, namely rooted trees, that are a generalisation of classical monomials as we show below. With the application to rough paths in mind [51], we denote by $\mathfrak{T}$ the set of all non-planar non-empty rooted trees with edges (not nodes) decorated with letters from a finite alphabet $\{1, \ldots, d\}$. We stress that all trees in $\mathfrak{T}$ have at least one node, the root.

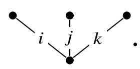
The set $\mathfrak{T}$ is a commutative monoid under the associative and commutative *tree product* $\odot$ given by the identifications of the roots, e.g.



(4.42)

see also [15, Definition 4.7].

The rooted tree $\bullet$ with a single node and no edge is the neutral element for this product. The set of monomials in d commuting variables $X_1, \ldots, X_d$ can be embedded in $\mathfrak{T}$ as follows: every primitive monomial X_i is identified with , and the product of monomials with the tree product. In this way every monomial is identified with a decorated corolla, for instance

$X_i X_j X_k \longrightarrow$  (4.43)

See the discussion around Lemma 4.9.3 below for more on this identification.

We denote by $T \subset \mathfrak{T}$ the set of all non-planar *planted* rooted trees. We recall that a rooted tree is planted if its root belongs to a single edge, called the trunk. For example, in the left-hand side of eq. (4.42), the first tree is not planted, while the second is.

We also denote by $\mathfrak{F}$ the set of non-planar rooted forests with edges (not nodes) decorated with letters from the finite alphabet $\{1, \ldots, d\}$, such that every non-empty connected component has at least one edge. On this space we define the product $\bullet$ given by the disjoint union, with neutral element the empty forest $\varnothing$.

We perform the identification

$$\bullet = \varnothing \tag{4.44}$$

between the rooted tree $\bullet \in \mathfrak{T}$ and the empty forest $\varnothing \in \mathfrak{F}$. Then we obtain canonical embeddings

$$T \hookrightarrow \mathfrak{T} \hookrightarrow \mathfrak{F} \tag{4.45}$$

and moreover

- $(\mathfrak{T}, \odot)$ is the free commutative monoid on T,
- $(\mathfrak{F}, \bullet)$ is the free commutative monoid on $\mathfrak{T}$.

In both cases the element $\bullet = \varnothing$ is the neutral element. We denote by

- V the vector space generated freely by T,
- C the vector space generated freely by $\mathfrak{T}$,
- $\hat{C}$ the vector space generated freely by $\mathfrak{F}$.

Then we have

- $(C, \odot)$ is the free commutative unital algebra generated by T,
- $(\hat{C}, \bullet)$ is the free commutative unital algebra generated by $\mathfrak{T}$,

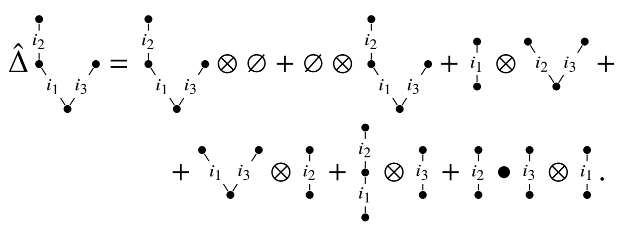
and again in both cases the element $\bullet = \varnothing$ is the neutral element. Finally, by eq. (4.44) and eq. (4.45) we also have canonical embeddings

$$V \hookrightarrow C \hookrightarrow \hat{C}. \tag{4.46}$$

On $\hat{C}$ we also define the coproduct $\hat{\Delta}$, given by the extraction-contraction operator of arbitrary subforests [17]:

$$\hat{\Delta}\tau = \sum_{\sigma \subseteq \tau} \sigma \otimes \tau/\sigma, \qquad \tau \in \mathfrak{F}, \tag{4.47}$$

where a subforest $\sigma \in \mathfrak{F}$ of τ is determined by a (possibly empty) subset of the set of edges of τ, and τ/σ is the tree obtained by contracting each connected component of σ to a single node. We recall that by eq. (4.44) the empty forest and the tree reduced to a single node are identified and called $\varnothing$. For example,



If $\hat{\varepsilon} : \hat{C} \to \mathbb{R}$ is the linear functional such that $\hat{\varepsilon}(\sigma) = \mathbf{1}(\sigma = \varnothing)$ for $\sigma \in \mathcal{F}$, then $(\hat{C}, \bullet, \varnothing, \hat{\Delta}, \hat{\varepsilon})$ is a Hopf algebra [17].

Note that, unlike $(\hat{H}, \hat{\Delta})$ in Section 4.7, $(\hat{C}, \hat{\Delta})$ is not co-commutative; moreover the canonical embedding $C \hookrightarrow \hat{C}$ in eq. (4.46) is not an algebra morphism from $(C, \odot)$ to $(\hat{C}, \bullet)$. We could also endow C with a coproduct Δ_C (the extraction-contraction operator of a subtree at the root, which plays an important role in [15] and is isomorphic to the classical Butcher–Connes–Kreimer coproduct), but we do not need this for what comes next.

We now go back to the construction of Section 4.8. With the embedding $C \hookrightarrow \hat{C}$, the coaction

$$\delta := \hat{\Delta} : C \mapsto \hat{C} \otimes C$$

makes C a left-comodule over $\hat{C}$ by an analogue of Proposition 4.7.3. Then we can define $\psi_\lambda : C \to C$ as in eq. (4.35), for $\lambda : \hat{C} \to \mathbb{R}$ unital, and a deformed product $\odot_\lambda$ on C as in eq. (4.36). As discussed before Definition 4.8.1, the coaction δ is not multiplicative with respect to $\odot$ and therefore $\odot_\lambda$ is in general truly different from $\odot$.

For $\Pi \in \mathrm{Lin}(V, \mathscr{C})$ and $\Gamma := R\Pi \in \mathrm{Lin}(C, \mathscr{C})$ as in Lemma 4.8.4, the map $\Gamma^\lambda = (R\Pi) \circ \psi_{\lambda^{-1}}$ defines by Lemma 4.8.8 an algebra isomorphism from $(C, \odot)$ to $(V_{\Gamma^\lambda}, \mathcal{M}_\Gamma^\lambda)$, so that in particular we have the analogue of eq. (4.41).

This idea is very important in regularity structures, where the pointwise product of explicit (random) distributions is ill-defined, while a suitable deformed product is well-defined as a (random) distribution. The above construction allows to recover a precise algebraic structure of such deformed pointwise products, in the same spirit as Theorem 4.5.4. See Section 4.10 below for a discussion.

We show now how these ideas can be implemented concretely, that is how a character λ can be constructed in practice in some interesting situation, generalising the construction of Wick polynomials in the previous sections of this chapter.

Let us now consider a $\mathrm{Lin}(V, \mathscr{C})$-valued random variable X, such that

- $\langle X, \varnothing \rangle = 1$,
- X is stationary, i.e., $\langle X, \tau \rangle(\cdot + x)$ has the same law as $\langle X, \tau \rangle$ for all $x \in \mathbb{R}^d$ and $\tau \in T$
- $\langle X, \tau \rangle(0)$ has finite moments of any order for all $\tau \in T$.

Then we can define

$$\mu : C \to \mathbb{R}, \qquad \mu(\tau) := \mathbb{E}(\langle RX, \tau \rangle(0)). \tag{4.48}$$

There is a unique extension of μ to a linear $\hat{\mu} : \hat{C} \to \mathbb{R}$ which is a character of $(\hat{C}, \bullet)$ and we denote by $\hat{\mu}^{-1} : \hat{C} \to \mathbb{R}$ its inverse with respect to the $\hat{\star}$ convolution product.

Theorem 4.9.1. *Let X be as above. The only character λ on $(\hat{C}, \bullet)$ such that*

$$\mathbb{E}(\langle RX, \psi_\lambda \tau\rangle(0)) = 0, \qquad \forall\, \tau \in \mathfrak{T} \setminus \{\varnothing\}$$

is equal to $\hat{\mu}^{-1}$.

Proof. We note first that for every character λ on $(\hat{C}, \bullet)$ we have

$$\mathbb{E}(\langle RX, \psi_\lambda \tau\rangle(x)) = (\lambda \otimes \mu)\hat{\Delta}\tau = (\lambda \,\hat{\star}\, \mu)\,\tau, \qquad \forall\, \tau \in \mathfrak{T}.$$

In particular

$$\mathbb{E}(\langle RX, \psi_{\hat{\mu}^{-1}} \tau\rangle(x)) = (\hat{\mu}^{-1} \otimes \hat{\mu})\hat{\Delta}\tau = 0, \qquad \forall\, \tau \in \mathfrak{T} \setminus \{\varnothing\}.$$

On the other hand, since λ and $\hat{\mu}$ are characters on $(\hat{C}, \bullet)$, if for all $\tau \in \mathfrak{T}$

$$(\lambda \,\hat{\star}\, \hat{\mu})\,\tau = \mathbf{1}(\tau = \varnothing),$$

then the same formula holds by multiplicativity for all $\tau \in \mathfrak{F}$ and we obtain that $\lambda = \hat{\mu}^{-1}$. $\square$

Remark 4.9.2. By stationarity, the function $RX \circ M_{\hat{\mu}^{-1}} \in \mathrm{Lin}(C, \mathscr{C})$ has the additional property

$$\mathbb{E}(\langle RX, M_{\hat{\mu}^{-1}} \tau\rangle(x)) = 0, \qquad \forall\, \tau \in \mathfrak{T} \setminus \{\varnothing\},\ x \in \mathbb{R}^d.$$

In other words, $RX \circ M_{\hat{\mu}^{-1}} : C \to \mathscr{C}$ gives a *centred* deformed product. $\triangle$

We now show that the construction on decorated rooted trees generalises in a very precise sense the Wick products of Section 4.4. We use the identification between monomials in d commuting variables $X_1, \ldots, X_d$ and corollas decorated with letters from $\{1, \ldots, d\}$ that we have explained in eq. (4.43). Choosing $A := \{1, \ldots, d\}$ we obtain a canonical embedding of $H \hookrightarrow C$, where H is the polynomial algebra defined in Section 4.3.1; we call $\mathrm{C}or$ the image of H in C by this embedding. Then a simple verification shows that

Lemma 4.9.3. *The embedding $H \hookrightarrow C$ is a Hopf algebra isomorphism between $(H, \cdot, \mathbf{1}, \Delta, \varepsilon)$ and $(\mathrm{C}or, \odot, \bullet, \hat{\Delta}, \hat{\varepsilon})$, where $\hat{\Delta}$ is defined in eq.* (4.47).

We obtain that every deformation $\odot_\lambda$ for a unital $\lambda : \hat{C} \to \mathbb{R}$ is a product on $\mathrm{C}or$ which is isomorphic to the deformed product defined in eq. (4.26) by restricting $\lambda : \mathrm{C}or \to \mathbb{R}$.

4.10 Connection with regularity structures

It would go beyond the scope of this work to introduce and explain the algebraic and combinatorial aspects of the seminal theory of Regularity Structures [56]; we want at least to explain how the concept of *renormalisation*, which plays such a prominent role there, is intimately related to the *deformation* of the standard pointwise product described in the previous sections. These ideas can also be found in the theory of *rough paths* [50, 51, 86, 87], which has largely inspired the theory of regularity structures.

The recent papers [15, 20] introduce a Hopf algebra $\hat{H}$ together with a linear space H, which is moreover a left-comodule over $\hat{H}$ with coaction $\delta : H \to \hat{H} \otimes H$. This framework is then used to describe in a compact way a number of complicated algebraic operations, related to the concept of *renormalisation*. The space H in [15] is an expanded version of the span of decorated rooted trees V defined in Section 4.9 above; more precisely it is the free vector space on a more complicated set of decorated rooted trees, which is aimed at representing monomials of generalised Taylor expansions. The space $\hat{H}$ in [15] is a Hopf algebra of decorated forests with a condition of *negative homogeneity*.

In [15, 20], the linear space H codes random distributions, which depend on a regularisation parameter $\epsilon > 0$. As one removes the regularisation by letting $\epsilon \to 0$, these random distributions do not converge in general. More precisely, we have (random) linear functions $\Pi_\epsilon : H \to \mathcal{D}'(\mathbb{R}^d)$ which are well defined for all $\epsilon > 0$, but for which there is in general no limit as $\epsilon \to 0$. In fact, we even have $\Pi_\epsilon : H \to \mathcal{C}(\mathbb{R}^d)$, and Π_ϵ is constructed in a multiplicative way as in Lemma 4.8.4 above. Indeed, although H is not an algebra, it is endowed with a *partial product*, i.e., some but not all pairs of its elements are supposed to be multiplied. We try to make this idea more precise in the next

Definition 4.10.1. *A* partial product *on H is a pair $(\mathcal{M}, S)$ where $S \subseteq H \otimes H$ is a linear space and $\mathcal{M} : S \to H$ is a linear function.*

Therefore, if τ and σ are elements of H, their product $\mathcal{M}(\tau \otimes \sigma)$ is well defined if and only if $\tau \otimes \sigma \in S$. For example, in regularity structures one has an element $\Xi \in H$ such that $\Pi_\epsilon \Xi = \xi_\epsilon := \rho_\epsilon * \xi$, where ξ is a white noise on $\mathbb{R}^d$ (a random distribution in $\mathcal{D}'(\mathbb{R}^d)$) and $(\rho_\epsilon)_{\epsilon>0}$ is a family of mollifiers. Although $(\xi_\epsilon)^2$ is well-defined as a pointwise product in $\mathcal{C}(\mathbb{R}^d)$, as $\epsilon \to 0$ there is no limit in $\mathcal{D}'(\mathbb{R}^d)$ and indeed, we do not expect to multiply ξ by itself in $\mathcal{D}'(\mathbb{R}^d)$. We express this by imposing that $\Xi \otimes \Xi \notin S$.

The divergences that arise in this context are due to ill-defined products; this is already clear in the example of $\Xi \otimes \Xi$ and $(\xi_\epsilon)^2$. Another more subtle example is the following: we consider $\xi_\epsilon := \rho_\epsilon * \xi$ again, and a (possibly random) function $f_\epsilon : \mathbb{R}^d \to R$ which, as $\epsilon \to 0$, tends to a non-smooth function f. Then the pointwise product $f_\epsilon \cdot \xi_\epsilon$ does not converge in general, since the product $f \cdot \xi$ is ill-defined in $\mathcal{D}'(\mathbb{R}^d)$. However a proper deformation of this pointwise product may still be well defined in the limit.

Let $(\tau_i, i \in I) \subset H$ be a family that freely generates H as a linear space. We can now give the following

Definition 4.10.2. *Let $\Pi : H \to \mathcal{D}'(\mathbb{R}^d)$ be a linear map and $(\mathcal{M}, S)$ a partial product on H. We define V_Π as the vector space freely generated by $\{\langle \Pi, \tau_i \rangle : i \in I\}$ as in Definition 4.8.2. Then we define a partial product on V_Π as follows:*

- $S_\Pi \subseteq V_\Pi \otimes V_\Pi := \{\Pi(\tau) \otimes \Pi(\sigma) : \tau \otimes \sigma \in S\}$, $\Pi(\tau) := \langle \Pi, \tau \rangle$

- $\mathcal{M}_\Pi : S_\Pi \to V_\Pi$, $\mathcal{M}_\Pi(\Pi(\tau) \otimes \Pi(\sigma)) := \Pi(\mathcal{M}(\tau \otimes \sigma))$.

We stress that this definition allows to define partial products of distributions in a very general setting. We are clearly inspired by the construction of the previous sections, by realising that we can even work on distributions rather than on continuous functions.

However the construction of interesting $\Pi : H \to \mathcal{D}'(\mathbb{R}^d)$ may not be simple. The method which is successfully used in a large class of applications in [15, 20, 56] is the following. We start from a linear $\Pi_\epsilon : H \to \mathcal{C}(\mathbb{R}^d)$ which is *canonical* in the sense that

$$\langle \Pi_\epsilon, \mathcal{M}(\tau \otimes \sigma) \rangle = \langle \Pi_\epsilon, \tau \rangle \cdot \langle \Pi_\epsilon, \sigma \rangle, \qquad \forall\, \tau \otimes \sigma \in S,$$

where $\cdot$ is the standard pointwise product in $\mathcal{C}(\mathbb{R}^d)$. In order to obtain a convergent limit as $\epsilon \to 0$, we try to deform this pointwise product, using the comodule structure of H over $\hat{H}$. For all unital multiplicative and linear $\lambda : \hat{H} \to \mathbb{R}$ we define $\psi_\lambda : H \to H$ as in eq. (4.35) and then we set as in eq. (4.40)

$$\Pi_\epsilon^\lambda(\tau) := \Pi_\epsilon(\psi_\lambda^{-1}\tau), \qquad \tau \in H.$$

Then we can define the deformed partial product on V_{Π_ϵ}:

$$\mathcal{M}^\lambda_{\Pi_\epsilon}(\Pi^\lambda_\epsilon(\tau), \Pi^\lambda_\epsilon(\sigma)) := \Pi^\lambda_\epsilon(\tau \otimes \sigma), \qquad \tau \otimes \sigma \in S.$$

If $\lambda = \lambda_\epsilon$ is chosen in such a way that $\Pi_\epsilon^{\lambda_\epsilon}$ converges to a well defined linear map $\hat{\Pi} : H \to \mathcal{D}'(\mathbb{R}^d)$, then we can define on $V_{\hat{\Pi}}$ the partial product

$$\mathcal{M}_{\hat{\Pi}}(\hat{\Pi}(\tau), \hat{\Pi}(\sigma)) := \hat{\Pi}(\tau \otimes \sigma), \qquad \tau \otimes \sigma \in S$$

which is the analogue of eq. (4.41) in this setting. We note that in general neither Π_ϵ nor λ_ϵ converge; indeed, λ_ϵ diverges exactly in a way that compensates the divergence of Π_ϵ, in such a way that $\Pi_\epsilon^{\lambda_\epsilon}$ converges.

The fact that the above construction can indeed be implemented in a large number of interesting situations is the result of [15, 20]. Those papers consider random maps Π_ϵ with suitable properties which resemble those of X in Theorem 4.9.1, namely Π_ϵ is supposed to be stationary and to possess finite moments of all orders. Then, as in Theorem 4.9.1, it is possible to choose a specific element $\lambda_\epsilon : \hat{H} \to \mathbb{R}$ which yields a *centred* family of functions $\Pi_\epsilon \circ \psi_{\lambda_\epsilon^{-1}}$, see [15, Theorem 6.17]. Under very general conditions, this special choice produces a converging family as $\epsilon \to 0$ [20].

Therefore the renormalised (converging) random distributions are a *centred* version of the original (non-converging) ones. The specific functional λ_ϵ^{-1} is equal to $\mu_\epsilon \circ \mathscr{A}$, where $\mu_\epsilon : \hat{H} \to \mathbb{R}$ is an expectation with respect to Π_ϵ as in eq. (4.48), and $\mathscr{A}$ is a *twisted antipode*; the functional $\mu_\epsilon \circ \mathscr{A}$ plays the role which is played by $\hat{\mu}^{-1}$ in Theorem 4.9.1.

Remark 4.10.3. We stress that the centred family $\Pi_\epsilon \circ \psi_{\lambda_\epsilon^{-1}}$ can not be in general reduced to the Wick polynomials of Theorem 4.4.2. This is because the coaction $\delta : H \to \hat{H} \otimes H$ in this context is significantly more complex than eq. (4.16) and eq. (4.47). $\triangle$

Chapter 5

Modifying Rough Paths

5.1 Introduction

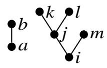
The theory of Rough Paths has been introduced by Terry Lyons in the ’90s with the aim of giving an alternative construction of stochastic integrations and stochastic differential equations. More recently, it has been expanded by Martin Hairer to cover stochastic partial differential equations, with the invention of regularity structures.

A rough path and a *model* of a regularity structure are mathematical objects which must satisfy some algebraic and analytical constraints. For instance, a rough path can be described as a Hölder function defined on an interval and taking values in a non-linear finite-dimensional Lie group; models of regularity structures are a generalisation of this idea. A crucial ingredient of regularity structures is the *renormalisation procedure*: given a family of models depending on a parameter $\varepsilon > 0$, which fails to converge in an appropriate topology as $\varepsilon \to 0$, one wants to modify it in a such a way that the algebraic and analytical constraints are still satisfied and the modified version converges. This procedure has been obtained in [15, 20] for a general class of situations with a *stationary character*.

The same question could have been asked much earlier about rough paths. Maybe this has not happened because the motivation was less compelling; although one can construct examples of rough paths depending on a parameter $\varepsilon > 0$ which do not converge as $\varepsilon \to 0$, this phenomenon is the exception rather than the rule. However the problem of characterizing the automorphisms of the space of rough paths is clearly of interest; one example is the transformation from Itô to Stratonovich integration, but our aim is to put this example in a much larger context.

We recall that there are several possible notions of rough paths; in particular we have *geometric* RPs and *branched* RPs, two notions defined respectively by Terry Lyons [86] and Massimiliano Gubinelli [51], see Sections 5.3 and 5.4 below. These two notions are intimately related to each other, as shown by Hairer and Kelly [57], see Section 5.4 below. We note that regularity structures [56] are a natural and far-reaching generalisation of branched RPs.

In this paper we concentrate on the automorphisms of the space of branched RPs, see remark 5.1.11 for a discussion of the geometric case. Let $\mathcal{F}$ be the collection of all non-planar rooted forests with nodes decorated by $\{1, \ldots, d\}$, see Section 5.4 below, and also Section 2.5. For instance the following forest



is an element of $\mathcal{F}$. We call $\mathcal{T} \subset \mathcal{F}$ the set of *rooted trees*, namely of non-empty forests with a single connected component. We grade elements $\tau \in \mathcal{F}$ by the number $|\tau|$ of their nodes.

Let now $\mathcal{H}$ be the linear span of $\mathcal{F}$. It is possible to endow $\mathcal{H}$ with a product and a coproduct $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ which make it a Hopf algebra, also known as the Butcher-Connes-Kreimer Hopf algebra. We let $\mathcal{G}$ denote the set of all *characters* over $\mathcal{H}$, that is, elements of $\mathcal{G}$ are functionals $X \in \mathcal{H}^*$ that are also multiplicative in the sense that

$$\langle X, \tau\sigma \rangle = \langle X, \tau \rangle \langle X, \sigma \rangle$$

for all forests (and in particular trees) $\tau, \sigma \in \mathcal{F}$. Furthermore, the set $\mathcal{G}$ can be endowed with a product $*$, dual to the coproduct, defined pointwise by $\langle X * Y, \tau \rangle = \langle X \otimes Y, \Delta\tau \rangle$. See Section 2.5 for further details. We work on the compact interval $[0, 1]$ for simplicity, and all results can be proved without difficulty on $[0, T]$ for any $T \geq 0$.

Definition 5.1.1. *Given $\gamma \in]0, 1[$, a branched γ-rough path is a path $X : [0, 1]^2 \to \mathcal{G}$ such that $X_{tt} = \varepsilon$, it satisfies Chen's rule*

$$X_{su} * X_{ut} = X_{st}, \qquad s, u, t \in [0, 1],$$

and the analytical condition

$$|\langle X, \tau \rangle| \lesssim |t - s|^{\gamma|\tau|}.$$

Setting $x^i_t := \langle X_{0t}, \bullet_i \rangle$, $t \in [0, 1]$, we say that X is a branched γ-rough path over the path $x = (x^1, \ldots, x^d)$.

Our main result is that we obtain a transitive action of an additive group of functions on branched rough paths, allowing to translate any given branched rough path into any other branched rough path by modifying its components. For a fixed $\gamma \in]0, 1[$, we define $N := \lfloor \gamma^{-1} \rfloor$. For instance, if $\gamma \in]1/3, 1/2]$ then $N = 2$. Let $\mathcal{C}^\gamma$ be the set of all collections of functions $(g^\tau : \tau \in \mathcal{T} : |\tau| \leq N)$ indexed by rooted trees with fewer than N nodes, such that $g^\tau \in C^{\gamma|\tau|}([0, 1])$, the classical space of real-valued Hölder functions on $[0, 1]$ with Hölder exponent $\gamma|\tau|$. Clearly this set is an abelian group under pointwise addition.

Theorem 5.1.2. *There is a transitive action of $\mathcal{C}^\gamma$ on branched γ-rough paths $(g, X) \mapsto gX$ such* *that for each $g, g' \in \mathscr{C}^\gamma$ and branched γ-rough path X the identity $g'(gX) = (g + g')X$ holds. For every pair of branched γ-rough paths X and X' there exists $g \in \mathscr{C}^\gamma$ such that $gX = X'$.*

The importance of Theorem 5.1.2 resides in the following remark. Chen's rule states that for any tree $\tau \in \mathfrak{T}$ the identity $\langle X_{su} * X_{ut}, \tau\rangle = \langle X_{st}, \tau\rangle$ holds. By introducing the *reduced coproduct* $\Delta'\tau := \Delta\tau - \tau \otimes \mathbf{1} - \mathbf{1} \otimes \tau$ and the function $F^\tau_{st} = \langle X_{st}, \tau\rangle$ this identity may we rewritten as

$$\delta F^\tau_{sut} = \langle X_{su} \otimes X_{ut}, \Delta'\tau\rangle.$$

where $\delta F^\tau_{sut} := F^\tau_{st} - F^\tau_{su} - F^\tau_{ut}$ is the second order finite increment considered by Gubinelli [50]. Therefore, if X and $\tilde{X}$ are two branched rough paths above the same base paths $x^i \in C^\gamma$, and $\tau = \overset{\bullet_j}{\underset{\bullet_i}{|}}$ we obtain that

$$\delta F^\tau_{sut} = \langle X_{su} \otimes X_{ut}, \Delta'\tau\rangle = (x^j_u - x^j_s)(x^i_t - x^i_u) = \delta \tilde{F}^\tau_{sut} \tag{5.1}$$

where $\tilde{F}^\tau$ is obviously defined.

The finite increment operator δ has the following property: if $F : [0,1]^2 \to \mathbb{R}$ is such that $\delta F = 0$ then there exists $f : [0,1] \to \mathbb{R}$ such that $F_{st} = f_t - f_s$. The proof of this fact is an easy exercise, and we remark that the function f is unique up to an additive constant shift, see also [51, formula (5)]. Thus, by this fundamental property there exists a function $g^\tau : [0,1] \to \mathbb{R}$ such that

$$F^\tau_{st} = \tilde{F}^\tau_{st} + g^\tau_t - g^\tau_s. \tag{5.2}$$

Moreover, since both F^τ and $\tilde{F}^\tau$ satisfy $|F^\tau_{st}| \lesssim |t-s|^{2\gamma}$ we have that g^τ is actually 2γ-Hölder. Hence the first level fixes the second up to the increment of a 2γ-Hölder function.

The same argument applies for any tree τ such that $2 \le |\tau| \le N := \lfloor \gamma^{-1} \rfloor$. Indeed, the reduced coproduct $\Delta'\tau$ is a sum of tensor products of trees with strictly less nodes, and therefore the values of $\langle X_{st}, \tau\rangle$ and $\langle \tilde{X}_{st}, \tau\rangle$ differ only by the increment of a $\gamma|\tau|$-Hölder function g^τ. As soon as $|\tau| \ge N + 1$, then by Gubinelli's *Sewing Lemma* [50] the function $(s,t) \mapsto \langle X_{st}, \tau\rangle$ is uniquely determined by the values of X on trees with strictly less than N nodes. More explicitly, the Sewing Lemma is an *existence and uniqueness result*; however, for $|\tau| \ge N + 1$ we have no uniqueness, as we have already seen, and existence is not trivial.

For instance, suppose that $N \ge 2$, namely $\gamma \le 1/2$; suppose we have a branched γ-rough path X and functions $g^\tau \in C^{\gamma|\tau|}$ for every tree $\tau \in \mathfrak{T}$ with at most N nodes. We set

$$\langle \tilde{X}_{st}, \bullet_i\rangle := \langle X_{st}, \bullet_i\rangle + g^{\bullet_i}_t - g^{\bullet_i}_s.$$

We need now to define $\langle \tilde{X}_{st}, \overset{\bullet_j}{\underset{\bullet_i}{|}}\rangle$. By Chen's rule this must satisfy

$$\delta\langle \tilde{X}, \overset{\bullet_j}{\underset{\bullet_i}{|}}\rangle_{sut} = \langle \tilde{X}_{su}, \bullet_i\rangle\langle \tilde{X}_{ut}, \bullet_j\rangle$$

and the analytical condition requires that $|\langle \tilde{X}_{st}, \overset{\bullet_j}{\bullet_i} \rangle| \lesssim |t-s|^{2\gamma}$. In this setting the Sewing Lemma does not apply and therefore does not give existence of a function $\langle \tilde{X}_{st}, \overset{\bullet_j}{\bullet_i} \rangle$ with the required properties, so that a different approach is necessary. The same problem applies for all trees $\tau \in \mathfrak{F}$ with at most N nodes. We note that, if $g^{\bullet_i} \in C^{N\gamma}([0,1])$ for all $i \in \{1, \ldots, d\}$, we have one and only one choice for $\langle \tilde{X}, \overset{\bullet_j}{\bullet_i} \rangle$ given by

$$\langle \tilde{X}_{st}, \overset{\bullet_j}{\bullet_i} \rangle := \langle X_{st}, \overset{\bullet_j}{\bullet_i} \rangle + \int_s^t \langle X_{su}, \bullet_j \rangle \, \mathrm{d}g_u^{\bullet_i} + (g_u^{\bullet_j} - g_s^{\bullet_j}) \, \mathrm{d}x_u^i \tag{5.3}$$

where the integrals are well-defined in the Young sense, see [50, section 3]. However for $g^{\bullet_i} \in C^{\gamma}([0,1])$, which is our assumption, the Young integral is not defined and any choice of $\langle \tilde{X}_{st}, \overset{\bullet_j}{\bullet_i} \rangle$ can be used to give a weak definition of the integral in the right hand side. Therefore this shows that our discussion involves algebraic and analytical considerations, as it is often the case for rough paths.

This above discussion shows that Theorem 5.1.2 yields the following

Theorem 5.1.3. *Given a branched γ-rough path X, the map $g \to gX$ yields a bijection between $\mathscr{C}^{\gamma}$ and the set of branched γ-rough paths.*

Therefore Theorem 5.1.3 yields a complete parametrisation of the space of branched rough paths. This result is somewhat surprising, since rough paths form a non-linear space, in particular because of the Chen relation; however Theorem 5.1.3 yields a natural bijection between the space of branched γ-rough paths and the linear space $\mathscr{C}^{\gamma}$.

Moreover, the fact that the above Young integrals are not well defined shows why existence of the map $X \to gX$ is not trivial.

Theorem 5.1.3 also gives a complete answer to the question of existence and characterization of branched γ-rough paths over a γ-Hölder path x. Unsurprisingly, for our construction we start from a result of T. Lyons and N. B. Victoir's [88] of 2007, which was the first general theorem of existence of a geometric γ-rough path over a γ-Hölder path x, see our discussion of Theorem 5.1.4 below.

5.1.1 Outline of the results

A "canonical" choice of a geometric rough path X over smooth x is given by its *signature* [23, 86]. Other cases where geometric rough paths have been constructed are Brownian motion and fractional Brownian motion (see [27] for the $H > \frac{1}{4}$ case and [96] for the general case) among others. However, until T. Lyons and N. B. Victoir's paper [88] in 2007, this question remained largely open in the general case. The precise result is as follows

Theorem 5.1.4 (Lyons–Victoir extension)**.** *If $p \in [1,\infty) \setminus \mathbb{N}$ and $\gamma := 1/p$, a γ-Hölder path $x : [0,1] \to \mathbb{R}^d$ can be lifted to a geometric γ-rough path. For any $p \geq 1$ and $\varepsilon \in]0,\gamma[$, a γ-Hölder path can be lifted to a geometric $(\gamma-\varepsilon)$-rough path.*

The construction of a geometric γ-rough path requires to satisfy at the same time analytical and algebraic constraints: on one hand, the algebraic conditions demand that at each step Chen's rule and the multiplicativity of the rough path with respect to the shuffle product must hold, and on the other hand the path so obtained should be γ-Hölder continuous with respect to a suitable metric, see Definition 5.3.1.

We follow this approach, using an explicit form of the Baker–Campbell–Hausdorff formula by Reutenauer [103] (formula (5.9)), as well as on analytic techniques akin to Gubinelli's Sewing Lemma [50], exploiting the fact that the group of characters over the shuffle Hopf algebra is in fact a Lie group whose topology may be metrized in various ways.

We now give a heuristic argument of how this construction works: a geometric rough path is a two-parameter family taking values in the character group $G^{(N)}$ of the truncated shuffle Hopf algebra $(H^{(N)}, \cdot, \bar{\Delta})$ over an alphabet A, subject to algebraic and analytical conditions (see Definition 5.3.1 below). Here the cutoff level N is given in terms of the regularity γ of the underlying path as $N := \lfloor \gamma^{-1} \rfloor$, i.e. it is the largest positive integer such that $N\gamma \leq 1$. We also recall that the space $\mathfrak{g}^{(N)}$ spanned by the *infinitesimal characters* over H forms a Lie algebra under the usual commutator bracket for the convolution product and that there is an exponential map $\exp_N : \mathfrak{g}^{(N)} \to G^{(N)}$ which is a bijection with inverse $\log_N : G^{(N)} \to \mathfrak{g}^{(N)}$ defined by the usual power series. The reader is referred to Section 2.4 for further details.

Definition 5.1.5. *A geometric rough path is a family $(X_{st})_{0\leq s,t\leq 1}$ such that $X_{st} \in G^N$, $X_{tt} = \varepsilon$ the counit of H. This family also satisfies Chen's rule $X_{st} = X_{su} \star X_{ut}$ and the analytical estimate $|\langle X_{st}, w\rangle| \lesssim |t-s|^{\gamma\ell(w)}$ for each word $w \in H$ with length $\ell(w) \leq N$.*

Therefore, for each $0 \leq s,t \leq 1$ we can find an appropriate logarithm $L_{st} = \log_N(X_{st})$, and the family $(L_{st})_{0\leq s,t\leq 1}$ is such that $L_{st} \in \mathfrak{g}^{(N)}$, $L_{tt} = 0$. Moreover, by Chen's rule and the Baker–Campbell–Hausdorff formula, see (5.7) below, we also have that

$$L_{st} = \log(X_{su} \star X_{ut}) = L_{su} + L_{ut} + \mathrm{BCH}(L_{su}, L_{ut})$$

where $\mathrm{BCH}(A,B)$ denotes the remaining terms in the BCH expansion of $\log(\exp(A) \star \exp(B))$. This last equality might be rewritten as the condition

$$\delta L_{sut} := L_{st} - L_{su} - L_{ut} = \mathrm{BCH}(L_{su}, L_{ut})$$

which is reminiscent of Gubinelli's Sewing Lemma. The analytical condition in Definition 5.1.5

translates into the bound

$$|\langle L_{st}, w\rangle| \leq \sum_{k=1}^{\ell(w)} \frac{1}{k} |\langle X_{st}^{\star k}, w\rangle| \lesssim |t-s|^{\gamma\ell(w)}.$$

Observe also that the condition $\langle X_{st}, a\rangle = x_t^a - x_s^a$ also fixes the first level of the logarithm as $\langle L_{st}, a\rangle = x_t^a - x_s^a$. Therefore, constructing the logarithms L should be equivalent –modulo some technical considerations– to constructing the geometric rough path X.

We will also make use of the following observation: the constraint imposed by Chen's relation allows us to reduce the two-parameter family $(X_{st})_{0\leq s,t\leq 1}$ to a one-parameter family $(X_t)_{0\leq t\leq 1}$. In fact, we have that since X_{tt} is the neutral element in $G^{(N)}$, equating two of the parametrs in Chen's rule we obtain that $\varepsilon = X_{st} \star X_{ts}$, i.e. $X_{st} = X_{ts}^{-1}$. Moreover, setting $u = 0$ in the same equation we find that $X_{st} = X_{s0} \star X_{0t} = X_{0s}^{-1} \star X_{0t}$, hence it suffices to find $X_t := X_{0t}$ with $X_0 = \varepsilon$. Since we are considering group-valued paths, the character $X_{st} = X_s^{-1} \star X_t$ naturally plays the role of the increments of the path $t \mapsto X_t$.

In their article, Lyons and Victoir use the fact that the BCH formula is easy to handle in some specific cases in order to provide a direct proof of their theorem, which we now shortly review. Take a pair (x^1, x^2) of γ-Hölder paths with $\frac{1}{3} \leq \gamma < \frac{1}{2}$ so that $N = 2$. We look for a logarithm $L_{st} \in \mathfrak{g}^{(2)}$ satisfying the conditions detailed above, with a fixed first level. In this case, the second order BCH formula is fully explicit

$$L_{st} = L_{su} + L_{ut} + \frac{1}{2}[L_{su}, L_{ut}] = L_{su} + L_{ut} + \frac{1}{2}(L_{su} \star L_{ut} - L_{ut} \star L_{su})$$

thus the second level components $Z^{ij} = \langle L, ij\rangle$ must satisfy $Z^{ii} = 0$, $Z^{ij} = -Z^{ji}$ and

$$\delta Z_{sut}^{12} = Z_{st}^{12} - Z_{su}^{12} - Z_{ut}^{12} = \frac{1}{2}\big[(x_u^1 - x_s^1)(x_t^2 - x_u^2) - (x_u^2 - x_s^2)(x_t^1 - x_u^1)\big]. \tag{5.4}$$

The right-hand side of this identity is bounded above by a constant times $|t-u|^\gamma |u-s|^\gamma$. Then, eq. (5.4) can be used, making some choices, to define the actual value of Z_{st}^{12} on the dyadic partition of $[0, 1]$, and this will satisfy the bound $|Z_{st}^{12}| \lesssim 2^{-2m\gamma}$ whenever $s = k2^{-m}$ and $t = k2^{-m}$ are two consecutive dyadics, as a consequence of (5.4). In particular, we must choose the initial value Z_{01}^{12}, and there is also some freedom in the choice of Z^{12}, since one can modify it by adding the increment of a $C^{2\gamma}$-Hölder function without affecting (5.4) as in (5.1)-(5.2). In any case, the construction depends heavily on these choices and thus the constructed object will neither unique nor canonical. Nevertheless in our later applications this will not be of great importance (see Theorem 5.1.2).

With this, we can build the character $Y_{st} = \exp_2(L_{st})$ on each pair of *consecutive* dyadics, and then extend the definition to arbitrary pairs by Chen's rule. In some way, this is as if we were doing a "telescopic sum" on the group $G^{(N)}$. Finally, one makes use of the following result, found in [88], to obtain a geometric rough path X as an extension of Y to arbitrary $s, t \in [0, 1]$.

Lemma 5.1.6. *Let (E, ρ) be a complete metric space and set $D = \{t_k^m := k2^{-m} : m \geq 0, k = 0, \ldots, 2^m - 1\}$. Suppose $y : D \to E$ is a path satisfying the bound $\rho(y_{t_k^m}, y_{t_{k+1}^m}) \lesssim 2^{-\gamma m}$ for some $\gamma \in (0, 1)$. Then, there exists a γ-Hölder path $x : [0, 1] \to E$ such that $x|_D = y$.*

The application of this lemma requires us to have a suitable notion of metric on $G^{(2)}$ such that the estimate satisfied by Z^{12} translates into the appropriate bound for Y, and such that the Hölder property of X with respect to this metric implies the desired analytic bounds in Definition 5.1.5.

In fact, we have the following result

Proposition 5.1.7. *There is a metric on G^N such that the analytical requirement in Definition 5.3.1 is equivalent to the fact that $t \mapsto X_t$ is γ-Hölder with respect to ρ.*

It has now become more apparent why one should have an explicit form of the terms appearing in the BCH expansion if one hopes to generalise this argument to arbitrary levels. In the above argument, some of the second level components were fixed by the first level, and in general some of the level n components will be given by linear combinations of the preceding ones. It turns out that in order to iterate this procedure, we must have an explicit enough BCH formula allowing us to prove estimates on the increment appearing in the right-hand side of Equation (5.4), which then translate into the correct bounds for the application of Lemma 5.1.6.

Using the same idea we extend this construction to the case where the collection $(x^1, \ldots, x^d)$ is allowed to have different regularities in each component, which we call *anisotropic (geometric) rough paths (aGRP)*.

Theorem 5.1.8. *Let $x^1, \ldots, x^d$ be a collection of real-valued paths such that x^i is γ_i-Hölder. There exists an anisotropic rough path X such that $\langle X_{st}, e_i \rangle = x_t^i - x_s^i$ for all $i = 1, \ldots, d$.*

The following property also holds: given a collection of functions $g^i \in C^{\gamma_i}$, let $\bar{x}_t^i = x_t^i + g_t^i$ and denote by gX the anisotropic geometric rough path above the path

$$\bar{x}_t = \sum_{i=1}^{d} \bar{x}_t^i e_i.$$

Then, for any two such functions g and g' we have that $g'(gX) = (g + g')X$.

This kind of extension to rough paths has already been explored in the papers [7, 52] in the context of isomorphisms between geometric and branched rough paths. It turns out that the additional property obtained by our method enables us to explicitly describe the propagation of suitable modifications from lower to higher levels. As a corollary, we devise a way to enlarge a given rough path by adding components to the base path.

Corollary 5.1.9. *Let $(x^1 : i = 1, \dots, d)$ be a collection of real-valued paths such that $x^i \in C^{\gamma_i}$, and let X be the anisotropic geometric rough path above $(x^1, \dots, x^d)$. If $x^0 \in C^{\gamma_0}$ is another path, let $V^0 = V \oplus \mathbb{R}e_0$ and X^0 be the anisotropic geometric rough path over $(x^i : i = 0, \dots, d)$. Then, the restriction of X^0 to V equals X.*

We then go on to describe the interpretation of the above results in the context of branched rough paths. We provide, in Proposition 5.6.5, an alternate description of the Hairer–Kelly map as a sum over a suitable set of partitions of the given tree, as opposed to the original iterative definition, which we then use to give a way to encode branched rough paths as anisotropic geometric rough paths, along the same lines as in [7].

Theorem 5.1.10. *Let X be a branched γ-rough path. There exists an anisotropic geometric rough path $\bar{X}$ indexed by words on trees, with exponents $\gamma_\tau = \gamma|\tau|$, and such that $\langle X, \tau\rangle = \langle \bar{X}, \psi(\tau)\rangle$, where ψ is the Hairer–Kelly map.*

The main difference of this result with [57, Theorem 1.9] is that we obtain an anisotropic geometric rough path instead of a classical geometric rough path. This means that we do not construct unneeded components, i.e. components with high regularity, and we also obtain the right Hölder estimates in terms of the size of the indexing tree. This addresses two problems mentioned in Hairer and Kelly's work, namely Remarks 4.14 and 5.9 in [57].

Our approach also does not make use of Foissy's and Chapoton's Hopf-algebra isomorphism [21, 43] between the Grossman–Larson Hopf algebra and the free algebra over a complicated set I of trees as is done in [7]. This allows us to construct an action of a *larger* group on the set of branched rough paths; indeed, using the above isomorphism one would obtain a transformation group parametrised by $(g^\tau)_{\tau\in I}$ where I it the abovementioned set of trees and $g^\tau \in C^{\gamma|\tau|}$; on the other hand our approach yields a transformation group parametrised by $(g^\tau)_{\tau\in T}$ where T is the set of all trees with at most N nodes. With the smaller set I, transitivity of the action $g \mapsto gX$ would be lost.

Remark 5.1.11. A similar argument cannot be used to give the analogous result for geometric rough paths. In the case of branched rough paths, constructing the functions g^τ is enough since, by the freeness of the Connes–Kreimer Hopf algebra, this fixes the values of $\langle gX, \sigma\rangle$ for all forests σ. On the contrary, the shuffle algebra is **not** free over the vector space spanned by words, but over a smaller subset called *Lyndon words*. Therefore, fixing functions g^w for all words w might give an inconsistent object gX in the sense that it might not respect the character property. △

Another conclusion that might be drawn from this remark is the following. Since H is not free over the vector space spanned by the free monoid $\mathrm{M}(A)$ over the alphabet A, if we want to define a character $X \in H^*$ is suffices to define its values $\langle X, w\rangle$ for all Lyndon words. As stated above, trying to define the value of $\langle X, w\rangle$ for all $w \in \mathrm{M}(A)$ might result in a linear map that is not multiplicative. The point is that the shuffle product imposes additional relations on these values that must be taken into consideration. For example, if we are trying to define $\langle X, ab\rangle$ for a two-letter word $ab \in \mathrm{M}(A)$ and we know the values of $\langle X, a\rangle$ for all $a \in A$, we have to consider the fact that

$$\langle X, a\rangle\langle X, b\rangle = \langle X, ab\rangle + \langle X, ba\rangle$$

as $a \sqcup\!\sqcup\, b = ab + ba$, so knowing one of $\langle X, ab\rangle$ or $\langle X, ba\rangle$ already fixes the other value given that we know $\langle X, a\rangle$ and $\langle X, b\rangle$. In [57] Hairer and Kelly introduce a geometric rough path as a path taking values in the full group G of characters over H, as opposed to the truncated group as we have done above which is the original definition by Lyons (c.f. Definition 5.3.1). In a subsequent remark they claim that these definitions are equivalent by invoking [86, Theorem 2.2.1]. This theorem states that a "multiplicative functional" on $G^{(N)}$ satisfying certain analytical conditions (those of the definition of a geometric rough path) can be extended in a unique way to all of G. The fact is that there is a confusion in the terminology used by both sets of authors, since what Lyons calls a multiplicative functional in fact refers to Chen's rule and to the character property, which he calls "group-like" (in reference to Proposition 2.4.6). Therefore, Lyons' extension theorem is about linear maps satisfying only Chen's rule and the analytical bound and does not take care of the character property. We address this issue and we prove that

Theorem 5.1.12. *Let X be a geometric γ-rough path. There exists a path $\hat{X} : [0,1]^2 \to G$ such that $\hat{X}_{tt} = \varepsilon$, satisfying Chen's rule and the analytic estimate with the additional property that its restriction to words of length less than or equal to $N = \lfloor \gamma^{-1} \rfloor$ equals X.*

Outline. We start by reviewing all the theoretical concepts needed to make the exposition in this section formal. In Section 5.3 we state and prove the main result of this chapter, namely we give an explicit construction of a geometric rough path above any given path $x \in C^\gamma(\mathbb{R}^d)$. Next, in Section 5.5 we extend this result to the wider class of anisotropic geometric rough paths. Finally, in Section 5.4 we connect our construction with M. Gubinelli's branched rough paths, and we extend M. Hairer and D. Kelly's work in Section 5.6.1. We also explore possible connections with renormalisation in Section 5.6.2 by studying how our construction behaves under modification of the underlying paths. Then, we connect this approach with a recent work by Bruned, Chevyrev, Friz and Preiß [13] in Section 5.6.3 who borrowed ideas from the theory of Regularity Structures [15, 56] and proposed a renormalisation procedure for geometric and branched rough paths [13] based on pre-Lie morphisms.

The main difference between our result and the BCFP procedure is that they consider translation only by time-independent factors, whereas –under reasonable hypotheses– we are also able to handle general translations depending on the time parameter. We also mention that some further algebraic aspects of renormalisation in rough paths have been recently developed in [14].

5.2 Preliminaries

For the rest of this section we fix a locally finite graded connected Hopf algebra $\mathcal{H}$, that is, $\mathcal{H}$ is a vector space endowed with a product $m : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$ and a coproduct $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ satisfying certain compatibility assumtions. There is also a unit $\mathbf{1} \in \mathcal{H}$, a counit $\varepsilon \in \mathcal{H}^*$ and an antipode $S : \mathcal{H} \to \mathcal{H}$ such that

$$m(\mathrm{id} \otimes S)\Delta x = \varepsilon(x)\mathbf{1} = m(S \otimes \mathrm{id})\Delta x$$

for all $x \in \mathcal{H}$. As usual we will denote the image $m(x \otimes y) = xy$ in order to reduce notation. The fact that $\mathcal{H}$ is locally finite graded connected means that it can be decomposed as a direct sum

$$\mathcal{H} = \bigoplus_{n=0}^{\infty} \mathcal{H}_{(n)}$$

with $\mathcal{H}_{(0)} = \mathbb{R}\mathbf{1}$, each $\mathcal{H}_{(n)}$ is finite-dimensional and

$$m(\mathcal{H}_{(n)} \oplus \mathcal{H}_{(m)}) \subset \mathcal{H}_{(n+m)}, \quad \Delta\mathcal{H}_{(n)} = \bigoplus_{p+q=n} \mathcal{H}_{(p)} \otimes \mathcal{H}_{(q)}.$$

The reader is referred to Chapter 2 for further details. Each element $x \in \mathcal{H}$ can thus be decomposed as a sum

$$x = \sum_{n=0}^{\infty} x_n$$

with $x_n \in \mathcal{H}_{(n)}$ where only a finite number of the summands are non-zero. We call each x_n the *homogeneous part of degree n* of x, and elements of $\mathcal{H}_{(n)}$ are said to be homogeneous of degree n. In this case we write $|x_n| = n$. From now on the grading $(\mathcal{H}_{(n)} : n \in \mathbb{N})$ will be called the *standard grading*.

The grading property of $\mathcal{H}$ implies in particular that for homogeneous elements $x \in \mathcal{H}_{(n)}$ the coproduct can be written as

$$\Delta x = x \otimes \mathbf{1} + \mathbf{1} \otimes x + \Delta' x$$

where

$$\Delta' x \in \bigoplus_{\substack{p+q=n \\ p,q\geq 1}} \mathcal{H}_{(p)} \otimes \mathcal{H}_{(q)} \tag{5.5}$$

is known as the *reduced coproduct*. Furthermore, the coassociativity of Δ allows to unambiguously define its iterates $\Delta_n : \mathcal{H} \to \mathcal{H}^{\otimes(n+1)}$ by setting

$$\Delta_0 = \mathrm{id}, \quad \Delta_n = (\mathrm{id} \otimes \Delta_{n-1})\Delta$$

and we have, for a homogenenous element $x \in \mathcal{H}_{(m)}$ of degree m

$$\Delta_n x = x \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} + \mathbf{1} \otimes x \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} + \cdots + \mathbf{1} \otimes \cdots \otimes x + \Delta'_n x$$

where now

$$\Delta'_n x \in \bigoplus_{\substack{p_1+\ldots+p_{n+1}=m \\ p_j \geq 1}} \mathcal{H}_{(p_1)} \otimes \cdots \otimes \mathcal{H}_{(p_{n+1})}$$

Remark 5.2.1. These facts about the iterated coproduct imply that the bialgebra $(\mathcal{H}, \Delta)$ is *conilpotent*, that is, for each homogeneous $x \in \mathcal{H}_{(m)}$ there is an integer $n \geq 1$ such that $\Delta'_n x = 0$ and we see that in fact $n = m$. △

Remark 5.2.2. From the above discusion we obtain in particular the inclusion

$$\Delta'_n \mathcal{H}_{(n+1)} \subset \mathcal{H}_{(1)}^{\otimes(n+1)},$$

that is, the n-fold reduced coproduct of a homogeneous element of degree $n + 1$ is a sum of $(n + 1)$-fold tensor products of homogeneous elements of degree 1. △

We recall that in general the dual space $\mathcal{H}^*$ carries an algebra structure given by the convolution product $\star$, dual to the coproduct Δ, defined by

$$\langle X \star Y, x \rangle = \langle X \otimes Y, \Delta x \rangle,$$

but in general $\mathcal{H}^*$ cannot be made into a coalgebra by dualising the product. In particular, $\mathcal{H}^*$ is commutative if and only if $\mathcal{H}$ is cocommutative. For a sequence of maps $X_1, \ldots, X_k \in \mathcal{H}^*$ we have the formula

$$X_1 \star \cdots \star X_k = m^{\otimes(k-1)}(X_1 \otimes \cdots \otimes X_k)\Delta_{k-1}. \tag{5.6}$$

A *character* on $\mathcal{H}$ is a linear map $X : \mathcal{H} \to \mathbb{R}$ such that

$$\langle X, xy \rangle = \langle X, x \rangle \langle X, y \rangle$$

for all $x, y \in \mathcal{H}$. An *infinitesimal character* (or derivation) on $\mathcal{H}$ is a linear map $\alpha : \mathcal{H} \to \mathbb{R}$ such that

$$\langle \alpha, xy \rangle = \langle \alpha, x \rangle \langle \varepsilon, y \rangle + \langle \varepsilon, x \rangle \langle \alpha, y \rangle.$$

We observe that $\langle X, \mathbf{1} \rangle = 1$ and $\langle \alpha, \mathbf{1} \rangle = 0$ for all $X \in G$ and $\alpha \in \mathfrak{g}$.

It is a known fact that the set G of characters on $\mathcal{H}$ is a group with unit ε and inverses $X^{-1} = S^* X = X \circ S$. The space $\mathfrak{g}$ of infinitesimal characters on $\mathcal{H}$ is a Lie algebra under the bracket $[\alpha,\beta] = \alpha \star \beta - \beta \star \alpha$. Moreover, there is an exponential map $\exp : \mathfrak{g} \to G$

$$\exp(\alpha) := \sum_{n=0}^{\infty} \frac{\alpha^{\star n}}{n!}$$

which is a bijection and its inverse is the map $\log : G \to \mathfrak{g}$

$$\log(X) := \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(X-\varepsilon)^{\star n}}{n}.$$

Remark 5.2.3. The above maps are actually defined over bigger spaces. These definitions make sense for maps $\alpha \in \hat{\mathfrak{g}}$ and $X \in \hat{G}$, where $\hat{\mathfrak{g}}$ is the Lie algebra of linear maps mapping $\mathbf{1} \to 0$ and $\hat{G}$ is the group of linear maps mapping $\mathbf{1} \to 1$. In fact, $\mathfrak{g}$ is a sub-Lie algebra of $\hat{\mathfrak{g}}$ and G is a sub-Lie group of $\hat{G}$. △

Remark 5.2.4. The conilpotency of $(\mathcal{H}, \Delta)$ implies that for each homogeneous element $x \in \mathcal{H}_{(n)}$ the above series defining exp and log terminate after a finite number of terms (precisely n, in fact). Therefore, for a general element $x \in \mathcal{H}$ and $\alpha \in \mathfrak{g}$ the value of $\langle \exp(\alpha), x \rangle$ is made up of a finite sum of finite sums, so there are no convergence issues involved whatsoever. △

The grading on $\mathcal{H}$ induces a grading on $\mathfrak{g}$ by restriction to $\mathcal{H}_{(n)}$ so we can write

$$\mathfrak{g} = \prod_{n=1}^{\infty} W_n,$$

i.e. the elements of $\mathfrak{g}$ correspond to formal series

$$\alpha = \sum_{n=1}^{\infty} \alpha_n$$

where $\alpha_n = \alpha|_{\mathcal{H}_{(n)}} \in W_n$.

Remark 5.2.5. The elements of $\mathfrak{g}$ correspond to infinite formal series and cannot in general be reduced to finite sums. The same goes for elements of the character group G and more in general for arbitrary elements of the dual space $\mathcal{H}^*$. This is one of the reasons why it is not possible to dualise the product on $\mathcal{H}$ to induce a coproduct on $\mathcal{H}^*$. △

The Baker–Campbell–Hausdorff formula below describes the group law on G in terms of the Lie bracketing on $\mathfrak{g}$ via the exponential map. See [60] for a proof.

Theorem 5.2.6 (Baker–Campbell–Hausdorff)**.** *Let* $\alpha, \beta \in \mathfrak{g}$*, then* $\log(\exp(\alpha) \star \exp(\beta)) \in \mathfrak{g}$.

We define the map $\mathrm{BCH}\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ by

$$\mathrm{BCH}(\alpha,\beta) := \log(\exp(\alpha) \star \exp(\beta)). \tag{5.7}$$

The main point of this theorem is that, even if one can compute explicitly all the terms appearing in the series defining $\log(\exp(\alpha) \star \exp(\beta))$ it is not immediately clear that each of these terms can be rewritten in terms of iterated commutators as the definition of $\mathfrak{g}$ requires. Another way to interpret this theorem is to say that there exists an element $\gamma = \mathrm{BCH}(\alpha, \beta) \in \mathfrak{g}$ such that $\exp(\alpha) \star \exp(\beta) = \exp(\gamma)$.

More can be said about the Lie series $\mathrm{BCH}(\alpha, \beta) \in \mathfrak{g}$ given by the previous theorem. In fact, it is possible to show that each term of the series is formed by iterated Lie brackets of α and β, where the first terms are

$$\mathrm{BCH}(\alpha, \beta) = \alpha + \beta + \frac{1}{2}[\alpha, \beta] + \frac{1}{12}[\alpha, [\alpha, \beta]] - \frac{1}{12}[\beta, [\alpha, \beta]] + \cdots ,$$

and the following terms are explicit but difficult to compute. Nevertheless, fully explicit formulas have been known since 1947 by Dynkin [35].

A more tractable description of the terms appearing in this series was given by Reutenauer in terms of permutations [103] and a proof in terms of Hopf algebras was given by Loday [80]. Let $\phi_k : (\mathcal{H}^*)^{\otimes k} \to \mathcal{H}^*$ be the linear map

$$\phi_k(\alpha_1 \otimes \cdots \otimes \alpha_k) = \sum_{\sigma \in S_k} a_\sigma \alpha_{\sigma(1)} \star \cdots \star \alpha_{\sigma(k)} \tag{5.8}$$

where S_k denotes the symmetric group of order k, and $a_\sigma := \frac{(-1)^{d(\sigma)}}{k} \binom{k-1}{d(\sigma)}^{-1}$ is a constant depending only on the *descent number* of the permutation $\sigma \in S_k$. One can then show that $\phi_k(\alpha_1, \ldots, \alpha_k) \in \mathfrak{g}$ if $\alpha_1, \ldots, \alpha_k \in \mathfrak{g}$, and that

$$\mathrm{BCH}_{(k)}(\alpha, \beta) = \sum_{i+j=k} \frac{1}{i!j!} \phi_k(\alpha^{\otimes i} \otimes \beta^{\otimes j}) \in W_k \tag{5.9}$$

where

$$\mathrm{BCH}(\alpha, \beta) = \sum_{k=1}^{\infty} \mathrm{BCH}_{(k)}(\alpha, \beta) \in \mathfrak{g}.$$

For example, we have that $\phi_1(\alpha) = \alpha$, $\phi_2(\alpha \otimes \beta) = \frac{1}{2}(\alpha \star \beta - \beta \star \alpha)$ and (omitting the $\star$ product)

$$\phi_3(\alpha_1 \otimes \alpha_2 \otimes \alpha_3) = \frac{1}{3}\alpha_1\alpha_2\alpha_3 - \frac{1}{6}(\alpha_2\alpha_1\alpha_3 + \alpha_1\alpha_3\alpha_2 + \alpha_2\alpha_3\alpha_1 + \alpha_3\alpha_1\alpha_2) + \frac{1}{3}\alpha_3\alpha_2\alpha_1.$$

Playing with these expressions we can observe that in fact $\phi_2(\alpha \otimes \beta) = \frac{1}{2}[\alpha, \beta]$ and

$$\phi_3(\alpha_1 \otimes \alpha_2 \otimes \alpha_3) = \frac{1}{6}[\alpha_1, [\alpha_2, \alpha_3]] - \frac{1}{6}[\alpha_3, [\alpha_1, \alpha_2]]$$

and then we recover the first three terms above after summation. This factorization can be formalized by means of the Dynkin operator and the Dynkin–Specht–Wever theorem [66]. The Dynkin operator $D : H^* \to \mathfrak{g}$ is defined as

$$D(x_1 \otimes \cdots \otimes x_k) = [x_1, \ldots [x_{k-1}, x_k] \ldots]$$

for $x_1, \ldots, x_k \in H^*_{(1)}$ and the DSW theorem asserts that a homogeneous element $X \in H^*_{(k)}$ is in $\mathfrak{g}$ if and only if $D(X) = kX$. A nice short proof of this fact in terms of Hopf algebras can be found in [118]. Since this is the case for each ϕ_k when evaluated in elements of $\mathfrak{g}$ we can replace every monomial $X_1 \star \cdots \star X_k$ by $\frac{1}{k}[X_1, \ldots, [X_{k-1}, X_k] \ldots]$. At this point there is still some more cancelling to be done using the antisymmetry of the Lie bracket and the Jacobi identity, but this already makes clear the connection between the two formulas. In any case, we will not use this and will only work with Reutenauer's formula.

From all these considerations we obtain

Lemma 5.2.7. *Let x be a homogeneous element of degree k such that*

$$\Delta'_{k-1} x = \sum_{(x)} x_{(1)} \otimes \cdots \otimes x_{(k)} \in \mathcal{H}_{(1)}^{\otimes k}$$

Then

$$\langle \phi_k(\alpha_1 \otimes \cdots \otimes \alpha_k), x \rangle = \sum_{(x)} \sum_{\sigma \in S_k} a_\sigma \prod_{j=1}^{k} \langle \alpha_{\sigma(j)}, x_{(j)} \rangle.$$

Proof. This follows directly from the definition of φ_k in eq. (5.8) together with eq. (5.6) and the fact that since $\alpha_j(\mathbf{1}) = 0$ we can write

$$\alpha_1 \star \cdots \star \alpha_k = m^{\otimes(k-1)}(\alpha_1 \otimes \cdots \otimes \alpha_k)\Delta'_{k-1}$$

instead. ◻

5.2.1 Nilpotent Lie algebras

From eq. (5.5) we have

Lemma 5.2.8. *For any $N \in \mathbb{N}$ the subspace*

$$\mathcal{H}_N := \bigoplus_{k=0}^{N} \mathcal{H}_{(k)} \subset \mathcal{H}$$

is a counital subcoalgebra of $(\mathcal{H}, \Delta, \varepsilon)$. The canonical projection $\pi_N : \mathcal{H} \to \mathcal{H}_N$ is coalgebra epimorphism.

It follows that we can consider the dual algebra $(\mathscr{H}_N^*, \star, \varepsilon)$ and the corresponding truncated Lie algebra

$$\mathfrak{g}^N = \bigoplus_{k=1}^{N} W_k$$

and Lie group $G^N = \exp(\mathfrak{g}^N)$. There is a canonical injection $\iota_N := \pi_N^* : \mathscr{H}_N^* \to \mathscr{H}^*$ such that $\langle \iota_N X, x\rangle = 0$ if $|x| > N$ so when working with elements of $\mathscr{H}_N^*$ or $\mathfrak{g}^N$ we will always assume that they satisfy this property. There are also restricted exponential and logarithm maps $\exp_N : \mathfrak{g}^N \to G^N$ and $\log_N : G^N \to \mathfrak{g}^N$ defined by the truncated sums

$$\exp_N(\alpha) := \sum_{n=0}^{N} \frac{\alpha^{\star n}}{n!}, \quad \log_N(X) := \sum_{n=1}^{N} (-1)^{n+1} \frac{(X-\varepsilon)^{\star n}}{n}. \tag{5.10}$$

Proposition 5.2.9. *The orthogonal subspace*

$$\mathscr{K}^N := \mathscr{H}_N^\perp = \{\, X \in \mathscr{H}^* : \langle X, x\rangle = 0, \forall x \in \mathscr{H}_N \}$$

is an ideal of the algebra $(\mathscr{H}^*, \star, \varepsilon)$. In particular, the quotient algebra $\mathscr{H}^*/\mathscr{K}^N$ is isomorphic to $(\mathscr{H}_N^*, \star, \varepsilon)$.

Remark 5.2.10. The canonical inclusion ι_N is an algebra monomorphism, being the dual map to a coalgebra epimorphism. Moreover, it is such that if $\varpi_N : \mathscr{H}^* \to \mathscr{H}_N^*$ is the canonical projection then $\varpi_N \circ \iota_N = \mathrm{id}_{\mathscr{H}_N^*}$. Note however that ι_N does not map G^N into G or any subgroup of it. In fact, we see that ι_N maps G^N to $\hat{G}$ and $\mathfrak{g}^N$ to $\hat{\mathfrak{g}}$ defined inside Remark 5.2.3. Regarding $\mathfrak{g}^N$ see however the following proposition. △

Proposition 5.2.11. *The canonical inclusion ι_N maps $\mathfrak{g}^N$ to $\mathfrak{g}$.*

Proof. We already know that $\iota_N : \mathfrak{g}^N \to \hat{\mathfrak{g}}$ so it only suffices to check that given $\alpha \in \mathfrak{g}^N$ we have that $\iota_N \alpha$ is an infinitesimal character. Let $x, y \in \mathscr{H}$, then

$$\begin{aligned}
\langle \iota_N \alpha, xy\rangle &= \langle \alpha, \pi_N(xy)\rangle \\
&= \sum_{n=0}^{N} \sum_{j=0}^{n} \langle \alpha, x_j y_{n-j}\rangle \\
&= \sum_{n=0}^{N} \sum_{j=0}^{n} (\langle \alpha, x_j\rangle\langle \varepsilon, y_{n-j}\rangle + \langle \varepsilon, x_j\rangle\langle \alpha, y_{n-j}\rangle) \\
&= \sum_{n=0}^{N} (\langle \alpha, x_n\rangle\langle \varepsilon, y_0\rangle + \langle \varepsilon, x_0\rangle\langle \alpha, y_n\rangle) \\
&= \langle \alpha, \pi_N x\rangle\langle \varepsilon, y\rangle + \langle \varepsilon, x\rangle\langle \alpha, \pi_N y\rangle,
\end{aligned}$$

hence $\iota_N \alpha \in \mathfrak{g}$. □

Remark 5.2.12. A similar statement cannot hold for G^N. For instance take $X \in G^N$ and an element $x \in \mathcal{H}_N \setminus \{\mathbf{1}\}$ such that $\langle X, x\rangle \neq 0$. Without loss of generality we may suppose that x is homogeneous. Take k large enough so that $k|x| \geq N+1$. Then

$$0 = \langle \iota_N X, x^k\rangle \neq \langle X, x\rangle^k.$$

△

Remark 5.2.13. All the above considerations work, with minor modifications, if linear maps are allowed to take values in an arbitrary commutative algebra instead of the ground field (in this case $\mathbb{R}$). That is, we may consider instead of the dual space $\mathcal{H}^*$, the space $\mathscr{L}(\mathcal{H}, A)$ of linear maps from $\mathcal{H}$ to a commutative unital algebra A. Even though this level of generality may seem superfluous, it could be interesting since the added structure may reveal some further connections with renormalisation via the Birkhoff decomposition of characters. Since our aim is to define paths taking values in the group G^N satisfying some extra properties, for example, we could make them depend on an extra parameter $\varepsilon > 0$ and A then could be taken to be the algebra of Laurent series in this extra parameter. See [89] for further details. △

Regarding the properties of the Lie algebra $\mathfrak{g}^N$ we can show that

Proposition 5.2.14. *The finite-dimensional Lie algebra $\mathfrak{g}^N$ is step N nilpotent.*

Proof. We recall from Chapter 2 that nilpotency means that the lower central series $\mathfrak{g}_1^N \supset \mathfrak{g}_2^N \supset \cdots$ defined inductively $\mathfrak{g}_1^N := \mathfrak{g}^N$, $\mathfrak{g}_{k+1}^N := [\mathfrak{g}^N, \mathfrak{g}_k^N]$ terminates after a finite number of steps, that is, there exists $k_0 \in \mathbb{N}$ such that $\mathfrak{g}_k^N = \{0\}$ for all $k \geq k_0$. The smallest number such that this happens is called the nilpotency step of $\mathfrak{g}^N$. Thus, we have to show that $\mathfrak{g}_{N+1}^N = \{0\}$ but $\mathfrak{g}_N^N \neq \{0\}$.

We first prove that $[W_k, W_j] \subset W_{k+j}$. Take $\alpha \in W_k, \beta \in W_j$ and an element $x \in \mathcal{H}$. Then

$$\langle [\alpha, \beta], x\rangle = \langle \alpha \otimes \beta - \beta \otimes \alpha, \Delta x\rangle = \sum_{(x)} \langle \alpha, x_{(1)}\rangle\langle \beta, x_{(2)}\rangle - \langle \beta, x_{(1)}\rangle\langle \alpha, x_{(2)}\rangle$$

where the only surviving terms are those with $|x_{(1)}| = k$ and $|x_{(2)}| = j$ or $|x_{(1)}| = j$ and $|x_{(2)}| = k$. In any case this is only possible if x is homogeneous of degree $k+j$.

Now, observe that using this fact and induction one can show that

$$\mathfrak{g}_k^N \subset \bigoplus_{j=k}^N W_j$$

thus $\mathfrak{g}_N^N \subset W_N$ and $\mathfrak{g}_{N+1}^N = \{0\}$. □

As a by-product of this proposition we obtain

Corollary 5.2.15. *The centre of* $\mathfrak{g}^N$ *is* W_N.

Finally we remark that the Baker–Campbell–Hausdorff is also valid in $\mathfrak{g}^N$ with the additional property that the sum is now finite due to the nilpotency. We define an operator $\mathrm{BCH}_N : \mathfrak{g}^N \times \mathfrak{g}^N \to \mathfrak{g}^N$ by

$$\mathrm{BCH}_N(\alpha, \beta) := \sum_{k=1}^{N} \mathrm{BCH}_{(k)}(\alpha, \beta) \tag{5.11}$$

where $\mathrm{BCH}_{(k)} \in W_k$ is was defined in eq. (5.9).

5.2.2 Homogeneous norms

Let $\mathfrak{L}$ be a finite-dimensional Lie algebra. A family of dilations on $\mathfrak{L}$ is a family $(\Omega_r)_{r>0}$ of automorphisms of $\mathfrak{L}$ such that $\Omega_r \Omega_s = \Omega_{rs}$. A homogeneous group is a connected simply connected Lie group whose Lie algebra is endowed with a family of dilations. If G is a homogeneous group, the map $\exp \circ \Omega_r \circ \log$ is a group automorphism of G and we also call them dilations.

Definition 5.2.16. *The element* $X \in \mathfrak{L}$ *is said to be an eigenvector of the dilation* Ω *with eigenvalue* $\alpha \in \mathbb{R}$ *if* $\Omega_r X = r^\alpha X$ *for all* $r > 0$*. For an eigenvalue* $\alpha \in \mathbb{R}$ *the eigenspace* E_α *is the subspace of* $\mathfrak{L}$ *spanned by all the eigenvectors of* Ω *with eigenvalue* α.

Since Ω_r is a Lie homomorphism we have that $[E_\alpha, E_\beta] \subset E_{\alpha+\beta}$.

Lemma 5.2.17. *A family of operators* (Ω_r) *is a dilation if and only if* $\Omega_r = \mathrm{e}^{\log(r)A}$ *for some matrix* A.

Proof. It suffices to observe that $f(r) := \Omega_{\mathrm{e}^r}$ satisfies $f(r+s) = f(r) \circ f(s)$. □

Thus, a dilation can only have a finite number of eigenvalues which correspond to eigenvalues of the matrix A. In the sequel we order the spectrum of Ω (or A) increasingly, i.e. $\alpha_1 \leq \cdots \leq \alpha_n$ where $n = \dim \mathfrak{L}$. Since if Ω is a dilation then $\tilde{\Omega}_r = \Omega_{r^\alpha}$ is also a dilation, we may and do suppose that $\alpha_1 \geq 1$.

In the following, we assume the matrix A to be diagonalizable. In this case, we may fix a basis $\{X_1, \ldots, X_n\}$ of $\mathfrak{L}$ such that $AX_j = \alpha_j X_j$. We use this basis to obtain a norm $\|\cdot\|$ on $\mathfrak{L}$ by requiring that this basis is orthonormal.

Definition 5.2.18. *Ahomogeneous norm on a homogeneous group* G *is a continuous function* $|\cdot| : G \to \mathbb{R}_+$ *which is of class* C^∞ *on* $G \setminus \{\mathbf{1}\}$ *and such that* $|X^{-1}| = |X|$*, and* $|\Omega_r X| = r|X|$*. The homogeneous norm* $|\cdot|$ *is said to be sub-additive if* $|XY| \leq |X| + |Y|$.

In case the homogeneous norm is sub-additive, we can induce a left-invariant metric on G by setting $\rho(X,Y) = |Y^{-1}X|$.

Lemma 5.2.19 ([45])**.** *Suppose G is a homogeneous group with Lie algebra $\mathfrak{L}$. Then there exist constants $C_1, C_2 > 0$ such that*

$$C_1\|\log X\| \leq |X| \leq C_2\|\log X\|^{1/\alpha_n}$$

for all $X \in G$ with $|X| \leq 1$.

A simple consequence of this lemma is the following

Corollary 5.2.20. *All homogeneous norms on G are equivalent.*

We can build a dilation on $\mathfrak{g}^N$ as follows: for $x \in \mathcal{H}_{(k)}$ set $\Omega_r x = r^k x$ and transpose this map to $\mathcal{H}^*$ by setting $\langle \Omega_r X, x\rangle = \langle X, \Omega_r x\rangle$.

Proposition 5.2.21. *The maps Ω_r are algebra automorphisms of $\mathcal{H}^*$.*

Proof. The map Ω_r is a coalgebra morphism of $\mathcal{H}$. Indeed,

$$\Delta(\Omega_r x) = r^{|x|}\Delta x = (\Omega_r \otimes \Omega_r)\Delta x$$

by Equation (5.5). □

In fact, the maps Ω_r are bialgebra automorphisms of $\mathcal{H}$, hence Hopf algebra automorphisms. Therefore, we obtain a dilation on $\mathfrak{g}^N$ by simple restriction, and we remark that the spaces W_k act as eigenspaces for Ω, with k as the associated eigenvalue.

We fix norms $|||\cdot|||_k$ on each of the dual spaces $\mathcal{H}^*_{(k)}$ which are compatible with the convolution product in the sense that

$$|||X \star Y|||_{n+m} \leq |||X|||_n |||Y|||_m \tag{5.12}$$

for all $X \in \mathcal{H}^*_{(n)}$ and $Y \in \mathcal{H}^*_{(m)}$. We obtain a homogeneous norm on G^N by setting

$$|X| := \max_{k=1,\dots,N}(k!|||X_k|||_k)^{1/k}, \tag{5.13}$$

where

$$X = \varepsilon + \sum_{k=1}^{N} X_k$$

The following formula for the components of the convolution product between two linear maps follows directly from eq. (5.5).

Lemma 5.2.22. *Given linear maps $X, Y \in \mathcal{H}_N^*$, write*

$$X = \varepsilon + \sum_{k=1}^{N} X_k, \quad Y = \varepsilon + \sum_{k=1}^{N} Y_k.$$

Then

$$X \star Y = \varepsilon + \sum_{k=1}^{N} \sum_{j=1}^{k} X_j \star Y_{k-j}.$$

Proposition 5.2.23. *The group $(G^N, |\cdot|)$ is homogeneous with $|\cdot|$ sub-additive.*

Proof. We only need to prove that the norm defined in Equation (5.13) is sub-additive, the other properties being clear. By Lemma 5.2.22 and the compatibility of the norms (5.12) we have that

$$\begin{aligned}
|||(X \star Y)_k|||_k &\leq \sum_{j=1}^{k} |||X_j|||_j |||Y_{k-j}|||_{k-j} \\
&\leq \frac{1}{k!} \sum_{j=1}^{k} \binom{k}{j} |X|^j |Y|^{k-j} \\
&= \frac{1}{k!} (|X| + |Y|)^{1/k}
\end{aligned}$$

whence the result. □

In particular we obtain a metric ρ_N on G^N which is left-invariant and such that the metric space (G^N, ρ_N) is complete. This distance may be explicitly computed by Equation (5.13) as

$$\rho_N(X, Y) = \max_{k=1,\dots,N} \left(k! |||(Y^{-1} \star X)|||_k \right)^{1/k}$$

Remark 5.2.24. In view of Corollary 5.2.20 we may obtain bounds over the distance $\rho_N(X, Y)$ by bounding first $p(Y^{-1} \star X)$ for any homgeneous norm p on G. The importance of $|\cdot|$ resides in that we know it to be sub-additive by Proposition 5.2.23 so we obtain a distance. On some concrete cases there might be other sub-additive homogenenous norms defined on G^N but we choose to work with the one defined in eq. (5.13) since it is closely related to rough paths, see Definition 5.3.1 and Proposition 5.4.3 △

5.3 Construction of Rough paths

As in the previous section, we fix a locally-finite graded connected Hopf algebra $\mathcal{H}$. Without loss of generality we assume that $\dim \mathcal{H}_{(1)} = d$ and we fix a linear basis $\{e_1, \dots, e_d\}$ of it. We also fix a number $\gamma \in]0,1[$ and let $N := \lfloor \gamma^{-1} \rfloor$ be the biggest integer such that $N\gamma \leq 1$.

Definition 5.3.1. *A* γ-rough path *is a path* $\mathbb{X} : [0,1] \to G^N$ *such that* $\mathbb{X}_0 = \varepsilon$ *which is* γ*-Hölder with respect to the metric* ρ_N *defined by the homogeneous norm in eq.* (5.13). *Setting* $x_t^i = \langle \mathbb{X}_t, e_i \rangle$ *we say that* $\mathbb{X}$ *is a* γ*-rough path over* $(x^1, \ldots, x^d)$.

Remark 5.3.2. By specializing this definition to different values of $\mathcal{H}$ we recover both *geometric rough paths* [86] where $\mathcal{H}$ is the shuffle Hopf algebra over an alphabet and *branched rough paths* [51] where $\mathcal{H}$ is the Butcher–Connes–Kreimer Hopf algebra on decorated non-planar rooted trees. △

Remark 5.3.3. The classical definitions of rough paths of various types consider functions $X : [0,1]^2 \to G$ with values in an appropriate group, satisfying Chen's rule

$$X_{su} * X_{ut} = X_{st}$$

and an analytical estimate. It turns out that one can reduce X to a one-variable path $\mathbb{X} : [0,1] \to G$ by noting that the above equation implies that $X_{st} = (X_{0s})^{-1} * X_{0t}$. The analytical estimate can be seen to be equivalent to requiring that the resulting path $\mathbb{X}_t = X_{0t}$ is γ-Hölder with respect to the corresponding homogeneous metric ρ_N. △

We now come to the problem of existence. Our construction of a rough path in the sense of Definition 5.3.1 over an arbitrary collection of γ-Hölder paths $(x^1, \ldots, x^d)$ relies in the following extension theorem. We note that the proof follows very closely the approach of Lyons-Victoir [88].

Theorem 5.3.4 (Character extension)**.** *Let* $1 \leq n \leq N-1$ *and* $\gamma \in]0,1[$ *such that* $(n+1)\gamma < 1$. *Suppose we have a* γ*-Hölder path* $\mathbb{X} : [0,1] \to (G^n, \rho_n)$. *There is a* γ*-Hölder path* $\tilde{\mathbb{X}} : [0,1] \to (G^{n+1}, \rho_{n+1})$ *extending* $\mathbb{X}$, *i.e. such that* $\tilde{\mathbb{X}}\big|_{\mathcal{H}_n} = \mathbb{X}$.

Proof. The construction of $\tilde{\mathbb{X}}$ is made in two steps.

Step 1. Let $D = \{t_k^m := 2^{-m}k \mid m \in \mathbb{N}_0, k = 0, \ldots, 2^m\}$ be the dyadics in $[0,1]$. Set $X_{st} = (\mathbb{X}_s^n)^{-1} \star \mathbb{X}_t^n \in G^n$ and $L_{st} = \log_n(X_{st}) \in \mathfrak{g}^n$ where $\log_n$ was defined in eq. (5.10). Then, the BCH formula eq. (5.11) and Chen's rule imply that

$$L_{st} = \mathrm{BCH}_n(L_{su}, L_{ut}) = L_{su} + L_{ut} + \mathrm{BCH}'_n(L_{su}, L_{ut}). \tag{5.14}$$

We look for $Z_{st} \in W_{n+1}$ such that the exponential $X_{st}^{n+1} = \exp_{n+1}(\iota_{n+1} L_{st} + Z_{st})$ still satisfies Chen's rule. To this effect, we first define Z on the dyadics, starting from a fixed element $w \in W_{n+1}$ putting $Z_{0,1} = w$ and

$$Z_{t_{2k}^m, t_{2k+1}^m} = Z_{t_{2k+1}^m, t_{2k+2}^m} := \frac{1}{2} Z_{t_k^{m-1}, t_{k+1}^{m-1}} - \frac{1}{2} \mathrm{BCH}_{(n+1)}(\iota_{n+1} L_{t_{2k}^m, t_{2k+1}^m}, \iota_{n+1} L_{t_{2k+1}^m, t_{2k+2}^m}) \tag{5.15}$$

where ι_{n+1} is the canonical inclusion and we have used Proposition 5.2.11 so that the right-hand side is well defined. To ease notation in the following we identify L with $\iota_{n+1}L$ where appropriate. Note that with this $Z_{st} \in W_{n+1}$ for each pair of consecutive dyadics.

We now look to extend this definition to more general pairs of dyadics $s,t \in D$. Set $Y_{0,1} := \exp_{n+1}(L_{0,1} + w)$. If $s = t^m_{2k}$, $u = t^m_{2k+1}$ and t^m_{2k+2} are consecutive dyadics then we define

$$Y_{su} := \exp_{n+1}(L_{su} + Z_{su}), \quad Y_{ut} := \exp_{n+1}(L_{ut} + Z_{ut})$$

and note that by eq. (5.15) we have

$$\begin{aligned}\log_{n+1}(Y_{su} \star Y_{ut}) &= L_{su} + L_{ut} + \mathrm{BCH}'_{n+1}(L_{su}, L_{ut}) + Z_{su} + Z_{ut} \\ &= L_{su} + L_{ut} + \mathrm{BCH}'_{n}(L_{su}, L_{ut}) + Z_{st} \\ &= L_{st} + Z_{st}\end{aligned}$$

by eq. (5.14), so that

$$Y_{su} \star Y_{ut} = Y_{st}.$$

We have also used the fact that ι_{n+1} is an algebra morphism for the $\star$ product and that W_{n+1} is in the centre of g^{n+1} by Corollary 5.2.15. Therefore, we may set

$$Y_{t^m_k, t^m_j} := Y_{t^m_k, t^m_{k+1}} \star Y_{t^m_{k+1}, t^m_{k+2}} \star \cdots \star Y_{t^m_{j-1}, t^m_j}$$

so that the identity $Y_{t^m_i, t^m_j} \star Y_{t^m_j, t^m_k} = Y_{t^m_i, t^m_k}$ is valid for any $0 \le i < j < k \le 2^m$.

Step 2. In order to have a γ-rough path, Definition 5.3.1 requires us to construct a γ-Hölder path with values in G^{n+1}, and for this we will use Lemma 5.1.6. Set

$$a_m := 2^{m(n+1)\gamma} \max_{k=0,\ldots,2^m-1} \left|\!\left|\!\left| Z_{t^m_k, t^m_{k+1}} \right|\!\right|\!\right|_{n+1}.$$

Then, if v is a basis element in $\mathcal{H}_{(n+1)}$ we have by Lemma 5.2.7 and eq. (5.9), for $s = t^m_k$, $u = t^m_{k+1}$ and $t = t^m_{k+2}$, that

$$|\langle \mathrm{BCH}_{(n+1)}(L_{su}, L_{ut}), v\rangle| \le \sum_{(v)} \sum_{i+j=n+1} \frac{1}{i!j!} \sum_{\sigma \in S_{n+1}} |a_\sigma| \prod_{p=1}^{i} |\langle L_{su}, v_{\sigma(p)}\rangle| \prod_{q=i+1}^{n+1} |\langle L_{ut}, v_{\sigma(q)}\rangle|.$$

Now, since $v_{(j)} \in \mathcal{H}_{(1)}$ for all $j = 1, \ldots, n+1$ we actually have that

$$|\langle L_{su}, v_{(j)}\rangle| \le \sum_{k=1}^{d} |x^k_u - x^k_s||v^k_{(j)}| \le 2^{-m\gamma} \sum_{k=1}^{d} |v^k_{(j)}|$$

for some coefficients $\upsilon^k_{(j)} \in \mathbb{R}$, and we have a similar estimate for L_{ut} instead of L_{su}. Therefore we obtain that

$$\left\lvert\!\left\lvert\!\left\lvert \mathrm{BCH}_{(n+1)}(L_{su}, L_{ut}) \right\rvert\!\right\rvert\!\right\rvert_{n+1} \leq C 2^{-m(n+1)\gamma}$$

where the constant

$$C = \max_{\upsilon} \sum_{(\upsilon)} \sum_{i+j=n+1} \frac{1}{i!j!} \sum_{\sigma \in S_{n+1}} |a_\sigma| \sum_{k_1,\dots,k_{n+1}=1}^{n+1} \prod_{\ell=1}^{n+1} |\upsilon^{k_\ell}_{(\ell)}|.$$

Therefore, from eq. (5.15) we get

$$2^{-(m+1)(n+1)\gamma} a_{m+1} \leq 2^{-m(n+1)\gamma} + C 2^{-m(n+1)\gamma},$$

hence there is another constant $C > 0$ such that

$$a_m \leq C \sum_{j=0}^{m-1} 2^{-j(1-(n+1)\gamma)}.$$

Since we are in the regime where $(n+1)\gamma < 1$ we obtain that

$$\sup_{m \geq 0} a_m \leq \frac{C}{2 - 2^{(n+1)\gamma}}.$$

To conclude, we observe that

$$L \mapsto \sum_{k=1}^{n} \lvert\!\lvert\!\lvert L \rvert\!\rvert\!\rvert_k^{1/k}$$

defines a norm on $\mathscr{H}^*_n$, thus by Lemma 5.2.19 there's a constant C_1 such that

$$\lvert\!\lvert\!\lvert L_{st} \rvert\!\rvert\!\rvert_k \leq \|L_{st}|^k \leq |(\mathbb{X}^n_s)^{-1} \star \mathbb{X}^n_t|^k = \rho_n(\mathbb{X}^n_s, \mathbb{X}^n_t)^k \leq C_1 2^{-mk\gamma}$$

for all $k = 1, \dots, n$ and $s = t^m_j$, $t = t^m_{j+1}$ since $\mathbb{X}^n$ is γ-Hölder with respect to ρ_n. This and the previous estimate provide the bound

$$|Y_{t^m_j, t^m_{j+1}}| \leq 2^{-m\gamma}.$$

By Chen's rule, the path $\mathbb{Y} : D \to G^{n+1}$ defined by $\mathbb{Y}_{t^m_j} := Y_{0,t^m_j}$ satisfies

$$\rho_{n+1}(\mathbb{Y}_{t^m_j}, \mathbb{Y}_{t^m_{j+1}}) \leq 2^{-m\gamma},$$

thus by Lemma 5.1.6 we obtain a path $\tilde{\mathbb{X}} : [0,1] \to G^{n+1}$. □

Remark 5.3.5. Our construction depends on a finite number of choices, namely the elements w used to start the recursion in eq. (5.15). Different choices give, *a priori*, different outcomes. △

Corollary 5.3.6. *Given $\gamma \in]0,1[$ and a collection of γ-Hölder paths $x^i : [0,1] \to \mathbb{R}$, there exists a γ-rough path $\mathbb{X}$ over $(x^1, \ldots, x^d)$ in the sense of Definition 5.3.1.*

Proof. We start with the following observation: for $n = 1$, the group $G^1 \subset \mathcal{H}^*_{(1)}$ is abelian, and isomorphic to the additive group $\mathcal{H}^*_{(1)}$. Indeed, let $X, Y \in G^1$ and $x \in \mathcal{H}_{(1)}$. Then, as $\Delta x = x \otimes \mathbf{1} + \mathbf{1} \otimes x$ by the grading, we have that

$$\langle X \star Y, x \rangle = \langle X, x \rangle + \langle Y, x \rangle,$$

that is, $X \star Y = X + Y$. Moreover, in $\mathcal{H}_1$ the product $xy = 0$. Therefore, we may set $\langle \mathbb{X}^1_t, e_i \rangle := x^i_t$ where $\{e_1, \ldots, e_d\}$ is a basis of $\mathcal{H}_{(1)}$ and this path is γ-Hölder with respect to ρ_1.

By Theorem 5.3.4 there is a γ-Hölder path $\mathbb{X}^2 : [0,1] \to (G^2, \rho_2)$ extending $\mathbb{X}^1$ so in particular $\langle \mathbb{X}^2_t, e_i \rangle = x^i_t$ also. Continuing in this way we obtain succesive γ-Hölder extensions $\mathbb{X}^3, \ldots, \mathbb{X}^N$ and we set $\mathbb{X} := \mathbb{X}^N$. □

The following result already appeared has already been proved in the case where the underlying Hopf algebra $\mathcal{H}$ is *combinatorial* by Curry, Ebrahimi-Fard, Manchon and Munthe-Kaas in [28]. We remark that their proof works without modifications in our context so we have

Theorem 5.3.7. *Let $\mathbb{X}$ be a γ-rough path. There exists a path $\hat{\mathbb{X}} : [0,1] \to G$ such that $\langle \hat{\mathbb{X}}^{-1}_s \star \hat{\mathbb{X}}_t, x \rangle \lesssim |t - s|^{\gamma|x|}$.*

Remark 5.3.8. The proof of this theorem is similar to that of Theorem 5.3.4. The main difference is that in the case of the components of lower regularity the sum defining A above does not converge in $\mathfrak{g}$ so this path has to be defined some other way, and this is the content of the proof of Theorem 5.3.4. △

Remark 5.3.9. In view of Theorem 5.3.7 we can replace the truncated group in Definition 5.3.1 by the full group of characters G. What this means is that γ-rough paths are uniquely defined once we fix the first N levels and since $\mathcal{H}$ is locally finite, this amounts to a finite number of choices. △

5.4 Applications

We now apply Theorem 5.3.4 to various kinds of Hopf algebras in order to link this results with the contexts already existing in the literature.

5.4.1 Geometric rough paths

In this setting we fix an alphabet $A := \{1, \ldots, d\}$. As a vector space H corresponds to the linear span of the *free monoid* $\mathrm{M}(A)$ generated by A. The product on H is the shuffle product $\sqcup\!\sqcup : H \otimes H \to H$ defined recursively by $\mathbf{1} \sqcup\!\sqcup v = v \sqcup\!\sqcup \mathbf{1} = v$ for all $v \in H$, where $\mathbf{1} \in \mathrm{M}(A)$ is the unit for the monoid operation, and

$$(au \sqcup\!\sqcup bv) = a(u \sqcup\!\sqcup bv) + b(au \sqcup\!\sqcup v)$$

for all $u, v \in H$ and $a, b \in A$, where au and bv denote the product of the letters a, b with the words u, v in $\mathrm{M}(A)$.

The coproduct $\bar{\Delta} : H \to H \otimes H$ is obtained by *deconcatenation* of words,

$$\bar{\Delta}(a_1 \cdots a_n) = a_1 \cdots a_n \otimes \mathbf{1} + \mathbf{1} \otimes a_1 \cdots a_n + \sum_{k=1}^{n-1} a_1 \cdots a_k \otimes a_{k+1} \cdots a_n.$$

It turns out that $(H, \cdot, \bar{\Delta})$ is a commutative unital Hopf algebra, and $(H, \bar{\Delta})$ is the cofree coalgebra over the linear span of A. The antipode is the linear map $S : H \to H$ given by

$$S(a_1 \cdots a_n) = (-1)^n a_n \ldots a_1.$$

Finally, we recall that H is graded by the length $\ell(a_1 \cdots a_n) = n$ and it is also connected. The homogeneous components $H_{(n)}$ are spanned by the sets $\{a_1 \cdots a_n : a_i \in A\}$. See Chapter 2 for further details.

Definition 5.3.1 specialises in this case to *geometric rough paths* (GRP) as defined in [57] (see just below for the precise definition) and Theorem 5.3.4 coincides with [88, Theorem 6].

Definition 5.4.1. *Let $\gamma \in]0, 1[$ and set $N := \lfloor \gamma^{-1} \rfloor$. A* geometric rough path *is a map $X : [0, 1]^2 \to G^N$ such that $X_{tt} = \varepsilon$, it satisfies Chen's rule*

$$X_{st} = X_{su} \star X_{ut}$$

for all $s, u, t \in [0, 1]$ and the analytic bound $|\langle X_{st}, v\rangle| \lesssim |t - s|^{\gamma\ell(v)}$ for all $v \in H_N$.

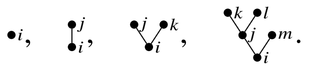
Remark 5.4.2. We note that in Definition 5.3.1 the path X (or rather $\mathbb{X}_t = X_{0t}$) takes values in the truncated character group G^N and not the full group G. This is in contrast with the definition of branched rough paths (Definition 5.4.4 below) where paths take values on the full Butcher group.

In [57] this definition is stated with G replacing the truncated group G^N and then it is said that in fact this is equivalent to Definition 5.3.1 (specialised to H) by invoking an "extension theorem" in [86]. The fact is that there is a confusion in terminology since what Lyons calls a "multiplicative functional" refers to Chen's rule and not the character property defining G, which he calls group-like elements (see [86, Definition 2.1.1, Definition 2.1.2]). In other words, Lyons shows that a collection of linear maps on H_N satisfying Chen's rule can be uniquely extended to a familiy of linear functionals on H keeping this identity. This is clearly weaker than what is required for the previous equivalence to be true since the extension given by Lyons may fail to lie in G. We correct this in Theorem 5.3.7. $\triangle$

Proposition 5.4.3. *A path $X : [0,1]^2 \to G^N$ is a geometric rough path if and only if $\mathbb{X}_t := X_{0t}$ is a rough path in the sense of Definition 5.3.1.*

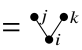
5.4.2 Branched rough paths

Let $\mathfrak{T}$ be the collection of all non-planar non-empty rooted trees with nodes decorated by $\{1,\dots,d\}$. Elements of $\mathfrak{T}$ are written as 2-tuples $\tau = (T,c)$ where T is a non-planar tree with node set N_T and edge set E_T, and $c : N_T \to \{1,\dots,d\}$ is a function. Edges in E_T are oriented away from the root, but this is not reflected in our graphical representation. Examples of elements of $\mathfrak{T}$ include the following



For $\tau \in \mathfrak{T}$ write $|\tau| = \#N_T$ for its number of nodes. Also, given an edge $e = (x,y) \in E_T$ we set $s(e) = x$ and $t(e) = y$. There is a natural partial order relation on N_T where $x \le y$ if and only if there is a path in T from the root to y containing x.

We denote by $\mathfrak{F}$ the collection of decorated rooted forests and we let $\mathcal{H}$ denote the vector space spanned by $\mathfrak{F}$. There is a natural commutative and associative product on $\mathfrak{F}$ given by disjoint union of forests, where the empty forest $\mathbf{1}$ acts as the unit. Then, $\mathcal{H}$ is the free commutative algebra over (the vector space spanned by) $\mathfrak{T}$ which is graded by $|\tau_1\cdots\tau_k| = |\tau_1| + \cdots + |\tau_k|$. Given a label $i \in \{1,\dots,d\}$ and a forest $\tau = \tau_1\cdots\tau_k$ we denote by $[\tau_1\cdots\tau_k]_i$ the tree obtained by grafting each of the trees $\tau_1,\dots,\tau_k$ to a new root labelled i, e.g.

$[\bullet_j]_i =$  , $\quad [\bullet_j\bullet_k]_i =$  .

The decorated Connes–Kreimer coproduct [24, 51] is the unique algebra morphism $\Delta : \mathfrak{F} \to \mathfrak{F} \otimes \mathfrak{F}$ such that

$$\Delta[\tau]_i = [\tau]_i \otimes \mathbf{1} + (\mathrm{id} \otimes [\cdot]_i)\Delta\tau.$$

This coproduct admits a representation in terms of *cuts*. An *admissible cut* C of a tree T is a non-empty subset of E_T such that any path from any vertex of the tree to the root contains at most one edge from C; we denote by $\mathfrak{A}(T)$ the set of all admissible cuts of the tree T. Any admissible cut C containing k edges maps a tree T to a forest $C(T) = T_1\cdots T_{k+1}$ obtained by removing each of the edges in C. Observe that only one of the remaining trees $T_1,\dots,T_{k+1}$ contains the root of T, which we denote by $R^C(T)$; the forest formed by the other k factors is denoted by $P^C(T)$. This naturally induces a map on decorated trees by considering cuts of the underlying tree, and restriction of the decoration map to each of the rooted subtrees $T_1, \ldots, T_{k+1}$. Then,

$$\Delta\tau = \tau \otimes \mathbf{1} + \mathbf{1} \otimes \tau + \sum_{C \in \mathfrak{A}(\tau)} P^C(\tau) \otimes R^C(\tau).$$

This, together with the counit map $\varepsilon : \mathcal{F} \to \mathbb{R}$ such that $\varepsilon(\tau) = 1$ if and only if $\tau = \mathbf{1}$ endows $\mathcal{F}$ with a connected graded commutative non-cocommutative bialgebra structure, hence a Hopf algebra structure [89].

As before we denote by $\mathcal{H}^*$ the linear dual of $\mathcal{H}$ which is an algebra via the convolution product $\langle X \star Y, \tau \rangle = \langle X \otimes Y, \Delta\tau \rangle$ and we denote by $\mathcal{G}$ the set of characters on $\mathcal{H}$, that is, linear functionals $X \in \mathcal{H}^*$ such that $\langle X, \sigma \cdot \tau \rangle = \langle X, \sigma \rangle \langle X, \tau \rangle$. For each $n \in \mathbb{N}$ the finite-dimensional vector space $\mathcal{H}_n$ spanned by the set $\mathcal{F}_n$ of forests with at most n nodes is a subcoalgebra of $\mathcal{H}$, hence its dual is an algebra under the convolution product, and we let $\mathcal{G}_n$ be the set of characters on $\mathcal{H}_n$. Likewise, we denote by $\mathcal{B}$ the vector space spanned by $\mathcal{T}$ and for each $n \in \mathbb{N}$ we denote by $\mathcal{B}_n$ the finite-dimensional vector space spanned by the set $\mathcal{T}_n$ of trees with at most n edges.

We recall the definition of branched rough paths from [51].

Definition 5.4.4. *Fix $\gamma > 0$ and let $N = \lfloor \gamma^{-1} \rfloor$. A* branched rough path *is a map $X : [0,1]^2 \to \mathcal{G}_N$ such that $X_{tt} = \varepsilon$ for all $t \in [0,1]$, it satisfies Chen's rule*

$$X_{st} = X_{su} \star X_{ut}$$

for all $s, u, t \in [0,1]$ and the analytic bound $|\langle X_{st}, \tau \rangle| \lesssim |t-s|^{\gamma|\tau|}$ for all $\tau \in \mathcal{H}_N$.

As stated in Remark 5.3.3 by passing to the one parameter path $\mathbb{X}_t := X_{0t}$ we can see this definition to be equivalent to Definition 5.3.1.

Remark 5.4.5. To our knowledge the branched version of the Extension Theorem Theorem 5.3.4 is new. However, a version Theorem 5.3.7 was already shown to hold (specifically for the branched case) by Gubinelli in [51] by means of his Sewing Map.

In this regard, the situation is different from that of geometric rough paths. The Butcher–Connes–Kreimer Hopf algebra is free as an algebra over the set of trees so defining characters over it is significantly easier than in the geometric case. To define an element $X \in \mathcal{G}$ it suffices to give the values $\langle X, \tau \rangle$ for all trees $\tau \in \mathcal{T}$; by freeness there is a unique multiplicative extension to all of $\mathcal{H}$. This is not at all the case for geometric rough paths: the algebra H above is **not** free over the linear span of words so if one is willing to define a character over H there are additional algebraic constraints that the values $\langle X, w \rangle$ on words must satisfy. See Remark 5.1.11 and the dicussion thereafter. △

As in the previous section we have

Proposition 5.4.6. *A path $X : [0,1] \to \mathcal{G}_N$ is a branched rough path if and only if it is a rough path in the sense of Definition 5.3.1.*

5.5 Anisotropic rough paths

We now apply our results to another class of rough paths which we call *anisotropic geometric rough paths* (aGRPs for short). As in the geometric case, fix an alphabet $A = \{1, \dots, d\}$ and denote by $\mathrm{M}(A)$ the free monoid generated by A. Let $(\gamma_a : a \in A)$ be a sequence of real numbers such that $0 < \gamma_a < 1$ for all a, and let $\hat{\gamma} = \min_{a \in A} \gamma_a$. For a word $v = a_1 \cdots a_k$ of length k define

$$\omega(v) = \frac{\gamma_{a_1} + \ldots + \gamma_{a_k}}{\hat{\gamma}} = \frac{1}{\hat{\gamma}} \sum_{a \in A}^{d} n_a(v) \gamma_a$$

where $n_a(v) = |\{j : v_j = a\}|$, and observe that ω is additive in the sense that $\omega(uv) = \omega(u) + \omega(v)$ for each pair of words $u, v \in \mathrm{M}(A)$. Given $\alpha > 0$, the set

$$\mathfrak{L}^\alpha := \{v \in \mathrm{M}(A) : \omega(v) \le \alpha \hat{\gamma}^{-1}\}$$

is finite; if $N_\alpha = \lfloor \alpha \hat{\gamma}^{-1} \rfloor$ then $\mathfrak{L}^\alpha \subset H_{N_\alpha}$ and

$$\#\mathfrak{L}^\alpha \le \frac{d^{N_\alpha} - 1}{d - 1}.$$

In analogy with Lemma 5.2.8, the additivity of ω implies

Lemma 5.5.1. *Fix $\alpha > 0$. The subspace $H_\alpha \subset H$ spanned by $\mathfrak{L}^\alpha$ is a subcoalgebra of $(H, \bar{\Delta}, \varepsilon)$.*

Consequently, we will consider the dual algebra $(H_\alpha^*, \star, \varepsilon)$. In this case, we *define* $\mathfrak{g}_{\mathrm{a}}^\alpha$ to be the space of infinitesimal characters on H_α and let $G_{\mathrm{a}}^\alpha = \exp(\mathfrak{g}_{\mathrm{a}}^\alpha)$. As before, there is a canonical injection $\iota_\alpha : H_\alpha^* \to H^*$ so we suppose that $\langle X, v \rangle = 0$ for all $X \in H_\alpha^*$ and $v \notin \mathfrak{L}^\alpha$. We grade H_α by word length and we observe that since $\omega(v) \ge \ell(v)$ we have that

Since H is cofree over the span of A, for each $\lambda > 0$ there is a unique coalgebra automorphism $\Omega_\lambda : H \to H$ such that $\Omega_\lambda a = \lambda^{\gamma_a / \hat{\gamma}} a$ for all $a \in A$. In a similar way as before we have that $(\Omega_\lambda)_{\lambda > 0}$ is a one-parameter family of automorphisms of H. There are also homogeneous norms

$$|X| := \max_{v \in \mathfrak{L}^\alpha} (\ell(v)! |\langle X, v \rangle|)^{1/\omega(v)} \tag{5.16}$$

and

$$\|X\| = \max_{v \in \mathfrak{L}^\alpha} |\langle \log X, v \rangle|^{1/\omega(v)}. \tag{5.17}$$

These homogeneous norms are symmetric, but neither is sub-additive thus they do not generate a metric on G^α_{a}.

Signatures

In order to have a useful metric on G^α_{a} we consider *signatures* of smooth paths. We now assume that $\alpha \geq 1$ so that $A \subset \mathfrak{L}^\alpha$. Let $x = (x^a : a \in A)$ be a collection of (piecewise) smooth paths, and define a map $S(x) : [0,1]^2 \to H^*$ by

$$\langle S(x)_{st}, v\rangle := \int_s^t \mathrm{d}x^{v_k}_{s_k} \int_s^{s_k} \mathrm{d}x^{v_{k-1}}_{s_{k-1}} \cdots \int_s^{s_2} \mathrm{d}x^{v_1}_{s_1}.$$

In his seminal work [23], K. T. Chen showed that $S(x)$ is a multiplicative functional, that is, $S(x)_{st} \in G_{\mathrm{a}}$. In particular $\log S(x)_{st} \in \mathfrak{g}$ thus its restriction to A_α is in $\mathfrak{g}^\alpha_{\mathrm{a}}$ and so we can consider $S(x)$ also as an element of G^α_{a}.

Consider the metric $d_{\mathrm{a}}(X, Y) = \sum_{a \in A} |\langle X - Y, a\rangle|^{\hat{\gamma}/\gamma_a}$ on $H^*_{(1)}$, where we recall that $H_{(1)}$ is the vector space spanned by A. The *anisotropic length* of a smooth curve $\theta : [0,1] \to H^*_1$ is defined to be its length with respect to this metric and will be denoted by $L_{\mathrm{a}}(\theta)$. Observe that since $d_{\mathrm{a}}(\Omega_\lambda X, \Omega_\lambda Y) = \lambda d_{\mathrm{a}}(X, Y)$ we have that $L_{\mathrm{a}}(\Omega_\lambda \theta) = \lambda L_{\mathrm{a}}(\theta)$. We now define another homogeneous norm $|||\cdot||| : G^\alpha_{\mathrm{a}} \to \mathbb{R}_+$, called the *anisotropic Carnot–Carathéodory norm*, by setting

$$|||X||| := \inf\{L_{\mathrm{a}}(x) : x^a \in C^\infty, S(x)_{0,1} = X\}.$$

Since curve length is invariant under reparametrization in any metric space we obtain, as in [46]

Proposition 5.5.2. *The infimum defining the anisotropic Carnot–Carathéodory norm is finite and attained at some minimizing path $\hat{x}$.*

Proposition 5.5.3. *The anisotropic Carnot–Carathéodory norm is homogeneous, that is,* $|||\Omega_\lambda X||| = \lambda |||X|||$.

Proof. Let $\hat{x}$ be the curve such that $|||X||| = L_a(\hat{x})$. For any $\lambda > 0$ and word $v \in \mathfrak{L}^\alpha$ we have

$$\begin{aligned}\langle S(\Omega_\lambda \hat{x})_{0,1}, v\rangle &= \lambda^{\omega(I)} \langle S(\hat{x})_{0,1}, v\rangle \\ &= \langle \Omega_\lambda S(\hat{x})_{0,1}, v\rangle \\ &= \langle \Omega_\lambda X, v\rangle,\end{aligned}$$

thus $|||\Omega_\lambda X||| \leq L_{\mathrm{a}}(\Omega_\lambda \hat{x}) = \lambda L_{\mathrm{a}}(\hat{x}) = |||X|||$. The reverse inequality is obtained by noting that $X = (\Omega_{\lambda^{-1}} \circ \Omega_\lambda) X$. □

The anisotropic Carnot–Carathéodory norm can also be seen to be symmetric and sub-additive, hence it induces a left-invariant metric ν_α on G^α_{a}.

Definition 5.5.4. *An* anisotropic geometric γ-rough path, *with* $\gamma = (\gamma_a,\, a \in A)$, *is a map* $X : [0,1]^2 \to G^1_{\mathrm{a}}$ *such that 1.* $X_{tt} = \varepsilon$, *2. it satisfies Chen's rule* $X_{su} \star X_{ut} = X_{st}$ *for all* $(s,u,t) \in [0,1]^3$, *and 3.* $|\langle X, v\rangle| \lesssim |t-s|^{\hat{\gamma}\omega(v)}$ *for all* $v \in \mathfrak{L}^1$.

Proposition 5.5.5. *Anisotropic geometric γ-rough paths are in one-to-one correspondence with $\hat{\gamma}$-Hölder paths* $X : [0,1] \to (G^1_{\mathrm{a}}, \nu_1)$.

Proof. Let X be an anisotropic geometric γ-rough path and v a word. By definition we have that $|\langle X_{st}, v\rangle| \lesssim |t-s|^{\hat{\gamma}\omega(v)}$, hence $|X_{st}| \lesssim |t-s|^{\hat{\gamma}}$. The equivalence between $|\cdot|$ and $|||\cdot|||$ implies that $\nu_1(X_s, X_t) = |||X_{st}||| \lesssim |t-s|^{\hat{\gamma}}$, hence $t \mapsto X_t$ is γ_1-Hölder with respect to ν_1. The other direction follows in a similar manner. □

Proceeding as in the proof of Theorem 5.3.4 we can show

Theorem 5.5.6. *Let* $(\gamma_a : a \in A)$ *be real numbers such that* $\gamma_a \in (0,1)$ *and* $1 \notin \sum_{a\in A} \gamma_a \mathbb{N}$, *and set* $\hat{\gamma} = \min_{a\in A} \gamma_a$. *Let* $(x^a : a \in A)$ *be a collection of real-valued paths such that* x^a *is* γ_a*-Hölder. There exists an anisotropic rough path* X *such that* $\langle X_{st}, a\rangle = x^a_t - x^a_s$ *for all* $a \in A$.

The following property also holds: given a collection of functions $(g^a : a \in A)$ *with* $g^a \in C^{\gamma_a}$, *let* $\bar{x}^a_t = x^a_t + g^a_t$ *and denote by* gX *the anisotropic geometric γ-rough path above the path*

$$\bar{x}_t = \sum_{a\in A} \bar{x}^a_t a \in H_1.$$

Then, for any two such functions g *and* g' *we have that* $g'(gX) = (g+g')X$.

Proof. The only difference in the proof of this theorem is in the analytical step, because now we have to show that

$$\nu_1(Y_{t^m_k, t^m_{k+1}}) \lesssim 2^{-m\hat{\gamma}}$$

in order to apply Lemma 2 of [88]. We recall that the metric space $(G^1_{\mathrm{a}}, \nu_1)$ only considers the regularity of the components $\langle X, v\rangle$ where the weight $\omega(v) < 1$ by the very definition of the homogeneous norms in eqs. (5.16) and (5.17). Looking at the proof of Theorem 5.3.4 we see that this bound comes from the bound on $\mathrm{BCH}_{(n+1)}(L^n_{su}, L^n_{ut})$ provided by Lemma 5.2.7. In this case we have, for a word $v \in \mathfrak{L}^1$ of length $n+1$ and $s = t^{m+1}_{2k}$, $u = t^{m+1}_{2k+1}$ and $t = t^{m+1}_{2k+2}$,

$$
\begin{aligned}
|\langle \mathrm{BCH}_{(n+1)}(L^n_{su}, L^n_{ut}), e_I\rangle| &\leq \sum_{p+q=n+1} \frac{1}{p!q!} \sum_{\sigma\in S_{n+1}} |a_\sigma| \prod_{r_1=1}^{p} |\langle L^n_{su}, v_{i_{\sigma(r_1)}}\rangle| \prod_{r_2=p+1}^{n+1} |\langle L^n_{ut}, v_{i_{\sigma(r_2)}}\rangle| \\
&= \sum_{p+q=n+1} \frac{1}{p!q!} \sum_{\sigma\in S_{n+1}} |a_\sigma| \prod_{r_1=1}^{p} |x_u^{v_{\sigma(r_1)}} - x_s^{v_{\sigma(r_1)}}| \prod_{r_2=p+1}^{n+1} |x_t^{v_{\sigma(r_2)}} - x_u^{v_{\sigma(r_2)}}| \\
&\leq \sum_{p+q=n+1} \frac{1}{p!q!} \sum_{\sigma\in S_{n+1}} |a_\sigma| \prod_{r_1=1}^{p} |u-s|^{\gamma_{v_{\sigma(r_1)}}} \prod_{r_2=p+1}^{n+1} |t-u|^{\gamma_{v_{\sigma(r_2)}}} \\
&= \sum_{p+q=n+1} \frac{1}{p!q!} \sum_{\sigma\in S_{n+1}} |a_\sigma| 2^{-m(\gamma_{v_{\sigma(1)}}+\cdots+\gamma_{v_{\sigma(n+1)}})} \\
&\lesssim 2^{-m\hat{\gamma}\omega(v)}.
\end{aligned}
$$

This implies that at each stage we have $\|Y_{t^m_k, t^m_{k+1}}\| \lesssim 2^{-m\hat{\gamma}}$ as before, and the equivalence of norms implies the desired bound. The rest of the proof follows through.

Now let g, g' be two collections of functions as in the statement of the theorem. We have the identity

$$\langle [g'(gX)]_t, a\rangle = \langle (gX)_t, a\rangle + (g')^a_t = x^a_t + g^a_t + (g')^a_t = \langle [(g'+g)X]_t, a\rangle.$$

Since both $g'(gX)$ and $(g'+g)X$ are constructed iteratively by adding at each step a function Z satisfying eq. (5.15) on the dyadics, if we let L^n and $\bar{L}^n$ denote the logarithms corresponding to $g'(gX)$ and $(g'+g)X$, Lemma 5.2.7 and the previous identity imply that

$$\mathrm{BCH}_{n+1}(L^n_{su}, L^n_{ut}) = \mathrm{BCH}_{n+1}(\bar{L}^n_{su}, \bar{L}^n_{ut})$$

and so $g'(gX) = (g'+g)X$. □

Corollary 5.5.7. *Let $(x^a : a \in A)$ be a collection of real-valued paths such that $x^a \in C^{\gamma_a}$, and let X be the anisotropic geometric γ-rough path on G^1 given by Theorem 5.5.6. If $b \notin A$ is a new letter, given $0 < \gamma_b < 1$ and $x^b \in C^{\gamma_b}$ is another path, let $A^0 = A \cup \{b\}$ and X^0 be the anisotropic geometric rough path on over $(x^a : a \in A^0)$ given by Theorem 5.5.6. Then, the restriction of X^0 to words from $\mathrm{M}(A)$ coincides with X.*

Proof. If L^n and $\bar{L}^n$ denote the logarithms used in the construction of X and X^0 respectively, and $v \in \mathrm{M}(A)$ is a word not containing b, the BCH formula gives

$$
\begin{aligned}
\langle \mathrm{BCH}_{(n+1)}(\bar{L}^n_{su}, \bar{L}^n_{ut}), v\rangle &= \sum_{p+q=n+1} \frac{1}{p!q!} \sum_{\sigma\in S_{n+1}} a_\sigma \prod_{r_1=1}^{p} (x_u^{v_{\sigma(r_1)}} - x_s^{v_{\sigma(r_1)}}) \prod_{r_2=p+1}^{n+1} (x_t^{v_{\sigma(r_2)}} - x_u^{v_{\sigma(r_2)}}) \\
&= \langle \mathrm{BCH}_{(n+1)}(L^n_{su}, L^n_{ut}), v\rangle.
\end{aligned}
$$

Therefore, at each step in the proof of Theorem 5.5.6 the paths X^n and $(X^0)^n$ are such that for words $v \in \mathrm{M}(A)$ one has

$$\langle (X^0)^n_{st}, v\rangle = \langle X^n_{st}, v\rangle$$

for all $s, t \in [0, 1]$. $\square$

5.6 More on branched rough paths

5.6.1 The Hairer–Kelly map

We recall the following result from [57].

Lemma 5.6.1. *There exists a graded morphism of Hopf algebras $\psi : \mathscr{H} \to T^c(\mathscr{B})$ satisfying $\psi(\tau) = \tau + \psi_{n-1}(\tau)$ for all $\tau \in \mathscr{T}_n$, where ψ_{n-1} denotes the projection of ψ onto $T^c(\mathscr{B}_{n-1})$.*

Observe that since ψ is a Hopf algebra morphism, in particular a coalgebra morphism, then

$$(\psi \otimes \psi)\Delta'\tau = \bar{\Delta}'\psi(\tau) = \bar{\Delta}'\psi_{n-1}(\tau)$$

for all $\tau \in \mathscr{B}_n$ since trees are primitive elements in $T^c(\mathscr{B})$. In fact, from the proof in [57] we are able to see that in fact ψ_{n-1} is given by the recursion $\psi_{n-1} = (\psi \otimes \mathrm{id})\Delta'$.

Example 5.6.2. Here are some examples of the action of ψ on some trees:

$$\begin{aligned}
\psi(\bullet_i) &= \bullet_i \\
\psi(\bullet_a\bullet_b) &= \psi(\bullet_a) \sqcup\!\sqcup\, \psi(\bullet_b) = \bullet_a \otimes \bullet_b + \bullet_b \otimes \bullet_a \\
\psi\big([\bullet_b]_a\big) &= [\bullet_b]_a + \bullet_b \otimes \bullet_a \\
\psi\big([[\bullet_d]_c\,\bullet_b]_a\big) &= [[\bullet_d]_c\,\bullet_b]_a + \bullet_b \otimes [[\bullet_d]_c]_a + \bullet_d \otimes [\bullet_c\bullet_d]_a + [\bullet_d]_c \otimes [\bullet_b]_a + \bullet_d \otimes \bullet_c \otimes [\bullet_b]_a + \bullet_d \otimes \bullet_b \otimes [\bullet_c]_a + \bullet_b \otimes \bullet_d \otimes [\bullet_c]_a \\
&\quad + [\bullet_d]_c \otimes \bullet_b \otimes \bullet_a + \bullet_b \otimes [\bullet_d]_c \otimes \bullet_a + \bullet_d \otimes \bullet_c \otimes \bullet_b \otimes \bullet_a + \bullet_d \otimes \bullet_b \otimes \bullet_c \otimes \bullet_d + \bullet_b \otimes \bullet_d \otimes \bullet_c \otimes \bullet_a.
\end{aligned}$$

$\diamond$

In order to describe the image of ψ we introduce the following extended labels on the nodes. Recall that a labeled forest is a pair $\tau = (F, c)$ where F is a non-planar forest $c : E_F \to \{1, \ldots, d\}$ is a function. An *extended label* is a function $\mathfrak{o} : N_F \to \mathbb{N}$, and we call a triple $(F, c, \mathfrak{o})$ an *extended labelled forest*; we also use the notation $\tau^{\mathfrak{o}}$ in order to stress the extended label. In particular, if $\tau^{\mathfrak{o}}$ is an extended labelled forest we denote the underlying labelled forest by τ.

Definition 5.6.3. *An extended label on a forest is said to be* admissible *if*

1. *$\mathfrak{o}(N_F)$ is an interval containing* 1*, we let $m = m_{\mathfrak{o}} := \max \mathfrak{o}(N_F) \le |\tau|$;*

2. *the function $\mathfrak{o}$ is increasing, that is, $\mathfrak{o}(x) \le \mathfrak{o}(y)$ whenever $x \le y$, and*

3. *for each $1 \le j \le m$ the set $O_j = \{x \in N_F : \mathfrak{o}(x) = j\}$ spans a subtree $\tau_j^{\mathfrak{o}} \subset \tau^{\mathfrak{o}}$.*

Note that condition (2) implies that $\mathfrak{o}$ must be increasing in each factor since nodes from different trees are not comparable, and condition (3) implies that labels must not appear twice in different factors since each set O_j must span a *subtree* and not a subforest. We denote by $\mathfrak{O}(\tau)$ the collection of all the admissible extended labels on the forest τ. It is fairly clear that for each admissible extended label the trees τ_j form a partition of τ into $m_{\mathfrak{o}}$ disjoint subtrees.

Let F be a forest. For $A \subset F$ a subforest we denote by

$$\partial_F A = \{e \in E_F : s(e) \in N_A, t(e) \notin N_A\}$$

the *boundary* of A in F.

Lemma 5.6.4. *Let F be a forest, $\mathfrak{o} \in \mathfrak{O}(F)$ an admissible extended label on F and let T_j be the subtree spanned by O_j as in Definition 5.6.3. Then $\partial_F T_j$ is an admissible cut.*

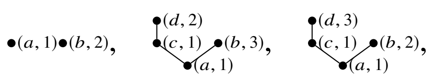
Proof. Denote by $F_>$ the subforest of F spanned by all $x \in N_F \setminus N_{F_j}$ such that $y < x$ for some $y \in T_j$. If $x \notin F_>$ then the unique path from the root to x does not contain edges in $\partial_F T_j$. If $x \in F_>$, suppose that the unique path from the root to x contains two or more edges from $\partial_F T_j$. Pick any two distinct such edges $e_1, e_2 \in E_{T_j}$ and let $y_1 = s(e_1), y_2 = s(e_2) \in N_{T_j}$. Then, the paths going from the root to x through y_1 and through y_2 form a cycle in F, which is a contradiction. □

Proposition 5.6.5. *We have the following representation:*

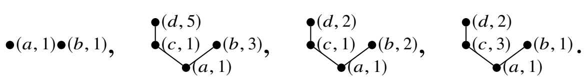
$$\psi(\tau) = \sum_{\mathfrak{o} \in \mathfrak{O}(\tau)} \tau_m \otimes \cdots \otimes \tau_1 \tag{5.18}$$

In particular, each term in this expansion satisfies $|\tau_i| > 0$ and $|\tau_1| + \cdots + |\tau_m| = |\tau|$.

As an example, observe that the following extended labellings are admissible:



whereas the following are not:



Observe that the two terms in the first example give exactly the terms

$$\bullet_b \otimes \bullet_d \otimes \overset{c}{\underset{a}{|}}, \quad \text{and} \quad \bullet_d \otimes \bullet_b \otimes \overset{c}{\underset{a}{|}}$$

in Example 5.6.2.

Proof of Proposition 5.6.5. The proof is by induction on the number of edges of τ, the base case being trivial (see Example 5.6.2).

Suppose identity eq. (5.18) is true for all forests with at most k edges, and let τ be a tree with $k+1$ edges. Let $\mathfrak{A}^*(\tau) = \mathfrak{A}(\tau) \cup \{\varnothing\}$ where $\mathfrak{A}$ is the set of admissible cuts of τ, and set $R^{\varnothing}(\tau) = \tau$, $P^{\varnothing}(\tau) = \mathbf{1}$. By definition

$$\psi(\tau) = \sum_{C \in \mathfrak{A}^*(\tau)} \psi(P^C(\tau)) \otimes R^C(\tau)$$

and by the induction hypothesis

$$\psi(\tau) = \sum_{C \in \mathfrak{A}^*(\tau)} \sum_{\mathfrak{o} \in \mathfrak{O}(P^C(\tau))} \tau_m \otimes \cdots \otimes \tau_1 \otimes R^C(\tau)$$

Given $\sigma \in \mathfrak{A}^*(\tau)$ and $\mathfrak{o} \in \mathfrak{O}(P^C(\tau))$ there is a unique extended label $\tilde{\mathfrak{o}} \in \mathfrak{O}(\tau)$ such that $\tau_1 = R^C(\tau)$. This extended label is defined by $\tilde{\mathfrak{o}}(x) = 1$ if $x \in N_{R^C(\tau)}$ and $\tilde{\mathfrak{o}}(x) = \mathfrak{o}(x) + 1$ if $x \in N_\tau \setminus N_{R^C(\tau)}$. Conversely, given $\tilde{\mathfrak{o}} \in \mathfrak{O}(\tau)$ we have that $\partial\tau_1 \in \mathfrak{A}(\tau)$ and the extended label such that $\mathfrak{o}(x) = \tilde{\mathfrak{o}}(x) - 1$ for $x \in N_\tau \setminus N_{\tau_1}$ belongs to $\mathfrak{O}(P^C(\tau))$. Therefore we have the identity

$$\psi(\tau) = \sum_{\tilde{\mathfrak{o}} \in \mathfrak{O}(\tau)} \tau_1 \otimes \cdots \otimes \tau_{m_{\tilde{\mathfrak{o}}}}.$$

□

The next theorem can be seen as an improvement of Theorem 4.10 in [57].

Theorem 5.6.6. *Let X be a branched γ-rough path, and set $N = \lfloor \gamma^{-1} \rfloor$. There exists an anisotropic geometric rough path $\bar{X}$ indexed by $\mathfrak{T}_N$ with exponents $\gamma_\tau = \gamma|\tau|$, and such that $\langle X, \tau \rangle = \langle \bar{X}, \psi(\tau) \rangle$.*

Proof. We construct $\bar{X}$ iteratively as follows. Let $\bar{X}^{(1)}$ be the (anisotropic) geometric rough path indexed by $\mathfrak{T}_1 = \{\bullet_1, \dots, \bullet_d\}$ over the paths $(x_t^i := \langle X_t, \bullet_i \rangle : i = 1, \dots, d)$ given by Theorem 5.5.6 (alternatively we could use have used Theorem 5.3.4 since all the exponents are equal).

Suppose we have constructed anisotropic geometrics rough paths $\bar{X}^{(k)}$, each indexed by $\mathfrak{T}_k$ over the paths $(x_t^\tau : \tau \in \mathfrak{T}_k)$ such that $x_t^\tau - x_s^\tau = \langle X_{st}, \tau \rangle - \langle \bar{X}_{st}^{(k-1)}, \psi_{k-1}(\tau) \rangle$ for $k = 1, \dots, n$. This is true for $n = 1$ by the previous paragraph. If we let $F_{st}^\tau = \langle X_{st}, \tau \rangle$ and $G_{st}^\tau = \langle \bar{X}_{st}^{(n)}, \psi_n(\tau) \rangle$

for $\tau \in \mathfrak{T}_{n+1}$ we have, by Chen's rule, that

$$\begin{aligned}\delta F^\tau_{sut} &= \langle X_{su} \otimes X_{ut}, \Delta' \tau\rangle \\ &= \langle \bar{X}^{(n)}_{su} \circ \psi \otimes \bar{X}^{(n)}_{ut} \circ \psi, \Delta'\tau\rangle\end{aligned}$$

Since ψ is in particular a coalgebra morphism between $(\mathscr{H}, \Delta)$ and $(H, \bar{\Delta})$ we obtain the identity $\delta F^\tau_{sut} = \langle \bar{X}^{(n)}_{su} \otimes \bar{X}^{(n)}_{ut}, \bar{\Delta}'\psi(\tau)\rangle$, which then, by Lemma 5.6.1 becomes

$$\delta F^\tau_{sut} = \langle \bar{X}^{(n)}_{su} \otimes \bar{X}^{(n)}_{ut}, \bar{\Delta}'\psi_n(\tau)\rangle = \delta G^\tau_{sut}.$$

Therefore there is a path x^τ_t such that $x^\tau_t - x^\tau_s = F^\tau_{st} - G^\tau_{st}$ and then

$$\begin{aligned}|x^\tau_t - x^\tau_s| &\le |\langle X_{st}, \tau\rangle| + |\langle \bar{X}^{(n)}_{st}, \psi_n(\tau)\rangle| \\ &\lesssim |t-s|^{\gamma|\tau|}\end{aligned}$$

since $\psi_n(\tau)$ preserves the number of nodes by Proposition 5.6.5.

We let $\bar{X}^{(n+1)}$ be the anisotropic geometric rough path over $(x^\tau_t : \tau \in \mathfrak{T}_{n+1})$ given by Theorem 5.5.6 and observe that they coincide on $\mathfrak{T}_n$. Finally notice that if τ is a tree then

$$\begin{aligned}\langle \bar{X}^{(n+1)}_{st}, \psi(\tau)\rangle &= \langle \bar{X}^{(|\tau|)}_{st}, \tau\rangle + \langle \bar{X}^{(|\tau|)}_{st}, \psi_{|\tau|-1}(\tau)\rangle \\ &= x^\tau_t - x^\tau_s + \langle X_{st}, \tau\rangle - (x^\tau_t - x^\tau_s) \\ &= \langle X_{st}, \tau\rangle\end{aligned}$$

and the corresponding identity for arbitrary forests follows by multiplicativity. The anisotropic geometric rough path sought for is $\bar{X} = \bar{X}^{(N)}$. □

5.6.2 Modification

In this section we prove Theorem 5.1.2.

Given $\gamma > 0$, let $N = \lfloor \gamma^{-1} \rfloor$ and denote by $\mathscr{C}^\gamma$ the set of functions $g : \mathfrak{T}_N \times [0,1] \to \mathbb{R}$ such that $g(\tau, \cdot) \in C^{\gamma|\tau|}$ and $g(\tau, 0) = 0$ for all $\tau \in \mathfrak{T}_N$. It is easy to see that $\mathscr{C}^\gamma$ is a group under pointwise addition in t, that is,

$$(g + g')(\tau, t) = g(\tau, t) + g'(\tau, t).$$

As a consequence of Theorem 5.5.6, $(g, X) \mapsto gX$ is an action of $\mathscr{C}^\gamma$ on the space of anisotropic geometric rough paths.

We use the map ψ in Lemma 5.6.1 to induce an action of $\mathscr{C}^\gamma$ on branched rough paths. Given a branched rough path X and $g \in \mathscr{C}^\gamma$ we let gX be the branched rough path defined by

$$\langle gX_{st}, \tau \rangle = \langle g\bar{X}_{st}, \psi(\tau) \rangle.$$

where $\bar{X}$ is the anisotropic geometric rough path given by Theorem 5.6.6. As a simple consequence of Theorem 5.5.6 we obtain

Corollary 5.6.7. *We have $g'(gX) = (g' + g)X$ for all $g, g' \in \mathscr{C}^\gamma$.*

Theorem 5.6.8. *The action of $\mathscr{C}^\gamma$ on branched γ-rough paths is transitive: for every pair of branched γ-rough paths X and X' there exists $g \in \mathscr{C}^\gamma$ such that $gX = X'$.*

Proof. We define $g \in \mathscr{C}^\gamma$ inductively by imposing the desired identity. For trees $\tau \in \mathfrak{T}_1 = \{\bullet_1, \ldots, \bullet_d\}$ we set $g(\tau, t) = \langle X'_{0t}, \tau \rangle - \langle X_{0t}, \tau \rangle \in C^\gamma$ so that

$$\begin{aligned}
\langle gX, \tau \rangle &= \langle g\bar{X}, \psi(\tau) \rangle \\
&= \langle g\bar{X}, \tau \rangle \\
&= \langle \bar{X}, \tau \rangle + g(\tau, t) - g(\tau, s) \\
&= \langle X', \tau \rangle
\end{aligned}$$

Suppose we have already defined $g(\tau, \cdot)$ for all $\tau \in \mathfrak{T}_n$ for some $n \geq 1$, satisfying the constraints in the definition of $\mathscr{C}^\gamma$. For a tree τ with $|\tau| = n + 1$ we define

$$F^\tau_{st} = \langle X'_{st}, \tau \rangle - \langle \bar{X}_{st}, \tau \rangle - \langle g\bar{X}_{st}, \psi_n(\tau) \rangle$$

and then

$$\begin{aligned}
\delta F^\tau_{sut} &= \langle X'_{su} \otimes X'_{ut}, \Delta' \tau \rangle - \langle g\bar{X}_{su} \otimes g\bar{X}_{ut}, \bar{\Delta}' \psi_n(\tau) \rangle \\
&= \langle X'_{su} \otimes X'_{ut}, \Delta' \tau \rangle - \langle g\bar{X}_{su} \otimes g\bar{X}_{ut}, \bar{\Delta}' \psi(\tau) \rangle \\
&= \langle X'_{su} \otimes X'_{ut}, \Delta' \tau \rangle - \langle g\bar{X}_{su} \circ \psi \otimes g\bar{X}_{ut} \circ \psi, \bar{\Delta}' \tau \rangle \\
&= \langle X'_{su} \otimes X'_{ut}, \Delta' \tau \rangle - \langle gX_{su} \otimes gX_{ut}, \bar{\Delta}' \tau \rangle \\
&= 0
\end{aligned}$$

by the induction hypothesis. Hence there is $g(\tau, \cdot) : [0, 1] \to \mathbb{R}$ such that $g(\tau, 0) = 0$ and

$$g(\tau, t) - g(\tau, s) = \langle X'_{st}, \tau \rangle - \langle \bar{X}_{st}, \tau \rangle - \langle g\bar{X}_{st}, \psi_n(\tau) \rangle$$

whence $g \in C^{\gamma|\tau|}$; by construction

$$\begin{aligned}
\langle gX, \tau \rangle &= \langle g\bar{X}, \psi(\tau) \rangle \\
&= \langle g\bar{X}, \tau \rangle + \langle g\bar{X}, \psi_n(\tau) \rangle \\
&= \langle \bar{X}, \tau \rangle + g(\tau, t) - g(\tau, s) + \langle g\bar{X}, \psi_n(\tau) \rangle \\
&= \langle X', \tau \rangle .
\end{aligned}$$

□

5.6.3 The BCFP renormalisation

In [13] a different kind of modification is proposed. There, a new label 0 is considered so rough paths –branched and geometric– are over paths taking values in $\mathbb{R}^{d+1}$. Recall that since branched rough paths are seen as Hölder paths taking values in the character group of the Connes–Kreimer Hopf algebra, we may think of them as an infinite forest series of the form

$$X_{st} = \sum_{\tau \in \mathscr{F}} \langle X_{st}, \tau \rangle \tau \tag{5.19}$$

where we regard τ as a linear functional on $\mathscr{H}(\mathbb{R}^{d+1})$, such that $\langle \tau, \sigma \rangle = 1$ if $\sigma = \tau$ and zero else. The aforementioned modification procedure then acts as a translation of the series (5.19). Specifically, for each collection $v = (v_0, \dots, v_d) \in (\mathscr{B}^*)^{d+1}$ an operator $M_v : \mathscr{H}(\mathbb{R}^{d+1})^* \to \mathscr{H}(\mathbb{R}^{d+1})^*$ is defined, such that for a γ-branched rough path, $(M_v X)_{st} := M_v(X_{st})$ is a γ/N-branched rough path.

In the particular case where $v_j = 0$ except for v_0, the action of this operator can be described in terms of an extraction/contraction map[1] $\Psi : \mathscr{H}(\mathbb{R}^{d+1}) \to \mathscr{H}(\mathbb{R}^{d+1}) \otimes \mathscr{H}(\mathbb{R}^{d+1})$. This map acts on a tree τ by extracting subforests and placing them in the left factor; the right factor is obtained by contracting the extracted forest labelling the resulting node with 0. As an example, consider

$$\begin{aligned}
\Psi([\bullet_j \bullet_k]_i) = \mathbf{1} \otimes [\bullet_j \bullet_k]_i + \bullet_i \otimes [\bullet_j \bullet_k]_0 + \bullet_j \otimes [\bullet_0 \bullet_k]_i + \bullet_k \otimes [\bullet_j \bullet_0]_i + [\bullet_j]_i \otimes [\bullet_k]_0 + [\bullet_k]_i \otimes [\bullet_j]_0 \\
+ \bullet_i \bullet_j \otimes [\bullet_0 \bullet_k]_0 + \bullet_i \bullet_k \otimes [\bullet_j \bullet_0]_0 + \bullet_j \bullet_k \otimes [\bullet_0 \bullet_0]_i + [\bullet_j]_i \bullet_k \otimes [\bullet_0]_0 + [\bullet_k]_i \bullet_j \otimes [\bullet_0]_0 + [\bullet_j \bullet_k]_i \otimes \bullet_0 .
\end{aligned}$$

Extending $v = v_0 \in \mathscr{B}^*$ to all of $\mathscr{H}^*$ as an algebra morphism it is shown that

$$\langle (M_v X)_{st}, \tau \rangle = \langle X_{st}, (v \otimes \mathrm{id}) \Psi(\tau) \rangle . \tag{5.20}$$

Furthermore, in this case $M_v X$ is a γ-branched rough path if coefficients corresponding to trees

[1] In [13] this map is named δ but we choose to call it Ψ in order to avoid confusion with the operator defined here.

with label zero are required to satisfy the stronger analytical condition

$$\sup_{0\le s,t\le 1} \frac{|\langle X_{st}, \tau\rangle|}{|t-s|^{(1-\gamma)|\tau|_0+\gamma|\tau|}} < \infty, \tag{5.21}$$

where $|\tau|_0$ counts the times the label 0 appears in τ. Essentially, this condition imposes that the components corresponding to the zero label be Lipschitz on the diagonal $s = t$.

We now show how this setting can be recovered from the results of Section 5.6.2. Let X be a γ-branched rough path on $\mathbb{R}^{d+1}$ satisfying (5.21). Since $M_v X$ is again a γ-branched rough path, by Theorem 5.6.8 there exists a collection of functions $g \in \mathscr{C}^\gamma$ such that $gX = M_v X$. Moreover, this collection is the unique one satisfying

$$g(\tau,t) - g(\tau,s) = \langle X_{st}, (v\otimes \mathrm{id})\Psi(\tau)\rangle - \langle X_{st}, \tau\rangle - \langle g\bar{X}_{st}, \psi_{|\tau|-1}(\tau)\rangle$$

for all $\tau \in \mathfrak{T}(\mathbb{R}^{d+1})$ where we have used eq. (5.20) in order to express $M_v X$ in terms of Ψ. Theorem 28 in [13] ensures that the first term on the right-hand side is in $C_2^{\gamma|\tau|}$ hence g is actually in $C^{\gamma|\tau|}$ as required.

The approach of [13] is based on pre-Lie morphisms and crucially on a *cointeraction property*, which has been explored by [17], see in particular [13, Lemma 18]. The cointeraction property can be used for *time-independent* modifications, indeed note that the functional v in [13] always constant.

Let us see why this is the case. The approach of [13] is based on a cointeraction property studied by [15, 17, 42] between the Connes-Kreimer coproduct and another *extraction-contraction* coproduct $\delta : \mathscr{H} \to \mathscr{H} \otimes \mathscr{H}$. The formula is the following

$$(\mathrm{id}\otimes\Delta)\delta = \mathscr{M}_{1,3}(\delta\otimes\delta)\Delta.$$

Let us consider now a character $v \in \mathscr{H}^*$. If we multiply both sides by $(v\otimes\mathrm{id}\otimes\mathrm{id})$ and set $M_v^* = (v\otimes\mathrm{id})\delta : \mathscr{H} \to \mathscr{H}$ as in[13, Proposition 17], then we obtain

$$\Delta\, M_v^* = (M_v^*\otimes M_v^*)\Delta,$$

namely M_v^* is a coalgebra morphism on $\mathscr{H}$. Then one can define a modified rough path as $vX := M_v X = X \circ M_v^*$. The crucial Chen property is still satisfied since

$$\begin{aligned}(vX)_{st} &= (v\otimes X_{st})\delta = (v\otimes X_{su}\otimes X_{ut})(\mathrm{id}\otimes\Delta)\delta \\ &= (v\otimes X_{su}\otimes X_{ut})\mathscr{M}_{1,3}(\delta\otimes\delta)\Delta = (v\otimes X_{su})\otimes(v\otimes X_{ut})(\delta\otimes\delta)\Delta = ((vX)_{su}\otimes(vX)_{ut})\Delta\end{aligned}$$

However this does not work if $v : [0,1]^2 \to \mathscr{H}^*$ is a time-dependent character. Indeed in this case we set $(vX)_{st} := (v_{st} \otimes \mathrm{id})\delta$ and we obtain

$$\begin{aligned}(vX)_{st} &= (v_{st} \otimes X_{st})\delta = (v_{st} \otimes X_{su} \otimes X_{ut})(\mathrm{id} \otimes \Delta)\delta \\ &= (v_{st} \otimes X_{su} \otimes X_{ut})\mathcal{M}_{1,3}(\delta \otimes \delta)\Delta = (v_{st} \otimes X_{su}) \otimes (v_{st} \otimes X_{ut})(\delta \otimes \delta)\Delta\end{aligned}$$

but we can not conclude that this is equal to $((vX)_{su} \otimes (vX)_{ut})\Delta$. Our construction, as explained after formula (5.3), is not purely algebraic but is based on a (non-canonical) choice of genealised Young integrals with respect to the rough path X.

Moreover our transformation group, infinite-dimensional, is much larger than that finite-dimensional group studied in [13].

Chapter 6

Directed Polymers

6.1 Introduction

In this section we review the theory of directed polymers and their relation with the domain of Stochastic Partial Differential Equations (SPDEs). We summarise some of the recent developments in the study of the Kardar–Parisi–Zhang (KPZ) equation and universality class. For the general theory of SPDEs the reader is referred to the book by Da Prato and Zabczyk [29].

The Kardar–Parisi–Zhang equation, first defined in [73], describes the evolution of a randomly growing interface. It is the simplest non-linear local model exhibiting diffusive behaviour describing the growth of an interface applicable to processes such as vapor deposition. The equation describing this model in one spatial dimension, known as the KPZ equation, is given by

$$\partial_t h = \nu \partial_x^2 h + \frac{\lambda}{2}(\partial_x h)^2 + \xi \tag{6.1}$$

where ξ is a space-time white noise, that is, a random Gaussian distribution with formal covariance structure

$$\mathbb{E}[\xi(t,x)\xi(s,y)] = \delta(y-x)\delta(t-s). \tag{6.2}$$

The parameters $\nu, \lambda \geq 0$ appearing in eq. (6.1) describe relaxation of the interface by surface tension and latteral growth due to material deposition, respectively. A formal computation (valid for example when ξ is replaced by a smooth potential) shows that the *Hopf–Cole transformation* $\mathscr{Z}(t,x) = \exp\big(\frac{\lambda}{2\nu} h(t,x)\big)$ maps solutions to eq. (6.1) to solutions to the Stochastic Heat Equation

$$\partial_t \mathscr{Z} = \nu \partial_x^2 \mathscr{Z} + \frac{\lambda}{2\nu} \mathscr{Z} \xi. \tag{6.3}$$

Observe that contrary to the case of the KPZ equation, eq. (6.3) is linear, hence one can attempt to find solutions even when ξ is as irregular as space-time white noise. For simplicity, we chose $\nu = \frac{1}{2}$ and $\lambda = 1$. It can be argued that the Hopf–Cole solution $h(t,x) = \log(\mathfrak{X}(t,x))$ is the (only) physically meaningful solution in that it has the correct properties expected from the situations it models.

The importance of eq. (6.1) is that it is in some sense universal. The precise meaning of this statement is somewhat involved, but in a very loose sense it means that there is a wide class of physical models –both discrete and continuous– where some quantities of interest converge under an appropriate scaling limit to a solution to eq. (6.1), independently of the precise structure of the model. That is to say that for each model in a certain class, called the *KPZ universality class*, depending on some parameters, one can rescale those parameters in such a way so that what one obtains in the limit can be identified as a solution to the KPZ equation. Another interesting feature of this equation in this situation is that the models exhibit non-Gaussian fluctuations, meaning that convergence to a non-trivial limit happens on a different scale than that of what we are used to in the Gaussian setting, and the limiting distributions are not Gaussian in general. Take the Central Limit Theorem as an example: in this case, if we have a collection $(X_n : n \in \mathbb{N})$ of i.i.d. random variables with mean μ and finite variance σ^2, then the partial sums $S_n = X_1 + \cdots + X_n$, properly rescaled as

$$\frac{S_n - n\mu}{\sqrt{n}} \to Z$$

where Z is a normally distributed random variable with mean μ and variance σ^2. That is to say, the non-trivial fluctuations of S_n are Gaussian at the scale of $n^{1/2}$, i.e.

$$S_n \sim n\mu + \sqrt{n}Z.$$

Now the universality statement can be made a bit more precise: it is expected that for spatial dimension $d = 1$, models belonging to the KPZ universality class have stochastic fluctuations of the form

$$h(t, t^{2/3}x) \sim \mu t + t^{1/3}Z \tag{6.4}$$

where the exact distribution of Z depends on the class of the initial condition but not on the details of the model. There are 4 such types or classes of initial conditions, which are detailed and examined in [101].

Evidence of why this should be true can be obtained by examining how eq. (6.1) rescales. Let h be a solution to eq. (6.1) and let $h_\epsilon(t,x) = \epsilon^b h(\epsilon^{-z}t, \epsilon^{-1}x)$. Space-time white noise satisfies the formal scaling relation $\xi(\epsilon^{-z}t, \epsilon^{-1}x) \stackrel{d}{=} \epsilon^{(z+1)/2}\xi(t,x)$ –the details of this equality are explained in section 6.2.1 below. Then, applying the chain rule gives that h_ϵ satisfies

$$\partial_t h_\epsilon = \frac{1}{2}\epsilon^{2-z}\partial_x^2 h_\epsilon + \frac{1}{2}\epsilon^{2-z-b}(\partial_x h)^2 + \epsilon^{b-(z-1)/2}\xi.$$

Since one expects the initial conditions to be locally Brownian, this forces $b = \frac{1}{2}$ and then one should take $z = \frac{3}{2}$ in order to avoid divergence of the non-linear term. Thus, one should see non-trivial behaviour at the scale of $h_\epsilon(t,x) = \epsilon^{1/2}h(\epsilon^{-3/2}t, \epsilon^{-1}x)$ or, taking $\epsilon = t^{-2/3}$, at the scales appearing in eq. (6.4).

The first model shown to be in the KPZ universality class was the *weakly asymmetric simple exclusion process* (WASEP). In 1999, L. Bertini and G. Giacomin showed that an appropriately rescaled height function converges to the Hopf–Cole solution to eq. (6.1). They also show that if one smooths out the space-time white noise ξ appearing on the right-hand side of eq. (6.1) in space –this is essentially to keep Itô's formula– by performing stochastic convolution against a suitable smooth mollifier, then the resulting solutions converge weakly to a well-defined stochastic process which is then identified as a Hopf–Cole solution to the KPZ equation. More precisely, let J_κ be a symmetric smooth function with compact support and let $\xi_t^\kappa(x) = \langle \xi_t, J_\kappa(x-\cdot)\rangle \in C^\infty$. The, if $\mathfrak{X}^\kappa$ is the solution to eq. (6.3) with ξ replaced with ξ^κ –which is a classical solution– and if $h^\kappa = \log(\mathfrak{X}^\kappa)$ then by Itô's formula

$$\partial_t h^\kappa = \frac{1}{2}\partial_x^2 h^\kappa + \frac{1}{2}[(\partial_x h^\kappa)^2 - C\kappa^{-1}] + \xi^\kappa + o(1).$$

where the constant $C > 0$ depends only on the mollifier. Observe that the non-linear term has changed. This is a simple instance of *Wick renormalisation*. The new non-linear term is in fact the Wick square : $(\partial_x h^\kappa)^2 := (\partial_x h^\kappa)^2 - \mathbb{E}(\partial_x h^\kappa)^2$. Other models belonging to this class are the *ballistic aggregation model* and the *Eden model*.

Finally, we describe the connection with *directed random polymers*. A discrete polymer path is a nearest-neighbour up-right path on $\mathbb{Z}_+^2$, started at the origin. Given $n, m \in \mathbb{N}$, denote by $\Pi_{m,n}$ the collection of all polymer paths ending at (m,n); clearly this set is finite. Suppose we are given a collection $(\omega_{i,j})_{i,j\geq 0}$ of i.i.d random variables. The weight of a path $\pi \in \Pi_{m,n}$ is defined as

$$H_{m,n}(\pi) := \sum_{k=1}^{m+n} \omega_{\pi_k}.$$

Let $\mathbf{P}_{m,n}$ denote the uniform measure on $\Pi_{m,n}$. Given a parameter $\beta > 0$, known as the *inverse temperature* of the model, define the discrete polymer measure

$$\mathbf{P}_{m,n}^\beta(\pi) := \frac{1}{Z_{n,m}^\beta} \mathrm{e}^{\beta H_{m,n}(\pi)}\,\mathbf{P}(\pi)$$

where the normalizing constant

$$Z_{m,n}^\beta := \sum_{\pi\in\Pi_{m,n}} \mathrm{e}^{\beta H_{m,n}(\pi)}$$

is known as the *partition function*. In the case $\beta = 0$ the polymer measure $\mathbf{P}_{m,n}^0$ corresponds to the simple random walk and so it should exhibit diffusive behaviour. On the contrary, in the limit $\beta \to \infty$ the measure concentrates on the single path maximizing the energy $H_{m,n}(\pi)$ and this maximum value plays the role of the free energy of the polymer; this model is known as *last passage percolation*. In some particular cases when the distribution of the random weights is known, the model becomes exactly solvable, meaning that some quantities key for the analysis become can be explicitly computed: this is the case when $\omega_{i,j}$ is a geometric or an exponential random variable. In 2000, K. Johansson showed that under this assumption the maximal weight chosen by $\mathbf{P}^{\infty}_{n,m}$ has the conjectured $\frac{1}{3}$ fluctuation exponent and that its properly rescaled fluctuations converge to the GUE Tracy–Widom distribution [68].

Some years later, T. Seppäläinen introduced the log-Gamma polymer which fixes a specific distribution for the weights –the Gamma distribution, so that $e^{\omega_{i,j}}$ has the log-Gamma distribution [109]. The model exactly solvable in this case and in that same article Seppäläinen shows that the partition function has right fluctuation exponent. Later, A. Borodin, I. Corwin and D. Remenik also proved that the random variable giving the fluctuations has the GUE Tracy-Widom distribution under the same scaling [8]. Finally, T. Alberts, K. Khanin and J. Quastel prove that there is another interesting regime, the *intermediate disorder regime*, where the size of the disorder is scaled alongside the polymer and where one can obtain results that are independent of the distribution of the weights [3]. In particular, they show that if the inverse temperature is rescaled as $\beta_n := n^{-1/4}\beta$ the free energy fluctuates as

$$\log Z^{\beta_n}_{n,n} \sim n\lambda(\beta_n) + \mathfrak{X}^{\beta}$$

where $\mathfrak{X}^{\beta}$ is a solution to section 6.4. They proved, among other things, the following

Theorem 6.1.1 (Alberts–Khanin–Quastel)**.** *Assume the weights $(\omega_{i,j})_{i,j\geq 0}$ are i.i.d. and such that $e^{\lambda(\beta)} := \mathbb{E}e^{\beta\omega} < \infty$ for all $\beta > 0$. Then, the rescaled partition function $e^{-n\lambda(n^{-1/4}\beta)}Z^{n^{-1/4}\beta}_{n,n}$ converges in distribution to a random variable $\mathfrak{X}$.*

In the same paper a heuristic explanation of why the $n^{-1/4}$ scaling should be the correct one. Using the rough approximation $e^x \approx 1 + x$ one obtains that, after performing a suitable change of coordinates, that

$$Z^{\beta_n}_{n,n} \approx 1 + \beta n^{-1/4} \sum_{i=1}^{n} \sum_{x\in\mathbb{Z}} \tilde{\omega}_{i,x}\mathbf{P}(S_i = x)$$

where the $\tilde{\omega}$ are still i.i.d, centred and with variance one, and S is a simple random walk under $\mathbf{P}$. Then, it follows that the right-hand side converges in distribution to a zero mean Gaussian random variable with variance $\frac{2\beta^2}{\sqrt{\pi}}$.

In a subsequent work, the same authors identified the limiting random variable with the partition function of a model they called the *continuum random polymer* [2]. This polymer model is such that its partition function is given by the solution to the Stochastic Heat Equation evaluated at a point. Essentially, given a realisation of the driving space-time white noise ξ, they use the solution to eq. (6.3) to construct a family of measures $\mathbf{P}^{\xi}_{\beta}$ over continuous paths defined on $[0,1]$, which turn out to be singular with respect to the standard Wiener measure on $C([0,1])$ for each $\beta > 0$ and almost all ξ. Nonetheless, the measure $Q_\beta(\mathrm{d}x\mathrm{d}\xi) := \mathbf{P}^{\xi}_{\beta}(\mathrm{d}x)\mathbb{P}(\mathrm{d}\xi)$ coincides with the Wiener measure for $\beta = 0$. Moreover, the measures $\mathbf{P}^{\xi}_{\beta}$ are a.s. supported on α-Hölder paths for all $\alpha < \frac{1}{2}$ and paths chosen under Q_β have $[X]_t = t$ as their quadratic variation. This indicates that polymer path measure Q_β should be just Brownian motion with a drift; this is more or less true, save for the fact that the drift is not in the Cameron–Martin class (hence the singularity of Q_β w.r.t. Q_0). Again, this is indication that there is need for *renormalisation*, as indicated above. In his breakthrough paper [55], M. Hairer made sense of this interpretation by using the theory of Rough Paths (see Chapter 5), and explicitly computes the required renormalisation constants in order to make sense of the (solution to the) KPZ equation eq. (6.1) as the limit of (the solutions to) classical PDEs. The same work would lead him to build his theory of Regularity Structures [57] a few years later, and for which he was awarded a Fields medal in 2014.

Using the same idea of the proof of Theorem 6.1.1, J. Quastel, G. Moreno-Flores and D. Remenik showed that a similar result holds true for a related model [93], called the *semi-discrete directed polymer in a Brownian environment*, introduced by N. O'Connell and M. Yor in 2001 [99]. In this model, the grid $\mathbb{Z}^2_+$ is replaced with the ensemble of lines $\mathbb{R}_+ \times \mathbb{N}$ (hence time is continuous but space is discrete) and polymer paths are now replaced by the paths of a Poisson process with rate 1, starting at 1 at time 0 and ending at N at a fixed time $T > 0$. The collection of such paths is denoted by $\Delta_N(T)$, and can be identified with the set of increasing $(N-1)$-tuples $0 < t_1 < \cdots < t_{N-1} < T$. The environment is replaced by a countable collection $(B^{(k)} : k \in \mathbb{N})$ of independent standard Brownian motions and the weight of a semi-discrete polymer path is now defined as

$$H_{N,T}(\boldsymbol{t}) := \sum_{k=1}^{N} B^{(k)}(t_{k-1}, t_k)$$

with the conventions $t_0 = 0$, $t_N = T$. Then, for a given $\beta > 0$ one defines the "Gibbs measure"

$$\mathbf{P}^{\beta}_{N,T}(\mathrm{d}\boldsymbol{t}) := \frac{1}{Z^{\beta}_{N,T}} \mathrm{e}^{\beta H_{N,T}(\boldsymbol{t})}\, \mathrm{d}\boldsymbol{t}$$

where

$$Z^{\beta}_{N,T} := \int_{\Delta_N(T)} \mathrm{e}^{\beta H_{N,T}(\boldsymbol{t})}\, \mathrm{d}\boldsymbol{t} \tag{6.5}$$

is the partition function of the model. Then one has

Theorem 6.1.2 (Quastel–Moreno-Flores–Remenik)**.** *As $N \to \infty$ the partition function*

$$\sqrt{N}\mathrm{e}^{-Nt+\sqrt{N}x-\beta^2/2(\sqrt{N}t-x)} Z^{N^{-1/4}\beta}_{Nt-\sqrt{N}x,Nt} \to \mathfrak{X}^{\beta}(t,x)$$

in L^2, uniformly on compact sets of $(0,\infty) \times \mathbb{R}$.

In order to further study the properties of the solution to eq. (6.1), I. Corwin and A. Hammond introduced, in 2013, a “multi-layer” extension of the KPZ equation which they called the *KPZ line ensemble* [25], mainly motivated by a similar construction made one year earlier by O’Connell and Warren [98], the *multilayer Stochastic Heat Equation*, discussed in section 6.4.1. This construction relies on the result by Moreno-Flores–Quastel–Remenik (Theorem 6.1.2), as well as formulas and results by O’Connell [97], and G. Amir, I. Corwin and J. Quastel [4]. The KPZ line ensemble consists on a collection of continuous curves $(h^{(n)} : n \in \mathbb{N})$ such that $h^{(1)}$ has the same distribution as the solution to the KPZ equation, the entire ensemble has a resampling invariance over finite intervals which they call the “Brownian Gibbs property”, and $h^{(1)}$ when properly rescaled are uniformly absolutely continuous with respect to Brownian bridges as $t \to \infty$. The multilayer extension to the SHE introduced by O’Connell and Warren is introduced by means of its chaos expansion

$$\mathscr{Z}_\beta^{(n)}(t,x) := p(t,x)^n \left(1 + \sum_{k=1}^{\infty} \beta^k \int_{\Delta_k(t)} \int_{\mathbb{R}^k} R_k^{(n)}(\boldsymbol{s}, \boldsymbol{y}; t, x) \right) \xi^{\otimes k}(\mathrm{d}\boldsymbol{s}, \mathrm{d}\boldsymbol{y}) \tag{6.6}$$

Currently, it is not known whether Corwin and Hammond’s KPZ line ensemble or O’Connell and Warren’s multilayer SHE are solutions to some SPDEs although there are some guesses in the latter case. In the same paper it is conjectured, among other things, that there is a relation between the KPZ line ensemble and the multilayer SHE in the spirit of the Hopf–Cole transformation. To state this conjecture, Corwin and Hammond introduce a semi-discrete analogue of the O’Connell–Warren multilayer SHE which they call the *O’Connell–Yor polymer partition function line ensemble* and state a scaling under which this process should converge to the multilayer SHE, thus giving a “multilayer” analogue of Theorem 6.1.2. In particular, this would imply that the processes $(h^{(n)}(t,x) : n \in \mathbb{N}, x \in \mathbb{R})$ and $\left(\log\left(\frac{\mathscr{Z}^{(n)}(t,x)}{\mathscr{Z}^{(n-1)}(t,x)}\right) : n \in \mathbb{N}, x \in \mathbb{R}\right)$ are equal for each fixed $t \geq 0$.

This chapter presents a proof of this conjecture which is based on the ideas of Moreno-Flores, Quastel and Remenik. Concretely we prove that

Theorem 6.1.3. *Let $Z_\beta^{(n)}(T,N)$ be the point-to-point n-layer semi-discrete polymer partition function and let $\mathscr{Z}_\beta^{(n)}(t,x)$ be the random field defined in eq.* (6.6)*. Then,*

$$\mathrm{e}^{-nN-\frac{1}{2}\beta^2 nN^{1/2}} Z^{(n)}_{\beta N^{-1/4}}(N,N) \to \mathscr{Z}_\beta^{(n)}(1,0)$$

in distribution.

Remark 6.1.4. While this results were being developed in the context of this thesis, M. Nica came up with a proof of the same result using similar techniques [95]. In fact, his approach relies in a coupling between the driving noise in Equation (6.3) and the Brownian motions appearing in Equation (6.5), as well as a shear transformation introduced in [93]. △

Outline. In the rest of this chapter we first review in brief detail some of the technical prerequisites needed for stating the main results presented. In Section 6.2 we present the basic theory of space-time white noise and its associated stochastic integration. In Section 6.3 we introduce the theory of determinantal processes and state some facts about how they may be analysed. Next, in Sections 6.3.2 and 6.3.3 we review the construction of "bridge processes" of a given stochastic process via Doob's h-transform. In Section 6.4 we use these facts to describe the solutions to the Stochastic Heat Equation as a chaos expansion with respect to space-time white noise and in Section 6.4.1 we present the multilayer extension of the SHE, introduced by N. O'Connell and J. Warren [98]. Finally, in the remaining sections we state and prove the main results of this chapter.

6.2 White noise

As stated before, space-time white noise is a Gaussian random distribution with covariance eq. (6.2). In this section, we describe one way to make this formal. Since it will be enough for our purposes, we restrict ourselves to the case of one spatial dimension.

Let $\mathcal{S}$ denote the Schwartz space of rapidly decreasing real-valued functions as defined and studied in Section 3.2.2. Without entering into too much detail, white noise can be realised as a continuous $\mathcal{S}'$-valued Gaussian process $(\xi_t : t \geq 0)$ such that

$$\mathbb{E}[\langle \xi_t, \phi\rangle\langle \xi_s, \psi\rangle] = t \wedge s \int_{-\infty}^{\infty} \phi(x)\psi(x)\,\mathrm{d}x \tag{6.7}$$

for all $\phi, \psi \in \mathcal{S}$. Here we use the notation $t \wedge s := \min\{t, s\}$. In particular, for each $\phi \in \mathcal{S}$ the process $t \mapsto \|\phi\|_{L^2}^{-1}\langle \xi_t, \phi\rangle$ is a standard Brownian Motion. Thus, the stochastic process ξ can be regarded as an infinite-dimensional analogue of Brownian Motion. More generally, if $\phi : \mathbb{R}_+ \to \mathcal{S}$ is a predictable process –with respect to an appropriate filtration, see [101, Section 2] for further details– such that

$$\mathbb{E}\left(\int_0^{\infty} \|\phi_s\|_{L^2}^2\,\mathrm{d}s\right) < \infty$$

then standard martingale techniques allow us to define a stochastic integral

$$\int_0^t \langle \psi_s, \mathrm{d}\xi_s\rangle \tag{6.8}$$

such that

$$\mathbb{E}\left(\int_0^t \langle \psi_s, \mathrm{d}\xi_s\rangle\right)^2 = \mathbb{E}\left(\int_0^t \|\psi_s\|_{L^2}^2\,\mathrm{d}s\right). \tag{6.9}$$

Note that the right-hand side of eq. (6.9) can be more explicitly written as

$$\int_0^t \int_{-\infty}^{\infty} \psi_s(x)^2 \, \mathrm{d}x\mathrm{d}s,$$

hence by abuse of notation we will write the integral in eq. (6.8) as

$$\int_0^t \int_{\mathbb{R}} \psi(s,x)\, \xi(\mathrm{d}s, \mathrm{d}x)$$

so eq. (6.9) reads

$$\mathbb{E}\Big(\int_0^t \int_{\mathbb{R}} \psi(s,x)\, \xi(\mathrm{d}s, \mathrm{d}x)\Big)^2 = \mathbb{E}\int_0^t \int_{\mathbb{R}} \psi(s,x)^2 \, \mathrm{d}x \, \mathrm{d}s$$

which is reminiscent of the classical Itô isometry property of the usual stochastic integral. Moreover, this isometry property allows us to extend the stochastic integral to functions $\psi \in L^2([0,\infty) \times \mathbb{R}, \mathrm{d}s\, \mathrm{d}x)$ by a density argument since the Schwartz space is dense in L^2 with respect to the L^2 norm. Finally, we stress that the process $t \mapsto \int_0^t \int_{\mathbb{R}} \psi(s,x)\, \xi(\mathrm{d}s, \mathrm{d}x)$ is an L^2 martingale with quadratic variation $\int_0^t \int_{\mathbb{R}} \psi(s,x)^2 \, \mathrm{d}s\mathrm{d}x$ again by eq. (6.9). In particular, for any $\psi \in L^2([0,\infty) \times \mathbb{R}, \mathrm{d}s\mathrm{d}x)$, the stochastic integral $\langle \xi, \psi \rangle$ is bounded in L^2, hence it has an almost sure limit.

One way to explicitly construct the process ξ is to take a collection $(\beta_n : n \in \mathbb{N})$ of independent standard real-valued Brownian Motions and $(e_n : n \in \mathbb{N})$ an orthonormal basis of $L^2(\mathbb{R}, \mathrm{d}x)$. Then, the series

$$\xi_t = \sum_{n \in \mathbb{N}} \beta_n(t) \hat{e}_n$$

defines a Wiener process with covariance eq. (6.7). Here $\hat{e}_n$ denotes the Fourier transform of e_n.

6.2.1 Scaling and stochastic convolutions

Given $\lambda > 0$, $b, x \in \mathbb{R}$ and a Schwartz function $\phi \in \mathcal{S}$ we denote by $\phi_x^{\lambda,b} \in \mathcal{S}$ the function defined by $\phi_x^{\lambda,b}(y) = \lambda^{-b}\phi(\lambda^{-1}(y-x))$. Observe that by definition $\|\phi_x^{\lambda,b}\|_{L^2} = \lambda^{1-2b}\|\phi\|_{L^2}$ for all $b, \lambda > 0$ and $x \in \mathbb{R}$.

Now, for given $\lambda > 0$ and $z, b, x \in \mathbb{R}$ and a space-time white noise ξ define a new $\mathcal{S}'$-valued stochastic process $(\xi_t^{\lambda;b,z} : t \geq 0)$ by $\langle \xi_t^{\lambda;b,z}, \phi \rangle := \langle \xi_{\lambda^z t}, \phi_0^{\lambda,b} \rangle$. Then,

$$\mathbb{E}[\langle \xi_t^{\lambda;b,z}, \phi \rangle \langle \xi_s^{\lambda;b,z}, \psi \rangle] = \lambda^{-2b+z+1} \cdot t \wedge s \int_{\mathbb{R}} \phi(x)\psi(x) \, \mathrm{d}x.$$

Hence, since Gaussian processes are characterised by their covariance kernels, we have that $\xi^{\lambda;b,z}$ has the same distribution as ξ whenever $b = (z+1)/2$. This can be written informally by saying that $\lambda^{(z+1)/2}\xi(\lambda^z t, \lambda x) \overset{d}{=} \xi(t,x)$ for all $z, \lambda > 0$.

Let ρ be a smooth compactly supported function defined on $\mathbb{R}_+ \times \mathbb{R}$, and define the function $\rho^{t,x}(s,y) := \rho(t-s, x-y)$. The *stochastic convolution* between ρ and ξ is the smooth function

$$\rho * \xi(t,x) := \int_0^t \langle \rho^{t.x}(s,\cdot), \mathrm{d}\xi_s \rangle \tag{6.10}$$

By definition the collection $(\rho * \xi(t,x) : (t,x) \in \mathbb{R}_+ \times \mathbb{R})$ is a Gaussian process with covariance

$$\mathbb{E}[\rho * \xi(t,x) \rho * \xi(s,y)] = \int_0^{t\wedge s} \int_{\mathbb{R}} \rho(t-s', x-y')\rho(s-s', y-y')\, \mathrm{d}s\mathrm{d}y'$$

6.2.2 Chaos expansions

In this section we adopt the point of view of [101]. By construction the stochastic integral defined in eq. (6.8) is progressively measurable, so one can consider iterated stochastic integrals of the type

$$\int_{0<s<t} \int_{\mathbb{R}^2} \phi(\mathrm{d}\boldsymbol{s}, \mathrm{d}\boldsymbol{x})\, \xi^{\otimes 2}(\mathrm{d}\boldsymbol{s}, \mathrm{d}\boldsymbol{x}) := \int_0^t \int_0^s \int_{\mathbb{R}^2} \phi(s_1, s_2, x_1, x_2)\, \xi(\mathrm{d}s_1, \mathrm{d}x_1)\xi(\mathrm{d}s_2, \mathrm{d}x_2)$$

for an appropriate class of (random) functions $\phi : \mathbb{R}_+^2 \times \mathbb{R}^2 \to \mathbb{R}$.

For a fixed $t > 0$ and an integer $k \geq 1$, let

$$\Delta_k(t) := \left\{ \boldsymbol{t} = (t_1, \dots, t_k) \in \mathbb{R}_+^k : 0 < t_1 < \cdots < t_k < t \right\}. \tag{6.11}$$

In the case $t = \infty$ we simply write Δ_k.

The above construction can be iterated to define multiple integrals of the form

$$I_k(\phi)_t := \int_{\Delta_k(t)} \int_{\mathbb{R}^k} \phi(\boldsymbol{s}, \boldsymbol{x})\, \xi^{\otimes k}(\mathrm{d}\boldsymbol{s}, \mathrm{d}\boldsymbol{x})$$

such that

$$\mathbb{E}\left[I_k(\phi) I_j(\psi) \right] = (\phi, \psi)_{L^2(\Delta_k \times \mathbb{R}^k)} \mathbf{1}_{k=j} \tag{6.12}$$

and they span $L^2(\Omega, \mathbb{P})$. In this way one obtains an isometry between $L^2(\Omega, \mathbb{P})$ and

$$H_\infty := \bigoplus_{k=0}^{\infty} L^2(\Delta_k \times \mathbb{R}^k)$$

where the sum is to be understood in the Hilbert space sense, i.e. it is the completion of the algebraic direct sum with respect to the inner product on the right-hand side of eq. (6.12).

The above isometry implies the following characterization of elements in $L^2(\Omega)$: for each $F \in L^2(\Omega)$ there exists a sequence of functions $(f_n : n \geq 1)$ with $f_n \in L^2(\Delta_n \times \mathbb{R}^n)$ such that

$$F = \mathbb{E}F + \sum_{n=1}^{\infty} I_n(f_n). \tag{6.13}$$

Moreover, the orthogonality of the iterated integrals implies that

$$\mathbb{E}[F - \mathbb{E}(F)]^2 = \sum_{n=1}^{\infty} \|f_n\|^2_{L^2(\Delta_n \times \mathbb{R}^n)}. \tag{6.14}$$

The representation in eq. (6.13) is called the *Wiener chaos decomposition* of F. Observe that by definition, a random variable F is in H_∞ if and only if the series in eq. (6.14) converges.

6.3 Determinantal processes

In this section we summarise the principal results in the theory of determinantal processes needed for the following sections. Generally speaking, a determinantal process is a stochastic process whose distribution is characterised by the determinant of a fixed kernel. A well studied of such a process is the collection eigenvalues of a random matrix drawn from the Gaussian Unitary Ensemble.

Let S be a measurable space and consider the space $\boldsymbol{S}_k^{(n)} \equiv \{1, \ldots, k\} \times S^n$. We denote configurations $x \in \boldsymbol{S}_k^{(n)}$ as $x = (x_j^i : i = 1, \ldots, n; j = 1, \ldots, k)$. Given two functions $\varphi, \psi : S^2 \to \mathbb{R}$ we define their convolution with respect to a measure μ over S by

$$(\varphi * \psi)(x, y) \equiv \int_S \varphi(x, z)\psi(z, y)\, \mu(\mathrm{d}z).$$

Suppose we are given a sequence of functions $(\varphi_{r,r+1} : r = 0, \ldots, k)$ from S^2 to $\mathbb{R}$. Let, for any $r < s \in \{1, \ldots, k\}$,

$$\varphi_{r,s} \equiv \varphi_{r,r+1} * \varphi_{r+1,r+2} * \cdots * \varphi_{s-1,s}$$

and define $\varphi_{r,s}$ to be the zero function if $s \leq r$. Fix two configurations $x_0, x_{k+1} \in S^n$ and consider the weight of a configuration $x \in \boldsymbol{S}_k^{(n)}$ to be

$$\omega(x) \equiv \prod_{r=0}^{k} \det\Big(\varphi_{r,r+1}\Big(x_r^i, x_{r+1}^j\Big)\Big)_{i,j=1}^{n}. \tag{6.15}$$

We assume that the partition function

$$Z_k^{(n)} \equiv \frac{1}{(n!)^k} \int_{\mathbb{R}^{kn}} \omega(x)\, \mu_k^{(n)}(\mathrm{d}x)$$

is non-vanishing. Here we used the notation $\mu_k^{(n)}(\mathrm{d}x) \equiv \prod_{r=1}^k \prod_{i=1}^n \mu(\mathrm{d}x_r^i)$.

The correlation function is defined to be

$$R_k^{(n)}(x_1^1, \ldots, x_k^1) \equiv \frac{1}{((n-1)!)^k Z_k^{(n)}} \int_{\mathbb{R}^{(n-1)k}} \omega(x) \prod_{r=1}^{k} \mu_*^{(n)}(\mathrm{d}x_r) \tag{6.16}$$

where this time $\mu_*^{(n)}(\mathrm{d}x) \equiv \prod_{i=2}^n \mu(\mathrm{d}x^i)$, that is, we integrate out all but the first variable of each vector configuration.

Let A be the $n \times n$ matrix with entries $A_{ij} = \varphi_{0,k+1}(x_0^i, x_{k+1}^j)$. It can be shown [69] that $\det A = Z_k^{(n)}$ so our assumptions imply that A is invertible. Define a kernel $K^{(n)} : (\{1, \ldots, k\} \times \mathbb{R})^2 \to \mathbb{R}$ by

$$K^{(n)}(r, x; s, y) = -\varphi_{r,s}(x, y) + \sum_{i,j=1}^{n} \varphi_{r,k+1}(x, x_{k+1}^i)(A^{-1})_{ij} \varphi_{0,s}(x_0^j, y). \tag{6.17}$$

Proposition 6.3.1. *Suppose that $|\varphi_{r,s}(x, y)| \leq c_{r,s}(x) b_{r,s}(y)$ for all $r, s \in \{0, \ldots, k+1\}$ and some positive functions $c_{r,s} \in L^p(\mathbb{R})$ and $b_{r,s} \in L^\infty(\mathbb{R})$. Then, there exists a constant $C > 0$ such that*

$$\|R_k^{(n)}\|_{L^p(\mu^{\otimes k})} \leq C^k k^{k(2/p-1)} \prod_{r=1}^{k} \sum_{s=1}^{k} \|c_{r,s}\|_p \tag{6.18}$$

Proof. First note that the hypothesis implies there exists a constant $D > 0$ such that

$$|K(r, x; s, y)| \leq Dc(x)b(y). \tag{6.19}$$

Indeed, by definition

$$\begin{aligned}
|K^{(n)}(r, x; s, y)| &\leq |\varphi_{r,s}(x, y)| + \sum_{i,j=1}^{n} |\varphi_{r,k+1}(x, x_{k+1}^i)(A^{-1})_{ij} \varphi_{0,s}(x_0^j, y)| \\
&\leq c_{r,s}(x) b_{r,s}(y) \left(1 + \sum_{i,j=1}^{n} b_{s,k+1}(x_{k+1}^j)(A^{-1})_{ij} c_{0,r}(x_0^j) \right) \\
&= Dc(x)b(y)
\end{aligned}$$

Let $B = (K^{(n)}(r, x_r; s, x_s))_{r,s=1}^k$. Then, by Hadamard's inequality we have that there is another constant $C > 0$ such that

$$\begin{aligned} R_k^{(n)}(x_1, \dots, x_k) &= \det B \\ &\le \prod_{r=1}^{k} \left(\sum_{s=1}^{k} |K^{(n)}(r, x_r; s, x_s)|^p \right)^{1/p} \\ &\le C^k k^{k/p} \prod_{r=1}^{k} \sum_{s=1}^{k} c_{r,s}(x_r)^p. \end{aligned}$$

Hence

$$\int_{S^k} R_k^{(n)}(x_1, \dots, x_k)^p \, \mu^{\otimes k}(\mathrm{d}x) \le C^{kp} k^{k(2-p)} \prod_{r=1}^{k} \sum_{s=1}^{k} \|c_{r,s}\|_p^p.$$

□

Next we cite a very general theorem from [70].

Theorem 6.3.2. *Let $A_0, A_1, \dots, A_m, A_{m+1}$ be subsets of $\mathbb{R}$, $\varphi_{r,r+1} : A_r \to A_{r+1}$ be given functions and μ_r a measure on A_r. Let $x^0 \in A_0^n$ and $x_{k+1} \in A_{k+1}^n$ be given vectors. For $r < s$ define*

$$\varphi_{r,s}(x, y) = \int \varphi_{r,r+1}(x, z_1)\varphi_{r+1,r+2}(z_1, z_2) \dots \varphi_{s-1,s}(z_{r-s-1}, y) \, \mu_{r+1}(\mathrm{d}z_1) \dots \mu_{s-1}(\mathrm{d}z_{r-s-1}). \tag{6.20}$$

Consider the probability measure on $A_1^n \times \cdots \times A_m^n$ with density given by

$$\frac{1}{Z_m^{(n)}} \prod_{r=0}^{m} \det(\varphi_{r,r+1}(x_i^r, x_j^{r+1}))_{i,j=1}^n \, \mu_1(\mathrm{d}x^1) \dots \mu_k(\mathrm{d}x^m). \tag{6.21}$$

Then, the k-point correlation function of this probability measure is given by $\det(K_{n,m}(z_i, z_j))_{i,j=1}^k$ *where the kernel $K_{n,m}$ is given by*

$$K_{n,m}(r, x; s, y) = -\varphi_{r,s}(x, y) + \sum_{i,j=1}^{n} \varphi_{r,m}(x, x_i^m)(A^{-1})_{ij} \varphi_{0,s}(x_j^0, y) \tag{6.22}$$

where $A_{ij} = (\varphi_{0,m}(x_i^0, x_j^m))_{i,j=1}^n$.

Remark 6.3.3. We observe that the kernel in eq. (6.22) is not unique. In fact, we may multiply it by $g(t, y)/g(s, x)$ for an arbitrary non-zero function and obtain the same correlation function. This is so because we can factor out each of these factors inside the determinant and they will cancel out in the end. △

Finally, we state the Karlin–MacGregor formula [74], which can be seen as a generalisation of the inclusion-exclusion formula for probabilities. Given $n \in \mathbb{N}$, let Λ_n denote the *Weyl chamber*

$$\Lambda_n := \{\boldsymbol{x} \in \mathbb{R}^n : x_1 < \cdots < x_n\}. \tag{6.23}$$

Theorem 6.3.4 (Karlin–MacGregor)**.** *Let X be a Markov process with transition kernel p, and consider n independent copies $X^1, \dots, X^n$ of X. For any $t > 0$ and $\boldsymbol{x} \in \Lambda_n$ we have that*

$$\mathbb{P}_{\boldsymbol{x}}(X_t^1 \in \mathrm{d}y_1, \dots, X_t^n \in \mathrm{d}y_n \mid \tau > t) = \det\big(p(t, x_i, y_j)\big)_{i,j=1}^n \mathrm{d}y_1 \cdots \mathrm{d}y_n \tag{6.24}$$

where τ is the hitting time of the boundary $\partial\Lambda_n$.

6.3.1 Correlation functions

For the following we fix integers $n \geq 2$ and $N \geq n$, and set $q = N - n + 1$. Consider an ensemble of n non-intersecting Poisson processes $\boldsymbol{X}$ and n non-intersecting Brownian bridges $\boldsymbol{W}$. Given an k-tuple of indices $\boldsymbol{i} \in [\![1, n]\!]^k$, semi-discrete space-time points $u_1 = (t_1, j_1), \dots, u_k = (t_k, j_k) \in \mathbb{R}_+ \times \mathbb{N}$ and space-time points $z_1 = (t_1, x_1), \dots, z_k = (t_k, x_k) \in \mathbb{R}_+ \times \mathbb{R}$, a terminal time $T > 0$ and starting and ending vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}_+ \times \mathbb{N}$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}_+ \times \mathbb{R}$, define the functions

$$\rho_{\boldsymbol{i}}^{(n)}(u_1, \dots, u_k; T, \boldsymbol{u}, \boldsymbol{v}) = \mathbf{E}_{\boldsymbol{u}}\left[\prod_{\ell=1}^k \mathbf{1}_{\left\{X_{t_\ell}^{i_\ell} = j_\ell\right\}} \middle| \boldsymbol{X}_T = \boldsymbol{v}\right] \tag{6.25}$$

and let $R_{\boldsymbol{i}}^{(n)}(z_1, \dots, z_k; T, \boldsymbol{x}, \boldsymbol{y})$ be the joint density of $(W_{t_1}^{i_1}, \dots, W_{t_k}^{i_k})$ evaluated at $(x_1, \dots, x_k)$. Finally, define $\rho_k^{(n)}$ and $R_k^{(n)}$ as the sum over all $\boldsymbol{i} \in [\![1, n]\!]^k$ of $\rho_{\boldsymbol{i}}^{(n)}$ and $R_{\boldsymbol{i}}^{(n)}$ respectively. Note that $\rho_{\boldsymbol{i}}^{(n)}$ vanishes if $u_i = u_j$ for some $i \neq j$; also note that the product inside the expectation is non-vanishing if and only $i_\ell \leq j_\ell \leq N - n + i_\ell$

6.3.2 Dyson's Brownian motion

Before we can present the multilayer extension of this equation, we must give a brief review of the construction of non-intersecting Brownian bridges. Let $\Lambda_n = \{\boldsymbol{x} \in \mathbb{R}^n : x_1 < \cdots < x_n\}$ be the n-dimensional Weyl chamber. Define the function $\rho_*^{(n)}$ as

$$\rho_*^{(n)}(t, \boldsymbol{x}, \boldsymbol{y}) = \det\big(p(t, x_i, y_j)\big)_{i,j=1}^n. \tag{6.26}$$

A Brownian motion started at some point $\boldsymbol{x} \in \Lambda_n$ and killed when it first enters Λ_n^c is known as *Dyson's Brownian motion*. This is a homogeneous Markov process with transition densities given by

$$Q^{(n)}(t, \boldsymbol{x}, \boldsymbol{y}) \equiv \frac{h(\boldsymbol{y})}{h(\boldsymbol{x})} \rho_*^{(n)}(t, \boldsymbol{x}, \boldsymbol{y}) \mathbf{1}_{\Lambda_n^2}(\boldsymbol{x}, \boldsymbol{y}), \tag{6.27}$$

where $h(\boldsymbol{x}) = \det\left(x_i^{j-1}\right)_{i,j=1}^n$ is the Vandermonde determinant, which is the solution to Laplace's equation $\Delta h \equiv 0$ on Λ_n with zero boundary condition.

As the next proposition shows, it is also possible to start this process from points of the form $(x, \ldots, x) \in \mathbb{R}^n$. This result is relatively well known but we give a proof as a way to introduce some techniques that will be useful later. For convenience we let $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^n$.

Proposition 6.3.5. *The transition densities in eq.* (6.27) *converge pointwise, as $\boldsymbol{x} \to x\mathbf{1}$, to*

$$Q^{(n)}(t, x\mathbf{1}, \boldsymbol{y}) = c_{n,t} h(\boldsymbol{y})^2 \prod_{i=1}^{n} p(t, x, y_i) \mathbf{1}_{\Lambda_n}(\boldsymbol{y})$$

where

$$c_{n,t} = \left(t^{\frac{n(n-1)}{2}} \prod_{j=1}^{n-1} j! \right)^{-1}.$$

Proof. The Harish-Chandra—Itzykson–Zuber formula [62, 64] allows us to rewrite (6.27) as an integral over the *Unitary Group* $U(n)$,

$$Q^{(n)}(t, \boldsymbol{x}, \boldsymbol{y}) = \frac{1}{(2\pi t)^{\frac{n}{2}}} c_{n,t} h(\boldsymbol{y})^2 \int_{U(n)} e^{-\frac{1}{2t} \operatorname{tr}(\mathbf{Y} - U\mathbf{X}U^*)^2} \, \mathrm{d}U$$

for any pair of n by n matrices $\mathbf{X}$ and $\mathbf{Y}$ with spectra $\sigma(\mathbf{X}) = \boldsymbol{x}$ and $\sigma(\mathbf{Y}) = \boldsymbol{y}$. In particular, we may choose $\mathbf{X}$ and $\mathbf{Y}$ to be diagonal. In this case, the bound

$$\sup_{U \in U(n)} e^{-\frac{1}{2t} \operatorname{tr}(\mathbf{Y} - U\mathbf{X}U^*)^2} \leq \prod_{i=1}^{n} e^{-\frac{(y_i - x_i)^2}{2t}}$$

found in [94] shows that the integrand remains uniformly bounded over $U(n)$. Hence, the point-wise convergence

$$e^{-\frac{1}{2t} \operatorname{tr}(\mathbf{Y} - U\mathbf{X}U^*)^2} \longrightarrow \prod_{i=1}^{n} e^{-\frac{1}{2t}(y_i - x)^2}$$

as $\boldsymbol{x} \to x\mathbf{1}$ is enough to show the point-wise convergence of $Q^{(n)}$. □

Given a final time $t > 0$ and initial and final points $\boldsymbol{x}, \boldsymbol{y} \in \Lambda_n$, there exists a canonical way [19, 41] of constructing a Markovian bridge, that is to say, one can find a probability measure $\mathsf{P}_{\boldsymbol{x}}^{t,\boldsymbol{y}}$ over the space of continuous paths on the interval $[0, t]$ such that under it, the canonical process $\boldsymbol{X}_s$ on $C([0, t])^n$ is a Markov process and

$$\mathsf{P}_{\boldsymbol{x}}^{t,\boldsymbol{y}}(\boldsymbol{X}_0 = \boldsymbol{x}) = \mathsf{P}_{\boldsymbol{x}}^{t,\boldsymbol{y}}(\boldsymbol{X}_t = \boldsymbol{y}) = 1.$$

Moreover, the transition densities of this new process can be computed explicitly and are given by

$$\rho^{(n)}(s, z; t, \boldsymbol{x}, \boldsymbol{y}) = \frac{Q^{(n)}(s, \boldsymbol{x}, z) Q^{(n)}(t - s, z, \boldsymbol{y})}{Q^{(n)}(t, \boldsymbol{x}, \boldsymbol{y})} \mathbf{1}_{\Lambda_n}(z). \tag{6.28}$$

Another object which is useful in the study of this process is the *k-point correlation function* defined as follows. Fix an integer $k \geq 1$ and a multi-index $(i_1, \ldots, i_k) \in \{1, \ldots, n\}^k$. For $\boldsymbol{s} \in \Delta_k(t)$ and $\boldsymbol{z} \in \mathbb{R}^k$ let $\varrho_{i_1,\ldots,i_k}(\boldsymbol{s}, \boldsymbol{z}; t, x, y)$ be the joint marginal density of the vector $(X^{i_1}_{s_1}, \ldots, X^{i_k}_{s_k})$ evaluated at $(x_1, \ldots, x_k)$ and define

$$R^{(n)}_k(\boldsymbol{s}, \boldsymbol{z}; t, x, y) = \sum_{i_1,\ldots,i_k=1}^{n} \varrho_{i_1,\ldots,i_n}(\boldsymbol{s}, \boldsymbol{z}; t, x, y). \tag{6.29}$$

This defines a determinantal process, that is, $R^{(n)}_k$ can be expressed as a determinant of a fixed kernel K which was explicitly calculated in [70].

Theorem 6.3.6. *There exists a kernel $K^{t,y}_x$ such that $R^{(n)}_k(\cdot, \cdot; t, x, y)$ is given by the determinant*

$$R^{(n)}_k(\boldsymbol{s}, \boldsymbol{z}; t, x, y) = \det\Big(K^{t,y}_x(s_i, z_i; s_j, z_j)\Big)^k_{i,j=1}. \tag{6.30}$$

The explicit form of this kernel is given by

$$\begin{aligned} K^{t,y}_x(s, z; u, z') := &-\frac{1}{\sqrt{2\pi(u-s)}} e^{-\frac{(z'-z)^2}{2(u-s)}} \mathbf{1}_{s \leq u} \\ &+ \sum_{j=0}^{n-1} \frac{1}{j!} \left(\frac{s(t-u)}{u(t-s)}\right)^{j/2} \sqrt{\frac{t}{2\pi u(t-s)}} H_j(r(s)(z-x)) H_j(r(u)(y-z')) e^{-\frac{(z-x)^2}{2(t-s)} - \frac{(y-z')^2}{2u}} \end{aligned} \tag{6.31}$$

where

$$r(s) = \sqrt{\frac{t}{s(t-s)}}$$

and H_j is the j-th Hermite polynomial.

6.3.3 The Poisson bridge and the Sawtooth walk

Now let X be a rate 1 Poisson process and let $\boldsymbol{X} = (X^1, \ldots, X^n)$ denote a vector of n independent copies of X. It was shown in [75] that the Vandermonde determinant is harmonic for $\boldsymbol{X}$ on the Weyl chamber Λ_n. We can thus define the process conditioned (in the sense of Doob) to stay in the interior Λ°_n, and started from any $x \in \Lambda^\circ_n$, via its h-transform which we still denote by $\boldsymbol{X}$. We have then, by the Karlin–McGregor formula [74], that for any $x, y \in \Lambda^\circ_n$ and $t > 0$

$$\mathbf{P}_{\boldsymbol{x}}(\boldsymbol{X}_t = \boldsymbol{y}) = \frac{h(\boldsymbol{y})}{h(\boldsymbol{x})} \det\left(\frac{\mathrm{e}^{-t} t^{(y^j - x^i)}}{(y^j - x^i)!} \mathbf{1}_{\mathbb{N}}(y^j - x^i)\right)^n_{i,j}. \tag{6.32}$$

In some special cases this determinant can be explicitly computed. For example, if we set

$\boldsymbol{x}^* = (1, 2, \dots, n)$ then it was shown in [75] that

$$\mathbf{P}_{\boldsymbol{x}_*}(\boldsymbol{X}_t = \boldsymbol{y}) = c_{n,t} h(\boldsymbol{y})^2 \mathrm{e}^{-nt} \prod_{i=1}^{n} \frac{t^{y^i}}{y^i!} \tag{6.33}$$

where $c_{n,t} = t^{-n(n-1)/2} \left(\prod_{j=0}^{n-1} j!\right)^{-1}$. We will also need the fact that if $\boldsymbol{y}^* = (N - n + 1, \dots, N)$ for some $N > n$ then

$$\mathbf{P}_{\boldsymbol{x}}(\boldsymbol{X}_t = \boldsymbol{y}_*) = c_{n,t} \mathrm{e}^{-nt} \left(\prod_{j=0}^{n-1} j! \right)^2 \prod_{i=1}^{n} \frac{t^{N-x^i}}{(N - x^i)!}. \tag{6.34}$$

In particular, from any of these two formulas we deduce that

$$\mathbf{P}_{\boldsymbol{x}_*}(\boldsymbol{X}_t = \boldsymbol{y}_*) = \mathrm{e}^{-nt} t^{n(N-n)} \prod_{j=0}^{n-1} \frac{j!}{(N-j)!}. \tag{6.35}$$

Given a fixed $N \geq n$ as above, consider the shear transformation

$$\varphi_N^{(n)}(t, \boldsymbol{x}) = \left(\frac{t}{N}, \frac{t - x_1}{\sqrt{N}}, \dots, \frac{t - x_n}{\sqrt{N}} \right)$$

which is the multidimensional analogue of the one introduced in [93]. The *multilayer sawtooth process* $\boldsymbol{S}$ is obtained as the image under $\varphi_N^{(n)}$. It is a Markov process with transition function given by

$$\rho_N(s, \boldsymbol{x}; t, \boldsymbol{y}) := \mathbf{P}(\boldsymbol{S}_t = \boldsymbol{y} \mid \boldsymbol{S}_s = \boldsymbol{x}) = \mathbf{P}_{\hat{\boldsymbol{x}}^s}(\boldsymbol{X}_{N(t-s)} = \hat{\boldsymbol{y}}^t)$$

where we have introduced the notation $\hat{\boldsymbol{x}}^s = (\lceil \sqrt{N} x^1 + Ns \rceil, \dots, \lceil \sqrt{N} x^n + Ns \rceil)$.

For $j \in \mathbb{N}$ and $0 < s < t$ let $I_N^{(j)}(s,t)$ denote the image of the rectangle $[j-1, j] \times [s, t]$ under φ_N. For $u \in [s/N, t/N]$ the intervals $I_N^{(j)}(u) := \left[\frac{j-1}{\sqrt{N}} - \sqrt{N} u, \frac{j}{\sqrt{N}} - \sqrt{N} u \right]$ are such that

$$I_N^{(j)}(s,t) = \bigcup_{u \in [s/N, t/N]} I_N^{(j)}(u).$$

Observe that these intervals are such that $\hat{x}^s = j$ if and only if $x \in I_N^{(j)}(s)$. We also have that $|I_N^{(j)}(s)| = N^{-1/2}$ for all $s > 0$ and $j \in \mathbb{Z}$.

From this process we can build a Markovian bridge between to given configurations $\boldsymbol{x}, \boldsymbol{y} \in \Lambda_n^\circ$ over the interval $(0, T)$. That is, we can construct a probability measure $\mathbf{P}_{\boldsymbol{x}}^{T,\boldsymbol{y}}$ over paths such that

$$\mathbf{P}_{\boldsymbol{x}}^{T,\boldsymbol{y}}(\boldsymbol{S}_0 = \boldsymbol{x}) = \mathbf{P}_{\boldsymbol{x}}^{T,\boldsymbol{y}}(\boldsymbol{S}_T = \boldsymbol{y}) = 1,$$

the process $\boldsymbol{S}$ is still a Markov process under $\mathbf{P}_{\boldsymbol{x}}^{T,\boldsymbol{y}}$ and moreover, its finite-dimensional distributions are

$$\mathbf{P}^{T,\boldsymbol{y}}_{\boldsymbol{x},N}(\boldsymbol{S}_{t_1} = \boldsymbol{x}_1, \ldots, \boldsymbol{S}_{t_k} = \boldsymbol{x}_k) = \frac{1}{\rho_N(0; \boldsymbol{x}, T, \boldsymbol{y})} \prod_{j=1}^{k+1} \rho_N(t_{j-1}, \boldsymbol{x}_{j-1}; t_j, \boldsymbol{x}_j) \tag{6.36}$$

where for convenience we have set $t_0 = 0$, $t_{k+1} = T$ and $\boldsymbol{x}_0 = \boldsymbol{x}$, $\boldsymbol{x}_{k+1} = \boldsymbol{y}$. We consider the rescaled distributions

$$\frac{N^{k/2}}{\rho_N(0; \boldsymbol{x}, 1, \boldsymbol{y})} \prod_{j=1}^{k+1} \rho_N(t_{j-1}, \boldsymbol{x}_{j-1}; t_j, \boldsymbol{x}_j) \tag{6.37}$$

where we have taken $T = 1$ for simplicity. There's also an associated correlation function

$$\varrho^{(n)}_{k,N}(\boldsymbol{s}, \boldsymbol{x}; 1, \boldsymbol{u}, \boldsymbol{v}) := N^{k/2} \rho^{(n)}_k((Ns_1, \hat{x}^{s_1}_1), \ldots, (Ns_k, \hat{x}^{s_k}_k); 1, \hat{\boldsymbol{u}}^0, \hat{\boldsymbol{v}}^1) \tag{6.38}$$

for $\boldsymbol{s} \in \Delta_k(1)$ and $\boldsymbol{x} \in \mathbb{R}^k$.

Now we derive a formula for the kernel describing eq. (6.38) as a determinant. Consider a sequence of times $0 < t_1 < \cdots < t_k < 1$ and configurations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \Lambda^\circ_n$. Then, eqs. (6.32) to (6.35) imply that in the case where we take $\boldsymbol{a}_N = \left(\frac{1}{\sqrt{N}}, \ldots, \frac{n}{\sqrt{N}}\right)$ and $\boldsymbol{b}_N = \left(-\frac{n-1}{\sqrt{N}}, \ldots, 0\right)$ so that $\hat{\boldsymbol{a}}^0_N = \boldsymbol{x}^*$ and $\hat{\boldsymbol{b}}^1_N = \boldsymbol{y}^*$, eq. (6.37) becomes an expression of the form eq. (6.21) provided that we set

$$\begin{aligned} \varphi_{0,1}(i, y) &= \sqrt{N} q_{i-1}(\hat{y}^{t_1}) \frac{t_1^{\hat{y}^{t_1} - i + 1}}{\hat{y}^{t_1}!}, \quad \varphi_{r,r+1}(x, y) = \sqrt{N} \frac{(t_{r+1} - t_r)^{\hat{y}^{t_{r+1}} - \hat{x}^{t_r}}}{(\hat{y}^{t_{r+1}} - \hat{x}^{t_r})!} \\ \varphi_{k,k+1}(x, j) &= \sqrt{N} \tilde{q}_{j-1}(\hat{x}^{t_k}) \frac{(N - j + 1)!(1 - t_k)^{N - \hat{x}^{t_k} - j + 1}}{(N - \hat{x}^{t_k})!}. \end{aligned} \tag{6.39}$$

for some families of monic polynomials $q_j, \tilde{q}_j$ such that $\det(q_j) = \deg(\tilde{q}_j) = j$. This is so because it is a standard fact (which follows by making row manipulations) that h can be written as $h(\boldsymbol{x}) = \det(q_{j-1}(x^i))$ for any such family.

Lemma 6.3.7. *For each $1 \leq r < s \leq k$ and $x, y \in \mathbb{R}$ we have that*

$$\varphi_{r,s}(x, y) = \sqrt{N} \frac{(t_s - t_r)^{\hat{y}^{t_s} - \hat{x}^{t_r}}}{(\hat{y}^{t_s} - \hat{x}^{t_r})!}.$$

Proof. We prove this formula by induction on $m = s - r \in \{1, \ldots, k-1\}$, the case $m = 1$ being true by definition. For the induction step, consider

$$\begin{aligned} \varphi_{r,s+1}(x, y) &= \int_{\mathbb{R}} \mathrm{d}z \ \ \varphi_{r,s}(x, z) \varphi_{s,s+1}(z, y) \\ &= N \int_{\mathbb{R}} \mathrm{d}z \ \ \frac{(t_s - t_r)^{\hat{z}^{t_s} - \hat{x}^{t_r}}}{(\hat{z}^{t_s} - \hat{x}^{t_r})!} \frac{(t_{s+1} - t_s)^{\hat{y}^{t_{s+1}} - \hat{z}^{t_s}}}{(\hat{y}^{t_{s+1}} - \hat{z}^{t_s})!}. \end{aligned}$$

Observe that for each fixed $s > 0$ the union

$$\bigcup_{k\in\mathbb{Z}} I_N^{(k)}(s) = \mathbb{R}.$$

Without loss of generality we may also assume that $x \in I_N^{(a)}(t_r)$ and $y \in I_N^{(b)}(t_{s+1})$ for some $a, b \in \mathbb{Z}$. Therefore,

$$\begin{aligned}
\varphi_{r,s+1}(x,y) &= \sqrt{N}\frac{1}{(b-a)!}\sum_{k=a}^{b}\binom{b-a}{k-a}(t_s-t_r)^{k-a}(t_{s+1}-t_s)^{b-k} \\
&= \sqrt{N}\frac{(t_{s+1}-t_r)^{b-a}}{(b-a)!} \\
&= \sqrt{N}\frac{(t_{s+1}-t_r)^{\hat{y}^{t_{s+1}}-\hat{x}^{t_r}}}{(\hat{y}^{t_{s+1}}-\hat{x}^{t_r})!}
\end{aligned}$$

which ends the proof. □

To compute the matrix A whose inverse appears in the expression for the kernel in eq. (6.17), we have to require some structure on the polynomials q_j and $\tilde{q}_j$. It turns out that the correct family is given by the *Krawtchouk polynomials* which we now define.

Given $N \in \mathbb{N}$ and $p, q \in (0,1)$ such that $p + q = 1$, consider the binomial weight

$$\mathrm{j}(x) = \binom{N}{x}p^x(1-p)^{N-x}, \quad x = 0, 1, \ldots, N.$$

The Krawtchouk polynomials $q_n(x; N, p)$ for $n = 0, \ldots, N$ are the orthogonal polynomial family with respect to the weight j. They may be obtained by means of their generating function [114]

$$G(x,z;N,p) \equiv \sum_{n=0}^{N}\frac{z^n}{n!}q_n(x;N,p) = (1+qz)^x(1-pz)^{N-x} \tag{6.40}$$

Note the symmetry $q_n(N-x; N, 1-p) = (-1)^n q_n(x; N, p)$ which follows from eq. (6.40).

Lemma 6.3.8. *Fix $r, s > 0$ such that $\frac{s}{r+s} > p$ and $y, n \in \{0, \ldots, N\}$, then*

$$\sum_{x=0}^{y}\binom{y}{x}s^x r^{y-x}q_n(x;N,p) = (r+s)^y\left(\frac{s}{s+r}\right)^n q_n\left(y;N,p\frac{r+s}{s}\right).$$

Proof. Consider the sum

$$\sum_{x=0}^{y}\binom{y}{x}s^x r^{y-x}G(x,z;N,p) = (1-pz)^{N-y}\left(1+z\frac{sq-rp}{s+r}\right)^y(s+r)^y = (s+r)^y G\left(y,\frac{s}{r+s}z;N,p\frac{r+s}{s}\right).$$

On the other hand, this sum equals

$$\sum_{n=0}^{N} \frac{z^n}{n!} \sum_{x=0}^{y} \binom{y}{x} s^x r^{y-x} q_n(x; N, p)$$

so the result follows by comparing the coefficients of z^n on both sides. □

By a similar technique we may also obtain

Lemma 6.3.9. *Fix $r, s > 0$ such that $\frac{r}{r+s} > p$ and $y, n \in \{0, \ldots, N\}$, then*

$$\sum_{x=y}^{N} \binom{N-y}{x-y} s^{x-y} r^{N-x} q_n(N-x; N, p) = (r+s)^y \Big(\frac{r}{s+r}\Big)^n q_n\Big(N-y; N, p\frac{r+s}{r}\Big).$$

Furthermore, these polynomials satisfy the orthogonality condition

$$\sum_{x=0}^{N} q_n(x; N, p) q_m(x; N, p)\, \mathrm{j}(x) = \binom{N}{n} (pq)^n \delta_{nm} \tag{6.41}$$

which follows again from eq. (6.40) by a similar argument.

Lemma 6.3.10 (Binomial local limit theorem)**.** *Given $N \in \mathbb{N}$, $p \in (0,1)$ and $x \in \mathbb{R}$ let $\hat{x} = \lceil \sqrt{N}x + Np \rceil$. Then*

$$\sqrt{N} \binom{N}{\hat{x}} p^{\hat{x}} (1-p)^{N-\hat{x}} \to \frac{1}{\sqrt{2\pi p}} e^{-\frac{x^2}{2p}}$$

Proof. Let W be a random variable having a binomial distribution with parameters N and p defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For $\vartheta \in \mathbb{R}$, the characteristic function $\phi(\vartheta) := \mathbb{E}\mathrm{e}^{\mathrm{i}\vartheta W} = (1 - p + p\mathrm{e}^{\mathrm{i}\vartheta})^N$ is π periodic since $W \in \mathbb{N}$ so that

$$\begin{aligned}
\mathrm{j}(x) = \mathbb{P}(W = x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 - p + p\mathrm{e}^{\mathrm{i}\vartheta})^N \mathrm{e}^{-\mathrm{i}\vartheta x} \, \mathrm{d}\vartheta \\
&= \frac{1}{2\pi\sqrt{N}} \int_{-\pi\sqrt{N}}^{\pi\sqrt{N}} \Big(1 - p + p\mathrm{e}^{\mathrm{i}\frac{\vartheta}{\sqrt{N}}}\Big)^N \mathrm{e}^{-\mathrm{i}\frac{\vartheta}{\sqrt{N}}x} \, \mathrm{d}\vartheta
\end{aligned}$$

Taylor expansion up to order two of the complex exponential gives there is a function $g(\vartheta)$ such that $\mathrm{e}^{\mathrm{i}\frac{\vartheta}{\sqrt{N}}} = 1 + \mathrm{i}\frac{\vartheta}{\sqrt{N}} - \frac{\vartheta^2}{2N} + g(\vartheta)$ and $\frac{g(\vartheta)}{\vartheta^2} \to 0$ as $\vartheta \to 0$. A similar argument gives that

$$\Big(1 - p + p\mathrm{e}^{\mathrm{i}\frac{\vartheta}{\sqrt{N}}}\Big)^N = \mathrm{e}^{\mathrm{i}\vartheta p\sqrt{N} - \frac{p\vartheta^2}{2}} \mathrm{e}^{g(\vartheta)}.$$

Without loss of generality we may assume that $-p\sqrt{N} \le x \le (1-p)\sqrt{N}$ so that $\hat{x} \in \{0, \dots, N\}$. Therefore,

$$\sqrt{N}\mathrm{j}(\hat{x}) \approx \frac{1}{2\pi}\int_{-\pi\sqrt{N}}^{\pi\sqrt{N}} \mathrm{e}^{-\frac{p\vartheta^2}{2}}\mathrm{e}^{-\mathrm{i}\vartheta x}\,\mathrm{d}\vartheta \to \frac{1}{\sqrt{2\pi p}}\mathrm{e}^{-\frac{x^2}{2p}}.$$

$\square$

In view of these properties we set $q_j = q_j(y; N, t_1/T)$ and $\tilde{q}_j = q_j(N - x; N, 1 - t_k)$. We have

Lemma 6.3.11. *For any $x, y \in \{1, \dots, N\}$ and $r, s \in \{1, \dots, k\}$,*

$$\varphi_{0,r}(i, y) = \sqrt{N}\frac{t_r^{\hat{y}^{t_r}-i+1}}{\hat{y}^{t_r}!(i-1)!}q_{i-1}(\hat{y}^{t_r}; N, t_r),$$
$$\varphi_{s,k+1}(x, j) = \sqrt{N}\frac{(N-j+1)!(1-t_s)^{N-\hat{x}^{t_s}-j+1}}{(N-\hat{x}^{t_s})!}q_{j-1}(N-\hat{x}^{t_s}; N, 1-t_s)$$

Proof. Without loss of generality we may assume that $y \in I_N^{(b)}(t_r)$ for some $b \in \mathbb{Z}$. By definition

$$\begin{aligned}\varphi_{0,r}(i, y) &= \int_{\mathbb{R}} \mathrm{d}z \;\; \varphi_{0,1}(i, z)\varphi_{1,r}(z, y) \\ &= \sqrt{N}\sum_{k=0}^{b} \frac{t_1^{k-i+1}}{(i-1)!k!}q_{i-1}(k; N, t_1)\frac{(t_r-t_1)^{b-k}}{(b-k)!}\end{aligned}$$

and by Lemma 6.3.8 this equals

$$\begin{aligned}\varphi_{0,r}(i, y) &= \sqrt{N}\frac{t_r^{b-i+1}}{b!(i-1)!}q_{i-1}(b; N, t_r) \\ &= \sqrt{N}\frac{t_r^{\hat{y}^{t_r}-i+1}}{\hat{y}^{t_r}!(i-1)!}q_{i-1}(\hat{y}^{\hat{y}^{t_r}} N, t_r).\end{aligned}$$

A similar computation gives the second identity, using Lemma 6.3.9. $\square$

Lemma 6.3.12. *For any $x, y \in \{1, \dots, N\}$,*

$$\varphi_{0,k+1}(i, j) = \frac{\sqrt{N}(-1)^{i-1}}{[(i-1)!]^2}\delta_{ij}.$$

Proof. Using the identities from the two previous Lemmas we get

$$
\begin{aligned}
\varphi_{0,k+1}(i,j) &= \int_{\mathbb{R}} \mathrm{d}z \;\; \varphi_{0,k}(i,z)\varphi_{k,k+1}(z,j) \\
&= \sqrt{N}\sum_{k=0}^{N} \frac{t_r^{k-i+1}}{(i-1)!k!} q_{i-1}(k;N,t_k) \frac{(N-j-1)!(1-t_k)^{N-k-j+1}}{(N-k)!} q_{j-1}(N-k;N,1-t_k) \\
&= \frac{\sqrt{N}(-1)^{j-1}(N-j+1)!}{N!(i-1)!t_k^{i-1}(1-t_k)^{j-1}} \sum_{k=0}^{N} \binom{N}{k} q_{i-1}(k;N,t_k) q_{j-1}(k;N,t_k)\mathrm{j}(k) \\
&= \frac{\sqrt{N}(-1)^{i-1}}{[(i-1)!]^2}\delta_{ij}
\end{aligned}
$$

□

In particular, Lemma 6.3.12 implies that the matrix A is diagonal so its inverse is trivial to compute. Hence the rescaled kernel for the Sawtooth bridge is given explicitly by

$$
\begin{aligned}
K_N^{(n)}(r,x;s,y) = {}& -\sqrt{N}\frac{(t_r-t_s)^{\hat{y}^{t_r}-\hat{x}^{t_s}}}{(\hat{y}^{t_r}-\hat{x}^{t_s})!}\mathbf{1}_{x\le y} \\
&+ \sum_{i=0}^{n-1} \sqrt{N}\frac{(N-i)!(1-t_r)^{N-\hat{x}^{t_r}-i}}{i!(N-\hat{x}^{t_r})!} \frac{t_s^{\hat{y}^{t_s}-i}}{\hat{y}^{t_s}!} q_i(\hat{x}^{t_r};N,t_r) q_i(\hat{y}^{t_s};N,t_s),
\end{aligned}
\tag{6.42}
$$

that is,

$$
N^{k/2}\varrho_{k,N}^{(n)}(\boldsymbol{s},\boldsymbol{x};1,\boldsymbol{a}_N,\boldsymbol{b}_N) = \det\Big(K_N^{(n)}(s_i,x_i;s_j,x_j)\Big)_{i,j=1}^{k}.
$$

In order to relate this kernel with the extended Hermite kernel appearing in eq. (6.31) we will need the following

Lemma 6.3.13. *Fix $x \in \mathbb{R}$ and $0 < t < 1$. Then*

$$
N^{-i/2}q_i(\hat{x};N,t) \to (t(1-t))^{i/2}H_i\left(\frac{x}{\sqrt{t(1-t)}}\right).
$$

Proof. Let $L_N(x,z) = \log G(\sqrt{N}x + Nt, z/\sqrt{N}; N,t)$ where G is the generating function in eq. (6.40). Then, by Taylor expanding the logarithm up to second order we have

$$
L_N(x,z;t) = xz - t(1-t)\frac{z^2}{2} + \mathrm{o}(1).
$$

Thus, since

$$
e^{L_N(x,z;t)} \le G(\hat{x}, z/\sqrt{N}; N,t) \le e^{L_N(x,z)}(1+\mathrm{o}(1))
$$

we get that

$$G(\hat{x}, z/\sqrt{N}; N, t) \to \exp\left(xz - t(1-t)\frac{z^2}{2}\right)$$

from where we deduce the result by identifying the limit as the generating function of the Hermite polynomials. □

Then we have,

Theorem 6.3.14. *For all* $0 < s < t$ *and* $x, y \in \mathbb{R}$ *we have*

$$\mathrm{e}^{-N(t-s)} N^{\sqrt{N}(y-x)+N(t-s)} K_N^{(n)}(s, x; t, y) \to K_0^{1,0}(s, x; t, y).$$

Proof. A simple computation using eq. (6.42) yields

$$\mathrm{e}^{-N(t-s)} N^{\sqrt{N}(y-x)+N(t-s)} K_N^{(n)}(s, x; t, y) = -A_N + B_N + \mathrm{o}(1)$$

where

$$A_N = \sqrt{N}\mathrm{e}^{-N(t-s)} \frac{[N(t-s)]^{\hat{y}^t - \hat{x}^s}}{(\hat{y}^t - \hat{x}^s)!}$$

and

$$B_N = \sum_{i=0}^{n-1} \frac{(N-i)!\mathrm{e}^N}{N^{N+i+1/2} \cdot i!} \cdot \frac{\sqrt{N}\mathrm{e}^{-N(1-s)}(N-Ns)^{N-\hat{x}^s}}{(N-\hat{x}^s)!} \cdot \frac{\sqrt{N}\mathrm{e}^{-Nt}(Nt)^{\hat{y}^t}}{\hat{y}^t!} \cdot \frac{N^{-i/2} q_i(\hat{x}^s; N, s)}{(1-s)^i} \cdot \frac{N^{-i/2} q_i(\hat{y}^t; N, t)}{t^i}$$

The local limit theorem for Poisson random variables implies that

$$A_N \to \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(y-x)^2}{2(t-s)}\right).$$

By Stirling's approximation we also have that

$$\frac{(N-i)!\mathrm{e}^N}{N^{N+i+1/2}} = \prod_{j=0}^{i-1}\left(1 - \frac{j}{N}\right) \cdot \frac{N! e^N}{N^{N+\frac{1}{2}}} \to \sqrt{2\pi}$$

and an application of Lemma 6.3.13 gives that

$$B_N \to \sum_{i=0}^{n-1} \frac{\mathrm{e}^{-\frac{x^2}{2(1-s)}} \mathrm{e}^{-\frac{y^2}{2t}}}{i!\sqrt{2\pi t(1-s)}} \left(\frac{s(1-t)}{t(1-s)}\right)^{i/2} H_i\left(\frac{x}{\sqrt{s(1-s)}}\right) H_i\left(\frac{y}{\sqrt{t(1-t)}}\right)$$

which is precisely eq. (6.31) evaluated at the appropriate points.

□

Finally, as a consequence we obtain

Theorem 6.3.15. *The rescaled correlation functions of the sawtooth walk converge to the correlation function of Dyson's Brownian motion; that is, for all $k \in \mathbb{N}$, $\boldsymbol{s} \in \Delta_k$ and $\boldsymbol{x} \in \mathbb{R}^k$ the convergence*

$$\varrho^{(n)}_{k,N}(\boldsymbol{s}, \boldsymbol{x}; N, \boldsymbol{a}_N, \boldsymbol{b}_N) \to R^{(n)}_k(\boldsymbol{s}, \boldsymbol{x}; 1, 0, 0)$$

holds, where $\boldsymbol{a}^*_N = (1/\sqrt{N}, \dots, n/\sqrt{N})$ and $\boldsymbol{b}^*_N = ((1-n)/\sqrt{N}, \dots, 0)$.

Proof. Following Remark 6.3.3 we may write

$$\varrho^{(n)}_{k,N}(\boldsymbol{s}, \boldsymbol{x}; 1, \boldsymbol{a}_N, \boldsymbol{b}_N) = \det\Big(\mathrm{e}^{-N(s_j - s_i)} N^{\sqrt{N}(x_j - x_i) - N(s_j - s_i)} K^{(n)}_N(s_i, x_i; s_j, x_j)\Big)^k_{i,j=1}$$

thus Theorem 6.3.14 gives the result. □

6.4 The Stochastic Heat Equation

Consider the classical *multiplicative stochastic heat equation* at inverse temperature $\beta > 0$

$$\partial_t \mathcal{Z}_\beta = \frac{1}{2}\Delta_y \mathcal{Z}_\beta + \beta \mathcal{Z}_\beta \xi, \quad \mathcal{Z}_\beta(0, x, y) = \delta_x(y)$$

where ξ is a space-time white noise in $1 + 1$ dimensions. We consider *mild solutions* to this equation which are given by

$$\mathcal{Z}_\beta(t, x, y) = p(t, x, y) + \beta \int_0^t \int_{\mathbb{R}} p(t - s, y', y)\mathcal{Z}_\beta(s, x, y')\, \xi(ds, dy') \tag{6.43}$$

where $p(t, x, y) = \frac{1}{\sqrt{2\pi t}} \exp\Big(-\frac{1}{2t}(y - x)^2\Big)\mathbf{1}_{[0,\infty)}(t)$ is the heat kernel associated to section 6.4. Upon iteration of eq. (6.43) it is possible to obtain a closed expression for the solution to section 6.4 as a polynomial chaos expansion,

$$\mathcal{Z}_\beta(t, x, y) = p(t, x, y)\left(1 + \sum_{k \geq 1} \beta^k \int_{\Delta_k(t)} \int_{\mathbb{R}^k} R_k(\boldsymbol{s}, \boldsymbol{x})\, \xi^{\otimes k}(d\boldsymbol{s}, d\boldsymbol{x})\right). \tag{6.44}$$

In this identity we defined the set $\Delta_k(t) = \big\{\boldsymbol{s} \in \mathbb{R}^k : 0 < s_1 < \cdots < s_k < t\big\}$ and R_k is the k-point correlation function of the Brownian bridge between $(0, x)$ and (t, y),

$$R_k(\boldsymbol{s}, \boldsymbol{x}) \equiv \frac{p(s_1, x, x_1)p(s_2 - s_1, x_1, x_2) \dots p(t - s_k, x_k, y)}{p(t, x, y)}.$$

By using standard analysis tools, the series in eq. (6.44) can be shown to be convergent in $L^2(\mathbb{P})$ [2] so this is a meaningful solution.

This suggests the interpretation of the solution of section 6.4 as the partition function for the continuum random polymer in a random environment given by the white noise ξ

$$\mathcal{Z}_\beta(t,x,y) = p(t,x,y)\mathsf{E}\colon \exp : \left(\beta \int_0^t \xi(s,X_s)\,\mathrm{d}s\right)$$

where X is a Brownian bridge under P (and E is the associated expectation), going from x at time 0 to y at time y. The appearance of the *Wick exponential* $:\exp:$ is, roughly, due to the time ordering of the iterated integrals appearing in the chaos expansion (6.44). Although not rigorous, this suggests a relationship between solutions of a SPDE and the partition function of a polymer model (the Albert–Khanin–Quastel continuum random polymer): in this model, one can think of Brownian paths as polymer paths to which one assigns the weight $\int_0^t \xi(s,X_s)\,\mathrm{d}s$, that is, space-time white noise acts as the environment.

The random field $\mathcal{Z}_\beta$ enjoys several properties, some of which are listed below.

Theorem 6.4.1. *Let $\beta > 0$. The random field $\mathcal{Z}_\beta$, which is the mild solution to section 6.4 satisfies:*

1. *$\mathcal{Z}_\beta$ has regularity $\frac{1}{2}-$ in space and $\frac{1}{4}-$ in time [101],*
2. *$\mathcal{Z}_\beta(t,x,y)$ has the same distribution as $\mathcal{Z}_\beta(t,0,y-x)$,*
3. *$\mathcal{Z}_\beta(t,0,y)$ has the same distribution as $\mathcal{Z}_1(\beta^2 t,0,\beta y)$*

6.4.1 The multilayer extension of the SHE

In [98] the following extension of eq. (6.44) is considered. Fix a parameter $\beta > 0$ and define, for each integer $n \geq 1$ the random field

$$\mathcal{Z}_\beta^{(n)}(t,x,y) := p(t,x,y)^n\left(1 + \sum_{k\geq 1}\beta^k \int_{\Delta_k(t)}\int_{\mathbb{R}^k} R_k^{(n)}(\boldsymbol{s},\boldsymbol{x};t,x,y)\,\xi^{\otimes k}(d\boldsymbol{s},d\boldsymbol{x})\right). \tag{6.45}$$

We can interpret this definition formally as follows: as in the $n=1$ case, the series appearing on the right-hand side can be seen as the Wick exponential

$$\mathsf{E}\colon\exp\colon\left(\beta\sum_{j=1}^n \int_0^t \xi(s,X_s^i)\,\mathrm{d}s\right)$$

where $(X^1,\dots,X^n)$ is a collection of non-intersecting Brownian bridges between $(0,x)$ and (t,y), see Section 6.3.2.

It can be shown [93, 98] that this series is convergent in $L^2(\xi)$ so this is a well defined object, called the *multilayer extension of the (solution of)* (6.4). It is immediate to check that $u_\beta^{(1)} = \mathcal{Z}_\beta$.

It can be shown [84] that $u_\beta^{(n)}$ has a jointly continuous version which is positive for all $t > 0$ and $x, y \in \mathbb{R}$ almost surely.

Sometimes, it is more useful to consider the following *intermediate objects*. Define

$$\mathscr{K}_\beta^{(n)}(t, \boldsymbol{x}, \boldsymbol{y}) \equiv \rho_*^{(n)}(t, \boldsymbol{x}, \boldsymbol{y}) \left(1 + \sum_{k \geq 1} \beta^k \int_{\Delta_k(t)} \int_{\mathbb{R}^k} R_k^{(n)}(\boldsymbol{s}, z; t, \boldsymbol{x}, \boldsymbol{y}) \, \xi^{\otimes k}(d\boldsymbol{s}, dz) \right) \tag{6.46}$$

and

$$\mathscr{M}_\beta^{(n)}(t, \boldsymbol{x}, \boldsymbol{y}) \equiv \frac{\mathscr{K}_\beta^{(n)}(t, \boldsymbol{x}, \boldsymbol{y})}{h(\boldsymbol{x}) h(\boldsymbol{y})}. \tag{6.47}$$

Here, $\rho_*^{(n)}$ is the transition kernel of a collection of n Brownian bridges startig from $\boldsymbol{x}$ and ending at $\boldsymbol{y}$, coditioned to non-intersection over the interval $[0, t]$. This will be carefully defined in Section 6.3.2 below. The field $\mathscr{K}_\beta^{(n)}$ corresponds to a multilayer polymer where the starting and end points are separated. Then, it can be shown that $\mathcal{K}_\beta^{(n)}$ is again well defined [98] and $\mathcal{M}_\beta^{(n)}$ has a continuous version such that

$$\mathcal{M}_\beta^{(n)}(t, \boldsymbol{x}, \boldsymbol{y}) \to c_{n,t} \mathscr{Z}_\beta^{(n)}(t, a, b) \tag{6.48}$$

as $\boldsymbol{x} \to a\mathbf{1}$ and $\boldsymbol{y} \to b\mathbf{1}$. In [83], C. H. Lun and J. Warren further show that $\mathscr{Z}_\beta^{(n)}$ has a continuous version which is a.s. positive for all $t > 0$ and $x, y \in \mathbb{R}$.

At present it is not known whether the processes $\mathscr{Z}_\beta^{(n)}$ satisfy a SPDE. However, some partial results have been obtained in this direction. In [98] it is proven that, at least when ξ is replaced by a smooth function $\phi : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ the ratios $u_\beta^{(n)} \equiv \frac{\mathscr{Z}_\beta^{(n)}}{\mathscr{Z}_\beta^{(n-1)}}$ satisfy

$$\partial_t u_\beta^{(n)} = \frac{1}{2} \Delta_y u_\beta^{(n)} + \beta \left(\phi + \Delta_y \log \left(\frac{\mathscr{Z}_\beta^{(n-1)}}{p^{n-1}} \right) \right) u_\beta^{(n)}, \quad u_\beta^{(n)}(0, x, y) = \delta(y - x) \tag{6.49}$$

Furthermore, the process $\mathscr{M}_\beta^{(n)}$ the process satisfies the mild equation

$$\mathscr{M}_\beta^{(n)}(t, \boldsymbol{x}, \boldsymbol{y}) = \frac{\rho_*^{(n)}(t, \boldsymbol{x}, \boldsymbol{y})}{h(\boldsymbol{x}) h(\boldsymbol{y})} + \frac{\beta}{(n-1)!} \int_0^t \int_{\mathbb{R}^n} Q^{(n)}(t - s, \boldsymbol{y}, \boldsymbol{y}') \mathscr{M}_\beta^{(n)}(s, \boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}'_* \xi(\mathrm{d}s, \mathrm{d}y'_1)$$

where $\mathrm{d}\boldsymbol{y}'_* = \mathrm{d}y'_2 \cdots \mathrm{d}y'_n$ and $Q^{(n)}$ is the kernel defined in eq. (6.27). In view of eq. (6.48), taking the limit in this formulation should provide us with a mild formulation of the stochastic version of eq. (6.49), but to our knowledge this has not been carried out. Finally, we also have that for each $n \geq 1$ and $x \in \mathbb{R}$, the collection $(\mathscr{Z}_\beta^{(1)}(t, x, \cdot), \ldots, \mathscr{Z}_\beta^{(n)}(t, x, \cdot) : t \geq 0)$ is a $C(\mathbb{R})^n$-valued Markov process [84].

6.5 Multilayer semi-discrete directed polymers

We now describe the multilayer version of the O'Connell–Yor directed polymer introduced by Corwin and Hammond in [25]. As data we are given a countable collection $(B^{(i)} : i \in \mathbb{N})$ of independent standard Brownian motions defined on a common probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. Given integers $1 \le n \le N$, and a fixed time $T > 0$ consider the set

$$\Delta_N^{(n)}(T) := \{(\boldsymbol{t}^1, \ldots, \boldsymbol{t}^n) \in \Delta_{N-n}(T)^n : t_j^{i+1} < t_j^i, i = 1, \ldots, n\} \subset [0,T]^{n(N-n)}.$$

For a fixed set of times $(\boldsymbol{t}^1, \ldots, \boldsymbol{t}^n) \in \Delta_N^{(n)}(T)$ and a realization ω of the Brownian motions, define the Hamiltonian

$$H_N^{(n)}(\boldsymbol{t}^1, \ldots, \boldsymbol{t}^n) := \sum_{i=1}^{n} \sum_{j=0}^{N-n-1} B^{(i+j)}(t_j^i, t_{j+1}^i)$$

and the "quenched" Gibbs measure

$$\mathbf{P}_\beta^\omega(\mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n) := \frac{1}{Z_\beta^{(n)}(T,N)} \mathrm{e}^{\beta H_N^{(n)}(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n)}\, \mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n$$

where the normalizing constant

$$Z_\beta^{(n)}(T,N) := \int_{\Delta_N^{(n)}(T)} \mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n\, \mathrm{e}^{\beta H_N^{(n)}(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n)} \tag{6.50}$$

is the polymer partition function. We also consider the "annealed" polymer measure $Q_\beta^{(n;T,N)}(\mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n \mathrm{d}\omega) := \mathbf{P}_\beta^\omega(\mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n)\mathbb{P}(\mathrm{d}\omega)$.

It is possible to identify the set $\Delta_N^{(n)}(T)$ with the event times of a collection of non-intersecting Poisson paths $\boldsymbol{X} = (X^1, \ldots, X^n)$, starting from the point $\boldsymbol{x}^* := (1, \ldots, n)$ and conditioned to the event $\{\boldsymbol{X}_T = \boldsymbol{y}^*\}$ where $\boldsymbol{y}^* := (N-n+1, \ldots, N)$. We use this interpretation to compute the measure of the set $\Delta_N^{(n)}(T)$.

Lemma 6.5.1. *Let $(\boldsymbol{T}^1, \ldots, \boldsymbol{T}^n)$ be the event times of an ensemble of n non-intersecting Poisson paths starting from $(1,2,\ldots,n)$. Then, conditionally on the event $\{\boldsymbol{X}_T = (N-n+1, \ldots, N)\}$, this vector is uniformly distributed on $\Delta_N^{(n)}(T)$.*

Proof. We first note that if $\boldsymbol{x} \in \mathbb{R}^n$ is any vector such that $x_1 < \cdots < x_n$ then the matrix

$$P(s, \boldsymbol{x}) = \left(\frac{e^{-s} s^{(x_j - x_i)}}{(x_j - x_i)!} \mathbf{1}_{\{x_i \le x_j\}} \right)_{i,j=1}^{n}$$

is upper triangular with diagonal entries equal to e^{-s}. Then, the Karlin–McGregor formula implies that the probability

$$\mathbf{P}_{\boldsymbol{x}}(\boldsymbol{X}_s = \boldsymbol{x}) = \det P(s, \boldsymbol{x}) = e^{-ns}$$

which is independent of x. Next, fix an index $k \in \{1, \dots, n\}$ and pick a vector $\boldsymbol{x} \in \mathbb{R}^n$ such that $x_1 < \cdots < x_k$ and $x_k + 1 < x_{k+1} < \cdots < x_n$. Then, the matrix

$$Q_k(s, \boldsymbol{x}) = \left(\frac{e^{-s} s^{(x_j + \delta_{kj} - x_i)}}{(x_j + \delta_{kj} - x_i)!} \mathbf{1}_{\{x_i \le x_j + \delta_{kj}\}} \right)_{i,j=1}^{n}$$

is again upper triangular with $n-1$ diagonal entries equal to e^{-s} and the k-th diagonal entry equal to se^{-s}. Again, the Karlin–McGregor formula implies that the probability

$$\mathbf{P}_{\boldsymbol{x}}(\boldsymbol{X}_s = \boldsymbol{x} + f_k) = \det Q_k(s, \boldsymbol{x}) = se^{-ns}$$

which is again independent of $\boldsymbol{x}$ and k. Here, f_k is the k-th canonical basis vector in $\mathbb{R}^n$.

Fix a vector $(\boldsymbol{t}^1, \dots, \boldsymbol{t}^n) \in \Delta_N^{(n)}(T)$ and real numbers h_j^i such that all the intervals $I_j^i \equiv (t_j^i, t_j^i + h_j^i]$ are disjoint. The probability

$$\mathbf{P}_{\boldsymbol{x}^*}\Big(T_j^i \in I_j^i \text{ for all } i, j \mid \boldsymbol{X}_T = \boldsymbol{y}^*\Big)$$

can be rewritten as the probability that there is one event (belonging to some of the n paths) in each interval I_j^i and no events elsewhere, divided by $\boldsymbol{P}_{\boldsymbol{x}^*}(\boldsymbol{X}_T = \boldsymbol{y}^*)$. By the preceding paragraph, we can disregard to which path each of these events belong because the probability does not depend on the index. Thus, as this last probability equals

$$e^{-nT} T^{n(N-n)} \prod_{j=0}^{n-1} \frac{j!}{(N-1-j)!}$$

the former equals

$$\left(\prod_{i=1}^{n} \prod_{j=1}^{N-n} h_j^i \right) T^{-n(N-n)} \prod_{j=0}^{n-1} \frac{(N-1-j)!}{j!}.$$

Hence, dividing by the product of the h_j^i variables and taking the limit as they all go to 0, we obtain that the density of the vector $(\boldsymbol{T}^1, \dots, \boldsymbol{T}^n)$ is constant and equal to

$$T^{-n(N-n)} \prod_{j=0}^{n-1} \frac{(N-1-j)!}{j!}.$$

□

Corollary 6.5.2. *The Lebesgue measure of the set $\Delta_N^{(n)}(T)$ is equal to*

$$T^{n(N-n)}\prod_{j=0}^{n-1}\frac{j!}{(N-1-j)!}.$$

6.5.1 An heuristic argument

We now describe a heuristic argument in order to guess what the right scaling for the partition function (6.50) should be. The argument is based in the one in the proof of theorem 6.1.2, which is itself based on the argument by Alberts, Khanin and Quastel described in Section 6.1.

Recall the definition of the n-layer point-to-point semi-discrete directed polymer partition function at inverse temperature $\beta > 0$

$$Z_\beta^{(n)}(T,N) = \int_{\Delta_N^{(n)}(T)} \mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n \;\; \mathrm{e}^{\beta H(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n)}$$

where the Hamiltonian

$$H(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n) := \sum_{i=1}^{n} H_i(\boldsymbol{t}^i)$$

and we have defined

$$H_i(\boldsymbol{t}) := \sum_{j=0}^{N-n-1} B^{(i+j)}(t_j, t_{j+1})$$

for $\boldsymbol{t} \in \Delta_{N-n}(T)$.

Making the approximation $\mathrm{e}^x \approx 1 + x$ we can see that $Z_\beta^{(n)}$ behaves, up to first order, like

$$Z_\beta^{(n)}(T,N) \approx |\Delta_N^{(n)}(T)| + \beta \int_{\Delta_N^{(n)}(T)} \mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n \;\; \sum_{i=1}^{n}\sum_{j=0}^{N-n-1} B^{(i+j)}(t_j^i, t_{j+1}^i)$$

thus

$$c_{T,N}^{(n)} Z_\beta^{(n)}(T,N) \approx 1 + \beta\, c_{T,N}^{(n)} \int_{\Delta_N^{(n)}(T)} \mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n \;\; \sum_{i=1}^{n}\sum_{j=0}^{N-n-1} B^{(i+j)}(t_j^i, t_{j+1}^i)$$

where the value of the normalising constant $c_{T,N}^{(n)} = |\Delta_N^{(n)}(T)|^{-1}$ has been already computed in Corollary 6.5.2. According to Lemma 6.5.1 the integral on the right is exactly the conditional expectation over a Poisson bridge conditioned ending at time T at positions $\boldsymbol{y}^*$, so that

$$c_{T,N}^{(n)} Z_\beta^{(n)}(T,N) \approx 1 + \beta \sum_{i=1}^{n}\sum_{j=0}^{N-n-1} \mathbf{E}_{\boldsymbol{x}^*}\left[B^{(i+j)}(T_j^i, T_{j+1}^i) \,\middle|\, \boldsymbol{X}_T = \boldsymbol{y}^* \right]. \tag{6.51}$$

Consider the shear transformation $\psi_N : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+ \to \mathbb{R}$ given by $\psi_N(t,x) = \left(\frac{t}{N}, \frac{x-t}{\sqrt{N}}\right)$ and, for $j \in \mathbb{N}$ and $0 < s < t$ let $I_N^{(j)}(s,t)$ denote the image of the rectangle $[j-1, j] \times [s,t]$ under ψ_N. For $u \in [s/N, t/N]$ the intervals $I_N^{(j)}(u) := \left[\frac{j-1}{\sqrt{N}} - \sqrt{N}u, \frac{j}{\sqrt{N}} - \sqrt{N}u\right]$ are such that

$$I_N^{(j)}(s,t) = \bigcup_{u\in[s/N,t/N]} I_N^{(j)}(u)$$

and since an easy computation gives $|I_N^{(j)}(s,t)| = N^{-3/2}(t-s)$ we have that the processes

$$W_N^{(j)}(t) = N^{3/4} \int_0^{t/N} \langle \mathbf{1}_{I_N^{(j)}(u)}, \mathrm{d}\xi_u \rangle$$

form a countable collection of independent standard Brownian motions. Thus, replacing these processes into eq. (6.51) we obtain that, in distribution, the rescaled partition function behaves approximately as

$$c_{T,N}^{(n)} Z_\beta^{(n)}(T,N) \approx 1 + \beta N^{3/4} \sum_{i=1}^{n} \sum_{j=0}^{N-n-1} \int_0^{T/N} \mathbf{E}_{\boldsymbol{x}^*}\left[\mathbf{1}_{\{X_s^i = i+j\}} \,\middle|\, \boldsymbol{X}_T = \boldsymbol{y}^*\right] \mathbf{1}_{I_N^{(j)}(s)}(x)\, \xi(\mathrm{d}s, \mathrm{d}x).$$

Recalling the correlation function from eq. (6.25) we see that

$$c_{T,N}^{(n)} Z_\beta^{(n)}(T,N) \approx 1 + \beta N^{3/4} \int_0^{T/N} \sum_{i=1}^{n} \sum_{j=0}^{N-n-1} \rho_i^{(n)}(s, i+j; T, \boldsymbol{x}^*, \boldsymbol{y}^*) \mathbf{1}_{I_N^{(i+j)}(s)}(x)\, \xi(\mathrm{d}s, \mathrm{d}x).$$

We also recall the Sawtooth walk and its rescaled correlation function defined in Section 6.3.3. Since by definition this rescaled version equals its unscaled counterpart in each of the $I_N^{(\ell)}(u)$ we have, as the union of the intervals $I_N^{(i+j)}(s)$ over $j = 0, \dots, N-n-1$ equals the interval $J_N^{(i)}(s) := \left[\frac{i-1}{\sqrt{N}} - \sqrt{N}s, \frac{i+N-n}{\sqrt{N}} - \sqrt{N}s\right]$, that

$$c_{T,N}^{(n)} Z_\beta^{(n)}(T,N) \approx 1 + \beta N^{1/4} \int_0^{T/N} \int_{\mathbb{R}} \sum_{i=1}^{n} \rho_{i,N}^{(n)}(s, x; T/N, \boldsymbol{x}_N, \boldsymbol{y}_N) \mathbf{1}_{J_N^{(i)}(s)}(x)\, \xi(\mathrm{d}s, \mathrm{d}x)$$

But the probabilities defining the correlation kernel $\rho_{i,N}^{(n)}$ vanishes if $x \notin J_N^{(i)}(s)$ hence the above approximation becomes

$$c_{T,N}^{(n)} Z_{\beta N^{-1/4}}^{(n)}(T,N) \approx 1 + \beta \int_0^{T/N} \int_{\mathbb{R}} \rho_{1,N}^{(n)}(s, x; T/N, \boldsymbol{x}_N, \boldsymbol{y}_N)\, \xi(\mathrm{d}s, \mathrm{d}x).$$

Therefore, by the convergence in Theorem 6.3.14, we see that at least to first order the partition properly rescaled partition function converges to the series defining the multilayer SHE process, thus obtaining Theorem 6.1.3.

6.6 Wick products

In order to turn this argument into an actual proof we have to obtain a description of the multilayer semi-discrete partition function as a series in β and prove that each of the coefficients of this series, when properly rescaled, converges appropriately to the stochastic integral appearing in eq. (6.45).

To achieve this, following Moreno-Flores, Quastel and Remenik we introduce the following modified version of the partition function,

$$\mathfrak{Z}_\beta^{(n)}(T,N) = c_{T,N}^{(n)} \int_{\Delta_N^{(n)}(T)} \mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n \prod_{i=1}^{n} \prod_{j=0}^{N-n-1} (1 + \beta B^{(i+j)}(t_j^i, t_{j+1}^i)) \tag{6.52}$$

which is a polynomial in β of degree $n(N-n)$, such that the coefficient of the linear part is precisely the one appearing in eq. (6.51). It is possible to show that the higher order coefficients admit similar expressions as ordered sums of nested products of the Brownian motions, and which can be related to the rescaled k-point correlation function of the sawtooth process just as we have done in the previous section. Therefore, at this stage Theorem 6.3.14 still applies and one can show that as $N \to \infty$ the coefficient accompanying β^k in the expansion of the double product in eq. (6.52) converges to the corresponding iterated stochastic integral in eq. (6.45). We note however that this process is complicated as it involves expanding a double product of monomials with intertwined coefficients.

The next step is to relate the convergence of an appropriate rescaling of $\mathfrak{Z}_\beta^{(n)}$ to the convergence of the actual partition function $Z_\beta^{(n)}$. In [93] this is done by considering an auxiliary collection of processes

$$U_N^{(i)}(s,t) := \frac{N^{1/4}}{\beta} \left[\mathrm{e}^{\beta N^{-1/4} W_N^{(i)}(s,t) - \frac{\beta^2}{2} N^{-1/2}(t-s)} - 1 \right] \tag{6.53}$$

so that

$$c_{N,N}^{(n)} \mathrm{e}^{-\frac{n\beta^2}{2} N^{1/2}} Z_{\beta N^{-1/4}}(N,N) = c_{N,N}^{(n)} \int_{\Delta_N^{(n)}(N)} \mathrm{d}\boldsymbol{t}^1 \cdots \mathrm{d}\boldsymbol{t}^n \prod_{i=1}^{n} \prod_{j=0}^{N-n-1} \left(1 + \beta N^{-1/4} U_N^{(i+j)}(t_j^i, t_{j+1}^i)\right). \tag{6.54}$$

This expression has the same form of eq. (6.52) so that the same techniques apply and one can obtain a similar convergence result over the coefficients of the polynomial appearing as the expansion of this new double product. At this stage some additional care has to be taken since both the coefficients of the polynomial and the degree of the polynomial itself are changing with N. Technically, what has to be done is that one should show convergence of the individual coefficients plus some uniform estimates in N for this to work.

But, if we observe the definition of $U^{(\ell)}(s,t)$ we can see that actually the term appearing above is, save for the indices, equal to

$$1+\beta N^{-1/4}U^{(\ell)}(s,t)=\mathrm{e}^{\beta N^{-1/4}W_N^{(\ell)}(s,t)-\frac{\beta^2}{2}N^{-1/2}(t-s)}.$$

The random variable $\beta N^{-1/4}W_N^{(\ell)}(s,t)$ is centred Gaussian with variance $\beta^2 N^{-1/2}(t-s)$ so that the exponential above is actually the *Wick exponential* $:\exp:\left(\beta N^{-1/4}W_N^{(\ell)}(s,t)\right)$. Moreover, these variables are centred independent for different values of ℓ and disjoint time intervals (s,t). Hence, the product appearing in eq. (6.54) is the Wick product

$$:\exp:\left(\beta N^{-1/4}\sum_{i=1}^{n}\sum_{j=0}^{N-n-1}W_N^{(i+j)}(t_j^i,t_{j+1}^i)\right)=\,:\exp:\left(\beta N^{-1/4}H_N^{(n)}(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n)\right)$$

where $H_N^{(n)}$ is the Hamiltonian in eq. (6.50). Again by independence, $H_N^{(n)}$ can be seen to have a centred Gaussian distribution with variance $n\beta^2\sqrt{N}$ for each $\Delta_N^{(n)}(N)$. Therefore,

$$:\exp:\left(\beta N^{-1/4}H_N^{(n)}(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n)\right)=\mathrm{e}^{\beta N^{-1/4}H_N^{(n)}-n\frac{\beta^2\sqrt{N}}{2}}$$

and this also explains the exponential factor in eq. (6.54). As a consequence, we get that

$$\mathrm{e}^{-n\frac{\beta^2\sqrt{N}}{2}}Z_{\beta N^{-1/4}}^{(n)}(N,N)=\int_{\Delta_N^{(n)}(N)}\mathrm{d}\boldsymbol{t}^1\cdots\mathrm{d}\boldsymbol{t}^n\;\;:\exp:\left(\beta N^{-1/4}H_N^{(n)}(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n)\right).$$

On the other hand, the increments $W_N^{(\ell)}(s,t)$ may be expressed as

$$W_N^{(\ell)}(s,t)=\int_s^t\mathrm{d}W_N^{(\ell)}(u)=\int_0^N\mathbf{1}_{[s,t)}(u)\,\mathrm{d}W_N^{(\ell)}(u)$$

thus

$$\begin{aligned}H_i(\boldsymbol{t})&=\sum_{j=0}^{N-n-1}\int_0^N\mathbf{1}_{[t_j,t_{j+1})}(u)\,\mathrm{d}W_N^{(i+j)}(u)\\&=\int_0^N\mathrm{d}W_N^{(X_u^i)}(u)\end{aligned}$$

where $\boldsymbol{X}=(X^1,\ldots,X^n)$ is a non-intersecting Poisson bridge starting at $(1,\ldots,n)$ at time 0 and ending at $(N-n+1,\ldots,N)$ at time N. Therefore, the Wick exponential of $H_N^{(n)}$ is given by the chaos series

$$\sum_{k=0}^{\infty} \beta^k N^{-\frac{k}{4}} \int_{\Delta_k(N)} \sum_{\boldsymbol{i}\in[\![n]\!]^k} \sum_{\boldsymbol{j}\in[\![N-n]\!]^k} \mathbf{E}_{\boldsymbol{x}^*}\left[\prod_{\ell=0}^{k} \mathbf{1}_{\{X^{i_\ell}_{s_\ell}=j_\ell\}}\right] \mathrm{d}W_N^{(j_1)}(s_1)\cdots \mathrm{d}W_N^{(j_k)}(s_k).$$

Define

$$J_k(N) := \int_{\Delta_k(N)} \sum_{\boldsymbol{i}\in[\![n]\!]^k} \sum_{\boldsymbol{j}\in[\![N-n]\!]^k} \mathbf{E}_{\boldsymbol{x}^*}\left[\prod_{\ell=1}^{k} \mathbf{1}_{\{X^{i_k}_{s_k}=j_k\}} \,\middle|\, \boldsymbol{X}_N = \boldsymbol{y}^*\right] \mathrm{d}W_N^{(j_1)}(s_1)\cdots \mathrm{d}W_N^{(j_k)}(s_k)$$

so that

$$:\exp\colon (H_N^{(n)}) = \mathbf{P}_{\boldsymbol{x}^*}[\boldsymbol{X}_N = \boldsymbol{y}^*] \sum_{k=0}^{\infty} \beta^k N^{-\frac{k}{4}} J_k(N).$$

Theorem 6.6.1. *For each $k \geq 1$ the convergence*

$$N^{-k/4} J_k(N) \to \int_{\Delta_k(1)} \int_{\mathbb{R}^k} R_k^{(n)}(\boldsymbol{s},\boldsymbol{x};1,0,0)\, \xi^{\otimes k}(\mathrm{d}\boldsymbol{s},\mathrm{d}\boldsymbol{x})$$

holds.

Proof. Using the definition of $W_N^{(\ell)}(s,t)$ we get that

$$N^{-k/4} J_k(N) = \int_{\Delta_k(1)} \int_{\mathbb{R}^k} \varrho_{k,N}^{(n)}(\boldsymbol{s},\boldsymbol{x};N,\boldsymbol{a}_N,\boldsymbol{b}_N)\, \xi^{\otimes k}(\mathrm{d}\boldsymbol{s},\mathrm{d}\boldsymbol{x})$$

whence the result follows after an application of Theorem 6.3.15 and an L^2 estimate found in [95]. ◻

Bibliography

1. ABOUD, Ammar; BULTEL, Jean-Paul; CHOURIA, Ali; LUQUE, Jean-Gabriel; MALLET, Olivier. *Bell polynomials in combinatorial Hopf algebras*. 2014. Available from arXiv: 1402.2960 [math.CO].
2. ALBERTS, Tom; KHANIN, Konstantin; QUASTEL, Jeremy. The Continuum Directed Random Polymer. *J. Stat. Phys.* 2014, vol. 154, no. 1-2, pp. 305–326. ISSN 00224715. Available from DOI: 10.1007/s10955-013-0872-z.
3. ALBERTS, Tom; KHANIN, Konstantin; QUASTEL, Jeremy. The intermediate disorder regime for directed polymers in dimension 1+1. *Ann. Probab.* 2014, vol. 42, no. 3, pp. 1212–1256. ISSN 00911798. Available from DOI: 10.1214/13-AOP858.
4. AMIR, Gideon; CORWIN, Ivan; QUASTEL, Jeremy. Probability distribution of the free energy of the continuum directed random polymer in 1+ 1 dimensions. *Commun. Pure Appl. Math.* 2011, vol. 64, no. 4, pp. 466–537. Available from DOI: 10.1002/cpa.20347.
5. AZIZNEJAD, Shayan; FAGEOT, Julien; UNSER, Michael. *Wavelet Analysis of the Besov Regularity of Lévy White Noises*. 2018. Available from arXiv: 1801.09245 [math.PR].
6. BERTINI, Lorenzo; GIACOMIN, Giambattista. Stochastic Burgers and KPZ Equations from Particle Systems. *Commun. Math. Phys.* 1997, vol. 183, no. 3, pp. 571–607. ISSN 0010-3616. Available from DOI: 10.1007/s002200050044.
7. BOEDIHARDJO, Horatio; CHEVYREV, Ilya. *An isomorphism between branched and geometric rough paths*. 2017. Available from arXiv: 1712.01965 [math.PR].
8. BORODIN, Alexei; CORWIN, Ivan; REMENIK, Daniel. Log-gamma polymer free energy fluctuations via a Fredholm determinant identity. *Commun. Math. Phys.* 2013, vol. 324, no. 1, pp. 215–232. Available from DOI: 10.1007/s00220-013-1750-x.
9. BOURBAKI, Nicolas. *Lie groups and Lie algebras, part I (chapters 1-3)*. Addison-Wesley, Hermann, 1975. Elements of mathematics. ISBN 9780201006438.
10. BOURBAKI, Nicolas. *Algèbre: Chapitres 1 à 3*. 2nd ed. Springer-Verlag Berlin Heidelberg, 2007. ISBN 978-3-540-33850-5. Original edition published by Masson, Paris, 1970.
11. BROUDER, C.; FAUSER, B.; FRABETTI, A.; OECKL, R. Quantum field theory and Hopf algebra cohomology. *J. Phys. A*. 2004, vol. 37, no. 22, pp. 5895–5927. ISSN 0305-4470. Available from DOI: 10.1088/0305-4470/37/22/014.
12. BROUDER, C.; PATRAS, F. One-particle irreducibility with initial correlations. In: EBRAHIMI-FARD, Kurusch; MARCOLLI, Matilde; SUIJLEKOM, Walter D. van (eds.). *Combinatorics and Physics*. American Mathematical Society, 2011, pp. 1–25. Contemporary Mathematics, no. 539.
13. BRUNED, Yvain; CHEVYREV, Ilya; FRIZ, Peter K.; PREISS, Rosa. *A Rough Path Perspective on Renormalization*. 2017. Available from arXiv: 1701.01152 [math.PR].
14. BRUNED, Yvain; CURRY, Charles; EBRAHIMI-FARD, Kurusch. *Quasi-shuffle algebras and renormalisation of rough differential equations*. 2018. Available from arXiv: 1801.02964 [math.PR].
15. BRUNED, Yvain; HAIRER, Martin; ZAMBOTTI, Lorenzo. *Algebraic renormalisation of regularity structures*. 2016. Available from arXiv: 1610.08468 [math.RA].
16. BUTCHER, John C. *Numerical Methods for Ordinary Differential Equations*. 3rd ed. John Wiley & Sons, Ltd, 2016. ISBN 9781119121534. Available from DOI: 10.1002/9781119121534.
17. CALAQUE, Damien; EBRAHIMI-FARD, Kurusch; MANCHON, Dominique. Two interacting Hopf algebras of trees: A Hopf-algebraic approach to composition and substitution of B-series. *Adv. Appl. Math.* 2011, vol. 47, no. 2, pp. 282–308. ISSN 01968858. Available from DOI: 10.1016/j.aam.2009.08.003.

18. CARTIER, Pierre. A Primer of Hopf Algebras. In: *Frontiers in Number Theory, Physics, and Geometry II: On Conformal Field Theories, Discrete Groups and Renormalization*. Ed. by CARTIER, Pierre; MOUSSA, Pierre; JULIA, Bernard; VANHOVE, Pierre. Springer-Verlag Berlin Heidelberg, 2007, pp. 537–615. ISBN 978-3-540-30308-4. Available from DOI: `10.1007/978-3-540-30308-4_12`.

19. ÇETIN, Umut; DANILOVA, Albina. Markov bridges: SDE representation. *Stoch. Process. Appl.* 2016, vol. 126, no. 3, pp. 651–679. ISSN 03044149. Available from DOI: `10.1016/j.spa.2015.09.015`.

20. CHANDRA, Ajay; HAIRER, Martin. *An analytic BPHZ theorem for regularity structures*. 2017. Available from arXiv: `1612.08138 [math.PR]`.

21. CHAPOTON, Frédéric. Free pre-Lie algebras are free as Lie algebras. *Canad. Math. Bull.* 2010, vol. 53, no. 3, pp. 425–437. ISSN 0008-4395. Available from DOI: `10.4153/CMB-2010-063-2`.

22. CHAPOTON, Frédéric; LIVERNET, Muriel. Pre-Lie algebras and the rooted trees operad. *Int. Math. Res. Notices*. 2001, vol. 2001, no. 8, pp. 395–408. Available from DOI: `10.1155/S1073792801000198`.

23. CHEN, Kuo-Tsai. Iterated path integrals. *Bull. Amer. Math. Soc.* 1977, vol. 83, no. 5, pp. 831–880. ISSN 0002-9904. Available from DOI: `10.1090/S0002-9904-1977-14320-6`.

24. CONNES, Alain; KREIMER, Dirk. Hopf Algebras, Renormalization and Noncommutative Geometry. *Commun. Math. Phys.* 1998, vol. 199, no. 1, pp. 203–242. ISSN 0010-3616. Available from DOI: `10.1007/s002200050499`.

25. CORWIN, Ivan; HAMMOND, Alan. KPZ line ensemble. *Probab. Theory Relat. Fields*. 2016, vol. 166, no. 1-2, pp. 67–185. ISSN 01788051. Available from DOI: `10.1007/s00440-015-0651-7`.

26. CORWIN, Ivan; NICA, Mihai. Intermediate disorder directed polymers and the multi-layer extension of the stochastic heat equation. *Electron. J. Probab.* 2017, vol. 22. ISSN 10836489. Available from DOI: `10.1214/17-EJP32`.

27. COUTIN, Laure; QIAN, Zhongmin. Stochastic analysis, rough path analysis and fractional Brownian motions. *Probab. theory Relat. fields*. 2002, vol. 122, no. 1, pp. 108–140. Available from DOI: `10.1007/s004400100158`.

28. CURRY, Charles; EBRAHIMI-FARD, Kurusch; MANCHON, Dominique; MUNTHE-KAAS, Hans Z. Planarly branched rough paths and rough differential equations on homogeneous spaces. 2018. Available from eprint: `1804.08515`.

29. DA PRATO, Giuseppe; ZABCZYK, Jerzy. *Stochastic Equations in Infinite Dimensions*. 2nd ed. Cambridge University Press, 2008. Encyclopedia of Mathematics and its Applications, no. 45.

30. DĂSCĂLESCU, Sorin; NĂSTĂSESCU, Constantin; RAIANU, Şerban. *Hopf algebras: An Introduction*. New York, NY: Marcel Dekker, Inc., 2001. Monographs and Textbooks in Pure and Applied Mathematics, no. 235. ISBN 978-0-8247-0481-0.

31. DAUBECHIES, Ingrid. Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.* 1988, vol. 41, no. 7, pp. 909–996. Available from DOI: `10.1002/cpa.3160410705`.

32. DAUBECHIES, Ingrid. *Ten Lectures on Wavelets*. SIAM, 1992. CBMS-NSF Regional Conference Series in Applied Mathematics, no. 61. ISBN 978-1-61197-010-4. Available from DOI: `10.1137/1.9781611970104`.

33. DAUBECHIES, Ingrid. Orthonormal bases of compactly supported wavelets II. Variations on a theme. *SIAM J. Math. Analysis*. 1993, vol. 24, no. 2, pp. 499–519. Available from DOI: `10.1137/0524031`.

34. DAVIE, A. M. Differential Equations Driven by Rough Paths: An Approach via Discrete Approximation. *Appl. Math. Res. eXPress*. 2008, vol. 2008, no. amb09. Available from DOI: `10.1093/amrx/abm009`.

35. DYNKIN, E. B. Calculation of the coefficients in the Campbell–Hausdorff formula. In: DYNKIN, E. B. *Selected Papers of E. B. Dynkin with Commentary*. Ed. by YUSHKEVICH, A. A.; SEITZ, G. M.; ONISHCHIK, A. L. American Mathematical Society and International Press, 2000, pp. 31–35. Collected Works. ISBN 978-0-8218-1065-1.

36. EBRAHIMI-FARD, Kurusch; PATRAS, Frederic. Cumulants, free cumulants and half-shuffles. *Proc. Royal Soc. Lond. A*. 2015, vol. 471, no. 2176. ISSN 1471-2946. Available from DOI: `10.1098/rspa.2014.0843`.

37. EBRAHIMI-FARD, Kurusch; PATRAS, Frederic. *Monotone, free, and boolean cumulants: a shuffle algebra approach*. 2017. Available from arXiv: `1701.06152 [math.CO]`.

38. EBRAHIMI-FARD, Kurusch; PATRAS, Frédéric. The Splitting Process in Free Probability Theory. *Int. Math. Res. Notices*. 2016, no. 9, pp. 2647–2676. ISSN 1073-7928. Available from DOI: `10.1093/imrn/rnv209`.

39. EBRAHIMI-FARD, Kurusch; PATRAS, Frédéric; TAPIA, Nikolas; ZAMBOTTI, Lorenzo. *Hopf-algebraic deformations of products and Wick polynomials*. 2017. Available from arXiv: `1710.00735 [math.PR]`.

40. FERRARI, Patrik L. *Dimers and orthogonal polynomials: connections with random matrices: Extended lecture notes of the minicourse at IHP*. 2013. Available from arXiv: `1004.3212 [math-ph]`.

41. FITZSIMMONS, Pat; PITMAN, Jim; YOR, Marc. Markovian Bridges: Construction, Palm Interpretation, and Splicing. In: *Seminar on Stochastic Processes, 1992*. Boston, MA: Birkhäuser Boston, 1993, pp. 101–134. Available from DOI: `10.1007/978-1-4612-0339-1_5`.

42. FOISSY, L. Commutative and non-commutative bialgebras of quasi-posets and applications to Ehrhart polynomials. *ArXiv e-prints*. 2016. Available from arXiv: `1605.08310 [math.RA]`.

43. FOISSY, Loıc. Finite dimensional comodules over the Hopf algebra of rooted trees. *J. Algebra*. 2002, vol. 255, no. 1, pp. 89–120. Available from DOI: `10.1016/S0021-8693(02)00110-2`.

44. FOISSY, Loıc. Les algèbres de Hopf des arbres enracinés décorés, II. *Bull. Sci. Math.* 2002, vol. 126, no. 4, pp. 249–288. Available from DOI: `10.1016/S0007-4497(02)01113-2`.

45. FOLLAND, Gerald B.; STEIN, Elias M. *Hardy Spaces on Homogeneous Groups*. 1st ed. Princeton University Press, 1982. Princeton mathematical notes, no. 28. ISBN 9780691083100.

46. FRIZ, Peter K.; VICTOIR, Nicolas. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. Cambridge: Cambridge University Press, 2010. Cambridge Studies in Advanced Mathematics, no. 120. ISBN 9780521876070.

47. GAUGER, Michael A. On the classification of Metabelian Lie Algebras. *Trans. Amer. Math. Soc.* 1973, vol. 179, pp. 293–329.

48. GROSSMAN, Robert L.; LARSON, Richard G. Hopf-algebraic structures of families of trees. *J. Algebra*. 1989, vol. 126, no. 1, pp. 184–210. Available from arXiv: `0711.3877 [math.RA]`.

49. GROSSMAN, Robert L.; LARSON, Richard G. Hopf-algebraic structure of combinatorial objects and differential operators. *Israel J. Math.* 1990, vol. 72, no. 1-2, pp. 109–117. ISSN 0021-2172. Available from DOI: `10.1007/BF02764614`.

50. GUBINELLI, M. Controlling rough paths. *J. Funct. Analysis*. 2004, vol. 216, no. 1, pp. 86–140. ISSN 0022-1236. Available from DOI: `10.1016/j.jfa.2004.01.002`.

51. GUBINELLI, Massimiliano. Ramification of rough paths. *J. Differ. Equ.* 2010, vol. 248, no. 4, pp. 693–721. ISSN 00220396. Available from DOI: `10.1016/j.jde.2009.11.015`.

52. GYURKÓ, Lajos Gegerly. Differential Equations Driven by Π-Rough Paths. *Proc. Edinb. Math. Soc.* 2016, vol. 59, no. 3, pp. 741–758. Available from DOI: `10.1017/S0013091515000474`.

53. HAAR, Alfréd. Zur Theorie der orthogonalen Funktionsysteme. *Math. Ann.* 1910, vol. 69, no. 3, pp. 331–371. Available from DOI: `10.1007/BF01456326`.

54. HAIRER, Ernst; LUBICH, Christian; WANNER, Gerhard. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. 2nd ed. Springer-Verlag Berlin Heidelberg, 2006. Springer Series in Computational Mathematics, no. 31. ISBN 978-3-540-30666-5. ISSN 0179-3632. Available from DOI: `10.1007/3-540-30666-8`.

55. HAIRER, Martin. Solving the KPZ equation. *Ann. Math.* 2013, vol. 178, no. 2, pp. 559–664. ISSN 0003-486X. Available from DOI: `10.4007/annals.2013.178.2.4`.

56. HAIRER, Martin. A theory of regularity structures. *Invent. Math.* 2014, vol. 198, no. 2, pp. 269–504. ISSN 0020-9910. Available from DOI: `10.1007/s00222-014-0505-4`.

57. HAIRER, Martin; KELLY, David. Geometric versus non-geometric rough paths. *Ann. de l'Institut Henri Poincaré - Probab.és et Stat.* 2014, vol. 198, no. 2, pp. 269–504. Available from arXiv: `1210.6294 [math.PR]`.

58. HAIRER, Martin; LABBÉ, Cyril. Multiplicative stochastic heat equations on the whole space. *J. European Math. Soc.* 2016. Available from arXiv: `1504.07162 [math.AP]`.

59. HAIRER, Martin; SHEN, Hao. A central limit theorem for the KPZ equation. *Ann. Probab.* 2017, vol. 45, no. 6B, pp. 4167–4221. Available from DOI: `10.1214/16-AOP1162`.

60. HALL, Brian C. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. 2nd ed. Springer, Cham, 2015. Graduate Texts in Mathematics, no. 222. ISBN 978-3-319-13467-3. ISSN 2197-5612. Available from DOI: 10.1007/978-3-319-13467-3.

61. HÄRDLE, Wolfgang; KERKYACHARIAN, Gerard; PICARD, Dominique; TSYBAKOV, Alexander. *Wavelets, Approximation, and Statistical Applications*. New York, NY: Springer, 1998. Lecture Notes in Statistics, no. 129. ISBN 978-1-4612-2222-4. Available from DOI: 10.1007/978-1-4612-2222-4_9.

62. HARISH-CHANDRA. Spherical Functions on a Semisimple Lie Group, I. *Amer. J. Math.* 1958, vol. 80, no. 2, pp. 241–310. Available from DOI: 10.2307/2372786.

63. HOFFMAN, Michael E. Combinatorics of Rooted Trees and Hopf Algebras. *Trans. Amer. Math. Soc.* 2002, vol. 355, pp. 3795–3811. ISSN 0002-9947. Available from DOI: 10.1090/S0002-9947-03-03317-8.

64. ITZYKSON, C.; ZUBER, J.-B. The planar approximation. II. *J. Math. Phys.* 1980, vol. 21, no. 3, pp. 411–421. Available from DOI: 10.1063/1.524438.

65. IZUKI, Mitsuo; SAWANO, Yoshihiro. Wavelet basis in the weighted Besov and Triebel–Lizorkin spaces with A_p^{loc}-weights. *J. Approx. Theory*. 2009, vol. 161, no. 2, pp. 656–673. Available from DOI: 10.1016/j.jat.2008.12.003.

66. JACOBSON, Nathan. *Lie Algebras*. 3rd ed. Interscience Publishers, 1966. Interscience Tracts in Pure and Applied Mathematics, no. 10. ISBN 978-0-470-43593-9.

67. JANSON, Svante. *Gaussian Hilbert Spaces*. Cambridge University Press, 2009. Cambridge Tracts in Mathematics, no. 129. Available from DOI: 10.1017/CBO9780511526169.

68. JOHANSSON, Kurt. Shape fluctuations and random matrices. *Commun. Math. Phys.* 2000, vol. 209, no. 2, pp. 437–476.

69. JOHANSSON, Kurt. Discrete Polynuclear Growth and Determinantal Processes. *Commun. Math. Phys.* 2003, vol. 242, no. 1-2, pp. 277–329. ISSN 0010-3616. Available from DOI: 10.1007/s00220-003-0945-y.

70. JOHANSSON, Kurt. Non-intersecting, simple, symmetric random walks and the extended Hahn kernel. *Ann. de l'Institut Fourier*. 2004, vol. 55, no. 6, pp. 2129–2145. ISSN 0373-0956. Available from DOI: 10.5802/aif.2155.

71. JONI, S A; ROTA, G.-C. Coalgebras and Bialgebras in Combinatorics. *Stud. Appl. Math.* 1979, vol. 61, no. 2, pp. 93–139. ISSN 00222526. Available from DOI: 10.1002/sapm197961293.

72. KABANAVA, Maryia. Tempered Radon Measures. *Rev. Mat.ática Comp.* 2008, vol. 21, no. 2, pp. 553–564.

73. KARDAR, Mehran; PARISI, Giorgio; ZHANG, Yi-Cheng. Dynamic Scaling of Growing Interfaces. *Phys. Rev. Lett.* 1986, vol. 56, no. 9, pp. 889–892. Available from DOI: 10.1103/PhysRevLett.56.889.

74. KARLIN, Samuel; MCGREGOR, J. Coincidence probabilities. *Pac. J. Math.* 1959, vol. 9, no. 4, pp. 1141–1164. ISSN 0030-8730. Available from DOI: 10.2140/pjm.1959.9.1141.

75. KÖNIG, Wolfgang; O'CONNELL, Neil; ROCH, Sébastien. Non-Colliding Random Walks, Tandem Queues, and Discrete Orthogonal Polynomial Ensembles. *Electron. J. Probab.* 2002, vol. 7, no. 2000, pp. 1–24. ISSN 1083-6489. Available from DOI: 10.1214/EJP.v7-104.

76. KRATTENTHALER, Christian. Advanced Determinant Calculus. *Séminaire Lotharingien de Comb.* 1999, vol. 42, no. B42q. Available from eprint: math/9902004 (math.CO).

77. LE DONNE, Enrico; RIGOT, Séverine. Besicovitch Covering Property on graded groups and applications to measure differentiation. *J. für die Reine und Angewandte Math. (Crelles J.)* 2016. Available from DOI: https://doi.org/10.1515/crelle-2016-0051.

78. LEONOV, V. P.; SHIRYAEV, A. N. On a method of semi-invariants. *Theory Probab. its Appl.* 1959, vol. 4, no. 3, pp. 319–329. Available from DOI: 10.1137/1104031.

79. LIU, Yanghui; TINDEL, Samy. *First-order Euler scheme for SDEs driven by fractional Brownian motions: the rough case*. 2017. Available from arXiv: 1703.03625 [math.PR].

80. LODAY, Jean-Louis. Série de Hausdorff, idempotents Eulériens et algèbres de Hopf. *Expo. Math.* 1994, vol. 12, no. 2, pp. 165–178.

81. LODAY, Jean-Louis; VALLETTE, Bruno. *Algebraic Operads*. Springer Berlin Heidelberg, 2012. Grundlehren der mathematischen Wissenschaften, no. 346. ISBN 978-3-642-30361-6. ISSN 0072-7830. Available from DOI: 10.1007/978-3-642-30362-3.

82. LUKKARINEN, J.; MARCOZZI, M. Wick polynomials and time-evolution of cumulants. *J. Math. Phys.* 2016, vol. 57, no. 8, pp. 083301, 27. ISSN 0022-2488. Available from DOI: 10.1063/1.4960556.

83. LUN, Chin Hang; WARREN, Jon. *Continuity and strict positivity of the multi-layer extension of the stochastic heat equation*. 2015. Available from arXiv: 1506.09030 [math.PR].

84. LUN, Chin Hang; WARREN, Jon. *The stochastic heat equation, 2D Toda equations and dynamics for the multilayer process*. 2016. Available from arXiv: 1606.05139 [math.PR].

85. LYONS, Terry. On the non-existence of path integrals. *Proc. Royal Soc. Lond. A: Math., Phys. Eng. Sci.* 1991, vol. 432, no. 1885, pp. 281–290. ISSN 0962-8444. Available from DOI: 10.1098/rspa.1991.0017.

86. LYONS, Terry. Differential equations driven by rough signals. *Rev. Mat.ática Iber.*. 1998, vol. 14, pp. 215–310. ISSN 0213-2230. Available from DOI: 10.4171/RMI/240.

87. LYONS, Terry J; CARUANA, Michael J. *Differential Equations Driven by Rough Paths*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. Lecture Notes in Mathematics, no. 1908. ISBN 978-3-540-71284-8. Available from DOI: 10.1007/978-3-540-71285-5.

88. LYONS, Terry; VICTOIR, Nicolas B. An extension theorem to rough paths. *Ann. de l'Institut Henri Poincare (C) Non Linear Analysis*. 2007, vol. 24, no. 5, pp. 835–847. ISSN 0294-1449. Available from DOI: 10.1016/j.anihpc.2006.07.004.

89. MANCHON, Dominique. Hopf Algebras in Renormalisation. In: HAZEWINKEL, M. (ed.). *Handbook of Algebra*. North-Holland, 2008, vol. 5, pp. 365–427. ISSN 1570-7954. Available from DOI: 10.1016/S1570-7954(07)05007-3.

90. MEYER, Yves. *Wavelets and Operators*. Trans. by SALINGER, D. H. Cambridge University Press, 1993. Cambridge Studies in Advanced Mathematics, no. 37. ISBN 0-521-45869-2.

91. MEYER, Yves. Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs. In: *Séminaire Bourbaki*. 1985–1986, vol. 28, pp. 209–223. Talk no. 662.

92. MILNOR, John W.; MOORE, John C. On the Structure of Hopf Algebras. *Ann. Math.* 1965, vol. 81, no. 2, pp. 211–264. ISSN 0003486X. Available from DOI: 10.2307/1970615.

93. MORENO FLORES, Gregorio; QUASTEL, Jeremy; REMENIK, Daniel. *Intermediate disorder limits for directed polymers with boundary conditions*. 2015.

94. MUNASINGHE, Ranjiva; RAJESH, R; TRIBE, Roger; ZABORONSKI, Oleg. Multi-Scaling of the n-Point Density Function for Coalescing Brownian Motions. *Commun. Math. Phys.* 2006, vol. 268, no. 3, pp. 717–725. ISSN 0010-3616. Available from DOI: 10.1007/s00220-006-0110-5.

95. NICA, Mihai. *Intermediate disorder limits for multi-layer semi-discrete directed polymers*. 2016. Available from arXiv: 1609.00298 [math.PR].

96. NUALART, David; TINDEL, Samy. A construction of the rough path above fractional Brownian motion using Volterra's representation. *Ann. Probab.* 2011, vol. 39, no. 3, pp. 1061–1096. Available from DOI: 10.1214/10-AOP578.

97. O'CONNELL, Neil. Directed polymers and the quantum Toda lattice. *Ann. Probab.* 2012, vol. 40, no. 2, pp. 437–458. Available from DOI: 10.1214/10-AOP632.

98. O'CONNELL, Neil; WARREN, Jon. A multi-layer extension of the stochastic heat equation. *Commun. Math. Phys.* 2016, vol. 341, no. 1, pp. 1–33. ISSN 14320916. Available from DOI: 10.1007/s00220-015-2541-3.

99. O'CONNELL, Neil; YOR, Marc. Brownian analogues of Burke's theorem. *Stoch. Process. Appl.* 2001, vol. 96, no. 2, pp. 285–304. ISSN 03044149. Available from DOI: 10.1016/S0304-4149(01)00119-3.

100. PECCATI, Giovanni; TAQQU, Murad S. *Bocconi and Springer Series*. Wiener Chaos: Moments, Cumulants and Diagrams. Milano: Springer Milan, 2011. Bocconi & Springer Series, no. 1. ISBN 978-88-470-1678-1. ISSN 2039148X. Available from DOI: 10.1007/978-88-470-1679-8.

101. QUASTEL, Jeremy. Introduction to KPZ. *Curr. Dev. Math.* 2011, vol. 2011, no. 1, pp. 1–54. ISSN 10896384. Available from DOI: 10.4310/CDM.2011.v2011.n1.a3.

102. QUILLEN, Daniel. Rational Homotopy Theory. *Ann. Math.* 1969, vol. 90, no. 2, pp. 205–295. ISBN 0387950680. ISSN 0003486X. Available from DOI: 10.2307/1970725.

103. REUTENAUER, Christophe. Theorem of Poincare-Birkhoff-Witt, logarithm and symmetric group representations of degrees equal to stirling numbers. In: LABELLE, Gilbert; LEROUX, Pierre (eds.). *Combinatoire énumérative*. Springer, Berlin, Heidelberg, 1986, vol. 1234. Lecture Notes in Mathematics. ISBN 978-3-540-47402-9. Available from DOI: `10.1007/BFb0072520`.

104. REUTENAUER, Christophe. Free lie algebras. In: HAZEWINKEL, M. (ed.). *Handbook of Algebra*. North-Holland, 2003, vol. 3, pp. 887–903. Handbook of Algebra, no. Supplement C. ISSN 1570-7954. Available from DOI: `https://doi.org/10.1016/S1570-7954(03)80075-X`.

105. RIORDAN, J. *Combinatorial identities*. Robert E. Krieger Publishing Co., 1979. ISBN 978-0-88275-829-9. Reprint of the 1968 original.

106. ROTA, G.-C.; SHEN, J. On the combinatorics of cumulants. *J. Comb. Theory, Ser. A*. 2000, vol. 91, no. 1-2, pp. 283–304. ISSN 0097-3165. Available from DOI: `10.1006/jcta.1999.3017`. In memory of Gian-Carlo Rota.

107. ROTA, Gian-Carlo; WALLSTROM, Timothy C. Stochastic integrals: a combinatorial approach. *Ann. Probab.* 1997, vol. 25, no. 3, pp. 1257–1283. Available from DOI: `10.1214/aop/1024404513`.

108. SCHWARTZ, Laurent. *Théorie des Distributions*. 2nd ed. Paris: Hermann, 1978. ISBN 978-2-7056-5551-8.

109. SEPPÄLÄINEN, Timo. Scaling for a one-dimensional directed polymer with boundary conditions. *Ann. Probab.* 2012, vol. 40, no. 1, pp. 19–73. Available from DOI: `10.1214/10-AOP617`.

110. SPEED, T. P. Cumulants and partition lattices. *Austral. J. Stat.*. 1983, vol. 25, no. 2, pp. 378–388. ISSN 0004-9581.

111. SPOHN, Herbert. KPZ Scaling Theory and the Semi-discrete Directed Polymer Model. In: *Random Matrix Theory, Interacting Particle Systems and Integrable Systems*. Ed. by DEIFT, Percy; FORRESTER, Peter. Cambridge University Press, 2012, pp. 483–494. MSRI Publications, no. 65. Available from arXiv: `1201.0645 [cond-math.stat-mech]`.

112. STRATONOVICH, R. L. A New Representation for Stochastic Integrals and Equations. *SIAM J. Control.* 1966, vol. 4, no. 2, pp. 362–371. Available from DOI: `10.1137/0304028`.

113. SWEEDLER, Moss E. *Hopf algebras*. W. A. Benjamin, 1969. Mathematics Lecture Note.

114. SZEGŐ, Gábor. *Orthogonal polynomials*. American Mathematical Society, 1939. Colloquium Publications, no. 23. ISBN 978-1-4704-3171-6.

115. TRACY, Craig A.; WIDOM, Harold. Nonintersecting Brownian excursions. *Ann. Appl. Probab.* 2006, vol. 17, no. 3, pp. 953–979. ISSN 1050-5164. Available from DOI: `10.1214/105051607000000041`.

116. TRIEBEL, Hans. *Theory of Function Spaces*. Basel: Springer Basel, 1983. ISBN 978-3-0346-0415-4. Available from DOI: `10.1007/978-3-0346-0416-1`.

117. TRIEBEL, Hans. *Theory of Function Spaces II*. Basel: Birkhäuser, 1992. Monographs in Mathematics, no. 84. ISBN 978-3-0346-0419-2. Available from DOI: `10.1007/978-3-0346-0419-2`.

118. WIGNER, David. An identity in the free Lie algebra. *Proc. Amer. Math. Soc.* 1989, vol. 106, no. 3, pp. 639–640. Available from DOI: `10.1090/S0002-9939-1989-0969322-5`.

119. WONG, Eugene; ZAKAI, Moshe. On the Convergence of Ordinary Integrals to Stochastic Integrals. *Ann. Math. Stat.* 1965, vol. 36, no. 5, pp. 1560–1564. ISSN 00034851.

120. WONG, Eugene; ZAKAI, Moshe. On the relation between ordinary and stochastic differential equations. *Int. J. Eng. Sci.* 1965, vol. 3, no. 2, pp. 213–229.

121. YOUNG, L. C. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.* 1936, vol. 67, no. 1, pp. 251. ISSN 1871-2509. Available from DOI: `10.1007/BF02401743`.