Renormalisation of Stochastic Partial Differential Equations

Yvain Bruned (University of Edinburgh, UK), Martin Hairer (Imperial College London, UK) and Lorenzo Zambotti (Sorbonne Université Paris, France)

We present the main ideas of the renormalisation of stochastic partial differential equations, as it appears in the theory of regularity structures. We informally discuss the regularisation of the noise, the transformation of the canonical model to the renormalised one, the space of the models and the underlying algebraic structure.

In the article [Hai14], the second author of this note introduced a theory of regularity structures (RS) in order to obtain for two important equations a notion of 'solution', as well as existence and uniqueness results, which had been open problems for decades. The first part of [Hai14] is a true *theory*, in the sense that it can be applied in the same way to a broad class of problems; however, the second part, that applies this theory to two concrete examples, contains more and more *ad hoc* arguments, which must be adapted if used in different contexts. Worse, for many other interesting equations, the approach of [Hai14] becomes intractable in practice because the combinatorial complexity of the objects involved can become arbitrarily large.

   Fortunately, the situation has changed recently. The quartet of articles [Hai14, BHZ19, CH16, BCCH17] builds a fully automatic *black box* to obtain results of (local) existence and uniqueness (modulo an element of the 'renormalisation group' associated to the equation in question) for a broad class of stochastic partial differential equations (SPDEs), which includes

$$\partial_t u = \Delta u + (\partial_x u)^2 + \xi, \quad x \in \mathbf{R}, \qquad \text{(KPZ)}$$

$$\partial_t u = \Delta u + u\,\xi, \qquad\qquad x \in \mathbf{R}^2, \qquad \text{(PAM)}$$

$$\partial_t u = \Delta u - u^3 + \xi, \qquad\quad x \in \mathbf{R}^3, \qquad (\Phi^4_3)$$

for $\xi \in \mathcal{D}'(\mathbf{R}^d)$ a random, stationary and possibly very irregular distribution (Schwartz generalised function). The main example of such a random distribution is given by *space-time white noise*, but the theory applies to a very large class of ξ.

   These equations are called *singular*. Why? We can notice that it is possible to multiply a distribution $T \in \mathcal{D}'(\mathbf{R}^d)$ and a smooth function $\psi \in C^\infty(\mathbf{R}^d)$ in a canonical way, defining the product $\psi T = T\psi \in \mathcal{D}'(\mathbf{R}^d)$ by

$$(\psi T)(\varphi) = (T\psi)(\varphi) := T(\psi\varphi), \qquad \varphi \in C_0^\infty(\mathbf{R}^d).$$

But if $\psi \notin C^\infty(\mathbf{R}^d)$, this product is in general not well defined. Now, each of these equations contains some products between a distribution in $\mathcal{D}'(\mathbf{R}^d)$ and another distribution or function that is not sufficiently regular. More precisely:

- in KPZ (*Kardar-Parisi-Zhang*), u is no better than Hölder continuous in space, so the derivative $\partial_x u$ is a distribution and $(\partial_x u)^2$ is not well defined.
- in PAM (*Parabolic Anderson Model*), ξ is a white noise in space, u is a non-smooth function, so $u\,\xi$ is not well defined.

- in (Φ^4_3), u is itself a distribution and so u^3 is not defined.

In these equations, the notion of solution is therefore problematic, even before speaking of existence and uniqueness results.

Regularisation

To get around this problem, we can try to regularise the noise, solve the equation and then pass to the limit: let $\xi_\varepsilon = \varrho_\varepsilon * \xi$ be a regularisation of ξ, with $(\varrho_\varepsilon)_{\varepsilon>0}$ a family of even space-time mollifers, and let u_ε the solution of

$$\partial_t u_\varepsilon = \Delta u_\varepsilon + F(u_\varepsilon, \nabla u_\varepsilon, \xi_\varepsilon) \qquad (1)$$

where F is a non-linear function belonging to a suitable class of nonlinearities, which includes the nonlinearities of three equations above. The natural question is: what happens when $\varepsilon \to 0$? In order to control this limit, a natural approach is to look for a topology on the noises such that
1. the solution map $\Phi \colon \xi_\varepsilon \mapsto u_\varepsilon$ is continuous
2. $\xi_\varepsilon \to \xi$ when $\varepsilon \to 0$.

The first point requires a sufficiently strong topology, while the second requires a sufficiently weak topology. In fact, no solution seems possible if the regularity of ξ is too low, and even in the simplest case of stochastic *ordinary* differential equations it is a theorem that it is impossible to find a Banach space containing samples of the noise ξ and making the solution map continuous [Lyo91]. The analytic part of the regularity structures theory (RS) provides a framework for solving this problem by constructing, for a given equation,

- a metric space $(\mathcal{M}, \mathrm{d})$ called *space of models*
- a *canonical lift* of any smooth ξ_ε to a model $\mathbf{X}^\varepsilon \in \mathcal{M}$
- a *continuous function* $\mathbf{\Phi} : \mathcal{M} \to \mathcal{D}'(\mathbf{R}^d)$ such that $u_\varepsilon = \mathbf{\Phi}(\mathbf{X}^\varepsilon)$ solves the regularised equation (1), i.e., $\mathbf{\Phi}(\mathbf{X}^\varepsilon) = \Phi(\xi_\varepsilon)$.

This scheme is inspired by the theory of rough paths, initiated by Terry Lyons [Lyo98] and then developed, among others, by Massimiliano Gubinelli, whose ideas of controlled rough paths [Gub04] and branching rough paths [Gub10] served as a direct inspiration in the elaboration of regularity structures.

   The RS theory identifies a class of equations, called *subcritical*, for which the canonical model $\mathbf{X}^\varepsilon \in \mathcal{M}$ encodes *a finite number of explicit multilinear functionals* obtained from the regularised noise ξ_ε by *pointwise multiplications* and *convolutions* with the heat kernel G or some derivatives of G, like for example $\partial_x G$. Among the components of the canonical model $\mathbf{X}^\varepsilon \in \mathcal{M}$ we can thus find

$$\xi_\varepsilon, \quad \xi_\varepsilon(G * \xi_\varepsilon), \quad (\partial_x G * \xi_\varepsilon)^2, \quad \xi_\varepsilon\left(G * (\partial_x G * \xi_\varepsilon)^2\right). \quad (2)$$

On the other hand, we do not have to consider *all* the possible functions of this type: for example, we typically do not ex-

pect (nor need) to make sense of ξ^2, so we do not consider ξ_ε^2 among the components of $\mathbf{X}^\varepsilon$.

To describe the functions that make up the components of $\mathbf{X}^\varepsilon$ we use a graphical notation: each function is represented by a rooted tree, where

- the edges correspond to convolutions with G (edges of type $|$) or $\partial_x G$ (edges of type $\mathbf{|}$),
- each branching point corresponds to the pointwise product of the functions represented by the subtrees above the node in question.
- the noises are represented by nodes of type $\circ$.

For example, the four functions in (2) are represented by the following trees:

$$\circ, \qquad \overset{\circ}{\diagup}, \qquad \mathbf{\vee}, \qquad \overset{\circ}{\mathbf{\vee}}. \qquad (3)$$

Formally, we see $\mathbf{X}^\varepsilon$ as a linear map sending a space $\mathcal{H}$ of formal linear combinations of such trees into a space of distributions by writing $\mathbf{X}^\varepsilon(\overset{\circ}{\diagup}) = \xi_\varepsilon(G * \xi_\varepsilon)$, etc. Note that the trees in (3) are naturally associated with a degree by applying the following rules: white noise has degree $-\frac{d}{2}$ with d the effective dimension of the corresponding space(-time), integration against the heat kernel increases degrees by 2, differentiation lowers degree by 1, and degrees are additive under multiplication. When $d = 3$ for example, we then have $\deg \circ = -\frac{3}{2}$, $\deg \overset{\circ}{\diagup} = \deg \mathbf{\vee} = -1$ and $\deg \overset{\circ}{\mathbf{\vee}} = -\frac{1}{2}$.

By simplifying a lot, we can say that convergence in $(\mathcal{M}, d)$ corresponds to the convergence of all these explicit functions as distributions. Note, however, that $\mathcal{M}$ is not a linear space: the topology of $\mathcal{M}$ encodes quantitative versions of statements of the type "close to the point z, the distribution $\mathbf{X}^\varepsilon(\overset{\circ}{\diagup})$ is well approximated by the distribution $\mathbf{X}^\varepsilon(\circ)\mathbf{X}^\varepsilon(\overset{\circ}{\diagdown})(z)$". (Note that the latter always makes sense since the argument of $\mathbf{X}^\varepsilon(\overset{\circ}{\diagdown})$ is 'frozen' at the value z.) A major problem that appears in the examples (2) is that the products appearing in these expressions may diverge in the limit $\varepsilon \to 0$, e.g.,

$$\mathbf{E}[\xi_\varepsilon(G * \xi_\varepsilon)] = (\varrho_\varepsilon * G * \varrho_\varepsilon)(0) \to G(0) = +\infty \,,$$

so that we do not expect in general that $\mathbf{X}^\varepsilon$ converges in $(\mathcal{M}, d)$ as $\varepsilon \to 0$.
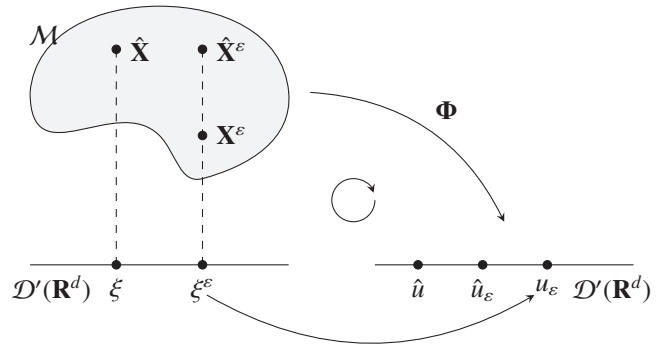
Renormalisation

To overcome this problem, we must accept that it is necessary to *modify* (*renormalise*) some components of $\mathbf{X}^\varepsilon$ and define a new lift $\hat{\mathbf{X}}^\varepsilon \in \mathcal{M}$ of ξ_ε. For example, the canonical (pointwise) product $\xi_\varepsilon(G * \xi_\varepsilon)$, which diverges when $\varepsilon \to 0$ as we have just seen, can be replaced by

$$\mathbf{X}^\varepsilon(\overset{\circ}{\diagup}) = \xi_\varepsilon(G * \xi_\varepsilon) \quad \mapsto \quad \xi_\varepsilon(G * \xi_\varepsilon) - \mathbf{E}[\xi_\varepsilon(G * \xi_\varepsilon)]$$
$$= \hat{\mathbf{X}}^\varepsilon(\overset{\circ}{\diagup}). \quad (4)$$

If, with appropriate modifications, we can build a lift $\hat{\mathbf{X}}^\varepsilon$ of ξ_ε such that

- we respect the non-linear constraints that define the space of models $\mathcal{M}$,
- the lift is 'admissible' in the sense that it respects the meaning of edges as convolution operators for planted trees, so one imposes for example that $\hat{\mathbf{X}}^\varepsilon(\mathbf{Y}) = \partial_x G * \hat{\mathbf{X}}^\varepsilon(\mathbf{\vee})$,
- we get a converging family in $(\mathcal{M}, d)$ when $\varepsilon \to 0$,

Figure 1. Illustration of the factorisation of the application $\xi_\varepsilon \mapsto u_\varepsilon$ into $\xi_\varepsilon \mapsto \mathbf{X}^\varepsilon \mapsto \Phi(\mathbf{X}^\varepsilon) = u_\varepsilon$. In the space of the models $\mathcal{M}$, there are many possible lifts of $\xi_\varepsilon \in \mathcal{D}'(\mathbf{R}^d)$, e.g. the canonical model $\mathbf{X}^\varepsilon$ and the renormalised model $\hat{\mathbf{X}}^\varepsilon$; it is only the latter which converges to a model $\hat{\mathbf{X}}$, giving a lift of ξ.

then we can use the continuity of the solution map Φ and get a family $\hat{u}_\varepsilon := \Phi(\hat{\mathbf{X}}^\varepsilon)$ converging in $\mathcal{D}'(\mathbf{R}^d)$ to some limit $\hat{u}$ which may be a reasonable candidate for being 'the' solution we are looking for.

The changes in the components of $\mathbf{X}^\varepsilon$ cannot, of course, be totally arbitrary: the non-linear structure that we have already mentioned must be preserved, see also the discussion on page 10. The *renormalisation group* $\mathcal{G}_-$ that we describe in [BHZ19] is precisely the group of transformations of $\mathcal{M}$ that respect this structure and that furthermore preserve stationarity.

This procedure can be summarised in four steps:

- *Analytical step*: Construction of the space of models $(\mathcal{M}, d)$ and continuity of the solution map $\Phi : \mathcal{M} \to \mathcal{D}'(\mathbf{R}^d)$, [Hai14].
- *Algebraic step*: Description of a group action on the space of models describing the transformation $\mathcal{M} \ni \mathbf{X}^\varepsilon \mapsto \hat{\mathbf{X}}^\varepsilon \in \mathcal{M}$ from the canonical model to the renormalised model, [BHZ19].
- *Probabilistic step*: Convergence in probability of the renormalised model $\hat{\mathbf{X}}^\varepsilon$ to a limit model $\hat{\mathbf{X}}$ in $(\mathcal{M}, d)$, [CH16].
- *Second algebraic step*: Identification of the renormalised equation satisfied by $\hat{u}_\varepsilon := \Phi(\hat{\mathbf{X}}^\varepsilon)$, [BCCH17].

The final result is a *renormalised solution* $\hat{u} := \Phi(\hat{\mathbf{X}})$, which is also the unique solution of a fixed point problem. Let us reiterate that all this works for very general noises, well beyond the Gaussian case.

Note here that the relation $\Phi(\mathbf{X}^\varepsilon) = \Phi(\xi_\varepsilon)$ is broken by renormalisation, i.e., one has in general $\Phi(\hat{\mathbf{X}}^\varepsilon) \neq \Phi(\xi_\varepsilon)$. At first glance one may be puzzled by this: have we really solved the original problem (5) or a completely different problem? The answer to this is somewhat subtle and requires us to realise that one rarely considers one single equation in isolation but is typically interested in solutions to a family of equations indexed by a number of constants (or possibly even functions). For example, in the case of the KPZ equation, we could consider the family of equations

$$\partial_t u = \partial_x^2 u + \lambda_1 (\partial_x u)^2 - \lambda_2 + \xi \,,$$

parametrised by $\lambda \in \mathbf{R}^2$. We should then view both the original 'naive' solution map Φ and the 'enhanced' solution map Φ as depending not only on the noise ξ_ε (or model $\mathbf{X}^\varepsilon$), but

also on the parameters λ describing a sufficiently large class of equations. It was then shown in [BCCH17] that the renormalisation group $\mathcal{G}_-$ already mentioned earlier does not only come with an action R on the space of models, but also with an action S on the parameter space of our class of equations, and these actions are intertwined in such a way that

$$\Phi(\lambda, R^g \mathbf{X}) = \Phi(S^g \lambda, \mathbf{X}) .$$

In particular, one can find elements $g_\varepsilon \in \mathcal{G}_-$ such that

$$\Phi(\lambda, \hat{\mathbf{X}}^\varepsilon) = \Phi(\lambda, R^{g_\varepsilon} \mathbf{X}^\varepsilon) = \Phi(S^{g_\varepsilon} \lambda, \mathbf{X}^\varepsilon) = \Phi(S^{g_\varepsilon} \lambda, \xi_\varepsilon) .$$

One way of interpreting this is that the renormalisation procedure is nothing but a *change in parametrisation* for the family of solutions $\lambda \mapsto \Phi(\lambda, \xi_\varepsilon)$. We should then interpret our convergence as $\varepsilon \to 0$ not as the convergence of a single solution in this family, but as the simultaneous convergence of the entire family of solutions. In this sense, the limiting solution family $\lambda \mapsto \Phi(\lambda, \hat{\mathbf{X}})$ should be viewed as the limit of the solution families $\lambda \mapsto \Phi(\lambda, \xi_\varepsilon)$ with the caveat that the parametrisation of this family has to be adjusted as $\varepsilon \to 0$ in order to get a non-degenerate parametrisation of the limiting family.

Note that this is precisely the same situation as arising in quantum field theory, where this change in parametrisation is the change from 'bare' to 'renormalised' coupling constants.

An example: KPZ

We consider the regularised version of the KPZ equation:

$$\partial_t u_\varepsilon = \partial_x^2 u_\varepsilon + \lambda_1 (\partial_x u_\varepsilon)^2 + \lambda_2 + \xi_\varepsilon. \tag{5}$$

The (minimal) list of the trees representing the components of a model in $\mathcal{M}$ is in this case

$$\circ, \quad \mathbf{Y}, \quad \mathbf{Y}, \quad \mathbf{Y}, \quad \mathbf{Y}, \quad \mathbf{Y}, \quad \mathbf{Y}, \quad \mathbf{Y}. \tag{6}$$

The renormalised version of the equation is then

$$\partial_t \hat{u}_\varepsilon = \partial_x^2 \hat{u}_\varepsilon + \lambda_1 (\partial_x \hat{u}_\varepsilon)^2 + \lambda_2 - \lambda_1^2 C_\varepsilon + \xi_\varepsilon,$$
$$C_\varepsilon = \mathbf{E}\big[(\partial_x G * \xi_\varepsilon)^2 \big] \sim \frac{1}{\varepsilon}. \tag{7}$$

This makes it plain that (7) is nothing but (5), but with the ε-dependent change of parameters $(\lambda_1, \lambda_2) \mapsto (\lambda_1, \lambda_2 - \lambda_1^2 C_\varepsilon)$. The first mathematical article on KPZ was [BG97], where the solution is built via the Hopf-Cole transform, which is the simple remark that $z_\varepsilon := \exp(\hat{u}_\varepsilon)$ solves the *linear* equation

$$\partial_t z_\varepsilon = \partial_x^2 z_\varepsilon + z_\varepsilon (\xi_\varepsilon - C_\varepsilon) . \tag{8}$$

For this equation, one can show, in the particular case of regularisations ξ_ε that are white in time but coloured in space, that z_ε converges when $\varepsilon \to 0$ to a random function z, solution of the Itô equation

$$\partial_t z = \partial_x^2 z + z \xi.$$

We can then *define* $\hat{u} := \log z$ (after showing that $z > 0$ everywhere almost surely). Obviously, it is $\hat{u}_\varepsilon = \log z_\varepsilon$, solution of (7), which converges to $\hat{u}$, and not u_ε.

It is not before [Hai13] that a direct approach to (5)–(7) has been obtained which then allows us to deal with a much larger class of approximating equations. The reason why mathematicians have not been able to solve this equation for fifteen years is that it is not easy to deal with the convergence of $(\partial_x \hat{u}_\varepsilon)^2 - C_\varepsilon$ when $\varepsilon \to 0$. Thanks to the RS theory,

we now know that it is enough to consider the convergence as a distribution of the family $\hat{\mathbf{X}}^\varepsilon(\tau)$ where τ varies over the family (6); for example

$$\hat{\mathbf{X}}^\varepsilon(\mathbf{Y}) = (\partial_x G * \xi_\varepsilon)^2 - \mathbf{E}[(\partial_x G * \xi_\varepsilon)^2] \tag{9}$$

which is the renormalised version of $(\partial_x G * \xi_\varepsilon)^2$. The continuity of the map Φ allows us to conclude the convergence of $\hat{u}_\varepsilon := \Phi(\hat{\mathbf{X}}^\varepsilon)$.

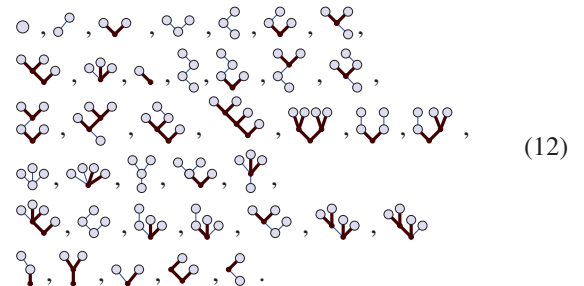## SPDEs with values in a manifold

A recent application of the RS theory is the following: in [BGHZ19] the authors of this note with F. Gabriel have constructed a natural random dynamic on the space of loops in a Riemannian manifold with metric g. This evolution can be viewed as the solution to the SPDE given in local coordinates by

$$\partial_t u^\alpha = \partial_x^2 u^\alpha + \Gamma^\alpha_{\beta\gamma}(u) \, \partial_x u^\beta \partial_x u^\gamma + \sum_{i=1}^m \sigma_i^\alpha(u) \, \xi_i , \tag{10}$$

see Figure 2. Here, Γ denotes the Christoffel symbols of the metric g while the σ_i are any finite collection of smooth vector fields such that

$$\sum_i \sigma_i^\alpha(u) \sigma_i^\beta(u) = g^{\alpha\beta}(u) . \tag{11}$$

The list of trees indexing the components of a model in the space $(\mathcal{M}, \mathrm{d})$ associated to this class of equations is much longer. For example, the most relevant trees of negative degree are the following:

$$\tag{12}$$

In [BGHZ19], natural geometric quantities such as the scalar curvature play an important and fascinating role in the study of the equation (10). It was shown there that it is possible to perform the renormalisation of this equation in such a way that solutions perform under changes of variables as expected from the naïve application of the rules of calculus and such that the law of these solutions is independent of the choice of vector fields σ_i satisfying (11).
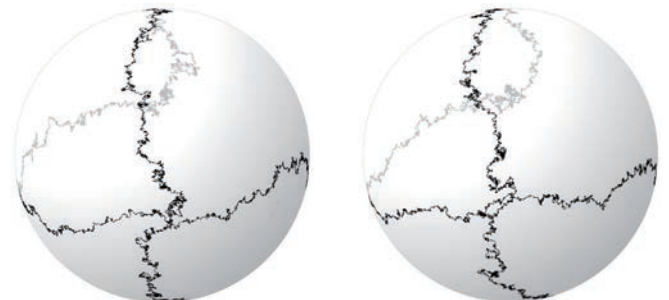

Figure 2. The solution of (10) **on the sphere at two successive times**

The algebraic structure

We can notice that the two examples of renormalised products that we discussed in (4)–(9) are simply given by the subtraction of a constant. In general, the renormalisation procedure (and therefore the transformation of $\mathbf{X}^\varepsilon$ to $\hat{\mathbf{X}}^\varepsilon$) is described in [BHZ19] by an operation of *recentering*. However, this recentering can be (much) more complicated than the simple subtraction of a constant; indeed, subtraction of a constant does not necessarily come from an 'admissible' transformation of the space of models, namely from the action of an element of the renormalisation group $\mathcal{G}_-$. It is shown in [BHZ19] that as long as the collection of trees $\mathcal{T}$ generating $\mathcal{H}$ has some properties natural in this context, there is a single (deterministic) $g_\varepsilon \in \mathcal{G}_-$, element of the renormalisation group such that if we set $\hat{\mathbf{X}}^\varepsilon(\tau) = \mathbf{X}^\varepsilon(g_\varepsilon\tau)$, then all components of $\hat{\mathbf{X}}^\varepsilon$ (corresponding to trees of negative degree, which are the only ones we ever considered in this note) have zero expectation. This is very similar to the 'BPHZ renormalisation' prescription found in the physics literature [BP57, Hep69, Zim69], so we call this particular choice of $\hat{\mathbf{X}}^\varepsilon$ the 'BPHZ lift' of the noise.

To describe the renormalisation group $\mathcal{G}_-$, we consider the algebra with unit $(\mathcal{H}_-, \cdot, \mathbf{1}_-)$ generated by the trees $\mathcal{T}_- \subset \mathcal{T}$ of negative degree (for example (6) for the KPZ equation or (12) for the loops in a manifold) and we realise $\mathcal{G}_-$ as the space of *characters* of $\mathcal{H}_-$, which are the algebra morphisms $g : \mathcal{H}_- \to \mathbf{R}$. To describe the group product in $\mathcal{G}_-$, we endow $\mathcal{H}_-$ with a structure of *coalgebra* with a *coproduct* $\Delta^- : \mathcal{H}_- \to \mathcal{H}_- \otimes \mathcal{H}_-$ which satisfies a property of *coassociativity*

$$(\Delta^- \otimes \mathrm{id})\Delta^- = (\mathrm{id} \otimes \Delta^-)\Delta^- \tag{13}$$

and a *counit* $\eta_- \in \mathcal{H}_-^*$ such that

$$(\eta_- \otimes \mathrm{id})\Delta^- = (\mathrm{id} \otimes \eta_-)\Delta^- = \mathrm{id} \tag{14}$$

on $\mathcal{H}_-$. The space $(\mathcal{H}_-, \cdot, \mathbf{1}_-, \Delta^-, \eta_-)$ is a *Hopf algebra*. The product in $\mathcal{G}_-$ is the dual of the coproduct in $\mathcal{H}_-$:

$$\mathcal{G}_- \times \mathcal{G}_- \ni (g_1, g_2) \mapsto g_1 \cdot g_2 \in \mathcal{G}_-,$$
$$(g_1 \star g_2)(h) := (g_1 \otimes g_2)\Delta^- h$$

for every $h \in \mathcal{H}_-$. The coassociativity (13) of Δ^- implies that this product is *associative*:

$$(g_1 \star g_2) \star g_3 = g_1 \star (g_2 \star g_3),$$

and the counit η_- is the neutral element such that $\eta_- \star g = g \star \eta_- = g$ for every $g \in \mathcal{G}_-$, thanks to (14). Moreover, it is possible to show that every element of $\mathcal{G}_-$ has an inverse.

We have seen that a model $\mathbf{X} \in \mathcal{M}$ is determined by a linear map $\mathbf{X} : \mathcal{H} \to \mathcal{D}'(\mathbf{R}^d)$. Now if $\mathbf{X} \in \mathcal{M}$ is a model and $g \in \mathcal{G}_-$ is an element of the renormalisation group, we can define a new model $\mathbf{X}^g = R^g\mathbf{X} \in \mathcal{M}$ by

$$\mathbf{X}^g : \mathcal{H} \to \mathcal{D}'(\mathbf{R}^d), \qquad \mathbf{X}^g(\tau) := (g \otimes \mathbf{X})\Delta^-\tau, \tag{15}$$

where $\Delta^- : \mathcal{H} \to \mathcal{H}_- \otimes \mathcal{H}$ is defined very similarly to the coproduct of $\mathcal{H}_-$.

As already alluded to on page 8, the renormalisation group $\mathcal{G}_-$, which acts on the space of models $\mathcal{M}$, must preserve another underlying algebraic structure, described by another group called $\mathcal{G}_+$ and which allows us to describe the topology of the space $\mathcal{M}$. Similar to the construction of $\mathcal{G}_-$ and $\mathcal{H}_-$, we have a Hopf algebra $(\mathcal{H}_+, \cdot, \mathbf{1}_+, \Delta^+, \eta_+)$ generated by a collection $\mathcal{T}_+$ of trees, this time of *positive* degree, and the group $\mathcal{G}_+$

is described as the character group of $\mathcal{H}_+$. The group $\mathcal{G}_+$ acts on $\mathcal{H}$ similarly to above by $\tau \mapsto (\mathrm{id} \otimes g)\Delta^+\tau$ with $\Delta^+ : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}_+$ given by a formula very similar to that defining Δ^+. A linear map $\mathbf{X} : \mathcal{H} \to \mathcal{D}'(\mathbf{R}^d)$ then defines a model if there exists a $\mathcal{G}_+$-valued function $\mathbf{R}^d \ni x \mapsto f_x \in \mathcal{G}_+$ such that, for every $x \in \mathbf{R}^d$, the 'recentred' model $\mathbf{X}_x = (\mathbf{X} \otimes f_x)\Delta^+$ satisfies a bound of the type $|\mathbf{X}_x(\tau)(\varphi_x^\lambda)| \lesssim \lambda^{\deg\tau}$, whenever φ_x^λ is a scale λ approximation of a Dirac δ-distribution centred at x. Making this quantitative yields a topology on the space of models.

The fact that this topology is preserved by $\mathcal{G}_-$ is encoded in an action of $\mathcal{G}_-$ on $\mathcal{G}_+$, that is, a group morphism of $\mathcal{G}_-$ into the (outer) automorphisms of $\mathcal{G}_+$. The action of $\mathcal{G}_-$ on $\mathcal{G}_+$ is described by a map $\Delta^- : \mathcal{H}_+ \to \mathcal{H}_- \otimes \mathcal{H}_+$ which satisfies a property called *cointeraction*:

$$\mathcal{M}^{(13)(2)(4)}(\Delta^- \otimes \Delta^-)\Delta^+ = (\mathrm{id} \otimes \Delta^+)\Delta^-, \tag{16}$$

where $\mathcal{M}^{(13)(2)(4)}(\tau_1 \otimes \tau_2 \otimes \tau_3 \otimes \tau_4) := (\tau_1 \cdot \tau_3 \otimes \tau_2 \otimes \tau_4)$.

We now define the action of $\mathcal{G}_-$ on $\mathcal{G}_+$ like this: for $g_- \in \mathcal{G}_-$ and $g_+ \in \mathcal{G}_+$, $g_- \bullet g_+ \in \mathcal{G}_+$ is given by

$$(g_- \bullet g_+)(h_+) = (g_- \otimes g_+)\Delta^- h_+, \quad \forall h_+ \in \mathcal{H}_+.$$

We can easily see that the cointeraction property (16) defines an action:

$$g_- \bullet (\bar{g}_- \bullet g_+) = (g_- \star \bar{g}_-) \bullet g_+, \quad g_-, \bar{g}_- \in \mathcal{G}_-, \quad g_+ \in \mathcal{G}_+.$$

Let us conclude by giving a simplified description of the operations $\Delta^\pm$. Recall that the spaces $\mathcal{H}, \mathcal{H}_+$ and $\mathcal{H}_-$ are realised as vector spaces generated by (possibly collections of) *rooted and decorated trees*, see for example (6) or (12) above. The operations Δ^- and Δ^+ on such trees are both constructed by using an operation *of extraction and contraction* of subforests:
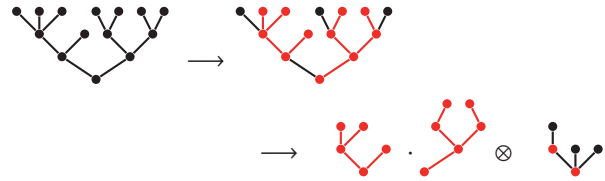
where

- we start from a tree or forest, drawn here in black on the left
- in the center, we select a subforest, colored in red
- on the right, the selected subforest is *extracted* in the left term of the tensor product, and *contracted* on the right. Note that in particular the total number of edges is always preserved by such operations.

The main difference between Δ^- and Δ^+ is in the selection of the subforests which are extracted: in the case of Δ^+, we extract only subforests consisting of a single tree that contains the root of the initial tree; in the case of Δ^-, we extract arbitrary subforests. In addition, the operation Δ^- only extracts subtrees of *negative* degree while Δ^+ only extracts those subtrees such that each 'trunk' adjacent to the root of the tree remaining on the right after contraction determines a subtree of *positive* degree.

Regarding the action of Δ^+, consider for example again the case of effective dimension 3, i.e., $\deg \circ = -\frac{3}{2}$. In this case, one has for example

$$\Delta^+\, {\overset{\circ}{\circ}}^{\circ} = {\overset{\circ}{\circ}}^{\circ} \otimes \mathbf{1} + \circ \otimes\, {\overset{\circ}{\,}}^{\circ}\,, \qquad \Delta^+ {\overset{\circ}{\,}}^{\circ} = {\overset{\circ}{\,}}^{\circ} \otimes \mathbf{1} + \mathbf{1} \otimes\, {\overset{\circ}{\,}}^{\circ}\,,$$

since these are the only ways in which we can extract/contract a subtree containing the root while being left on the right with a tree whose 'branches' touching the root are all of positive degree. A model $\mathbf{X}$ then must be such that there exists a function $x \mapsto f_x^{\bullet}$ with the property that

$$\left|(\mathbf{X}_{\circ}^{\circ} + f_x^{\circ}\mathbf{X}_\circ)(\varphi_x^{\lambda})\right| \lesssim \lambda^{-1}\,, \qquad \left|(\mathbf{X}_{\bullet}^{\circ} + f_x^{\circ})(\varphi_x^{\lambda})\right| \lesssim \lambda^{1/2}\,.$$

Note now that by admissibility, one must have $\mathbf{X}_{\bullet}^{\circ} = G * \mathbf{X}_\circ = G * \xi$, which is a Hölder continuous function. Since the exponent $\frac{1}{2}$ appearing in the second bound above is positive, this *forces* to have $f_x^{\circ} = -(G * \xi)(x)$. The first identity is then precisely of the type "near x, $\mathbf{X}_{\circ}^{\circ}$ can be approximated by $(\mathbf{X}_\circ) \cdot (\mathbf{X}_{\bullet}^{\circ})(x)$" as mentioned earlier on page 8.

Regarding the action of Δ^-, still in the same context, one has for example

$$\Delta^-{}_{\circ}^{\circ} = \mathbf{1} \otimes {}_{\circ}^{\circ} + {}_{\circ}^{\circ} \otimes \mathbf{1}\,, \quad \Delta^- {}_{\circ}^{\circ\circ} = \mathbf{1} \otimes {}_{\circ}^{\circ\circ} + {}_{\circ}^{\circ\circ} \otimes \mathbf{1} + 2\,{}_{\circ}^{\circ} \otimes {}^{\bullet}\,.$$

(In principle, according to the description given above, one should add additional terms obtained by extracting and contracting single instances of the noise $\circ$, but we can ignore these since we will always consider centred noise.) We then see that if we want to construct the BPHZ lift of a noise ξ_ε, the first identity, combined with the BPHZ prescription that $\mathbf{E}\mathbf{X}^\varepsilon\tau = 0$ for $\deg \tau < 0$, forces us to choose a character g_ε such that $g_\varepsilon({}_{\circ}^{\circ}) = -\mathbf{E}\mathbf{X}^\varepsilon{}_{\circ}^{\circ}$, while the second identity then forces us to choose $g_\varepsilon({}_{\circ}^{\circ\circ}) = -\mathbf{E}\mathbf{X}^\varepsilon{}_{\circ}^{\circ\circ}$, yielding

$$\hat{\mathbf{X}}^{\varepsilon}{}_{\circ}^{\circ\circ} = \mathbf{X}^{\varepsilon}{}_{\circ}^{\circ\circ} - \mathbf{E}\mathbf{X}^{\varepsilon}{}_{\circ}^{\circ\circ} - 2\mathbf{X}^{\varepsilon}{}^{\bullet} \cdot \mathbf{E}\mathbf{X}^{\varepsilon}{}_{\circ}^{\circ}\,.$$

The coassociativity and cointeraction properties seen above have a natural interpretation in terms of combinatorial operations on these trees and forests. Note that an algebraic structure very similar to our construction is known to arise in the numerical analysis of ordinary differential equations. There, this approach was pioneered by J. Butcher [But72] who pointed out that the natural composition operation for Runge–Kutta methods can be described by a Hopf algebra very similar to $\mathcal{H}_+$. More recently, it was pointed out by E. Hairer and his collaborators [CHV10] that an analogue of the Hopf algebra $\mathcal{H}_-$ has a natural interpretation as a 'substitution operation' for Runge–Kutta methods. The Hopf algebra $\mathcal{H}_+$ is also a generalisation of the so-called *Connes–Kreimer algebra* which was introduced in the 1990s to describe algebraically renormalisation in quantum field theory [CK98]. A review of this algebraic structure in various contexts is carried out in [CEFM11].

Bibliography

[BCCH17] Y. Bruned, A. Chandra, I. Chevyrev, and M. Hairer. Renormalising SPDEs in regularity structures. *ArXiv e-prints* (2017). To appear in J. Eur. Math. Soc. arXiv:1711.10239.

[BG97] L. Bertini and G. Giacomin. Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.* **183**, no. 3, (1997), 571–607.

[BGHZ19] Y. Bruned, F. Gabriel, M. Hairer, and L. Zambotti. Geometric stochastic heat equations. *arXiv e-prints* (2019). arXiv:1902.02884.

[BHZ19] Y. Bruned, M. Hairer, and L. Zambotti. Algebraic renormalisation of regularity structures. *Invent. Math.* **215**, no. 3, (2019), 1039–1156. arXiv:1610.08468. doi:10.1007/s00222-018-0841-x.

[BP57] N. N. Bogoliubow and O. S. Parasiuk. Über die Multiplikation der Kausalfunktionen in der Quantentheorie der Felder. *Acta Math.* **97**, (1957), 227–266. doi:10.1007/BF02392399.

[But72] J. C. Butcher. An algebraic theory of integration methods. *Math. Comp.* **26**, (1972), 79–106. doi:10.2307/2004720.

[CEFM11] D. Calaque, K. Ebrahimi-Fard, and D. Manchon. Two interacting Hopf algebras of trees: a Hopf-algebraic approach to composition and substitution of B-series. *Adv. in Appl. Math.* **47**, no. 2, (2011), 282–308. arXiv:0806.2238. doi:10.1016/j.aam.2009.08.003.

[CH16] A. Chandra and M. Hairer. An analytic BPHZ theorem for Regularity Structures. *ArXiv e-prints* (2016). arXiv:1612.08138v5.

[CHV10] P. Chartier, E. Hairer, and G. Vilmart. Algebraic structures of B-series. *Found. Comput. Math.* **10**, no. 4, (2010), 407–427. doi:10.1007/s10208-010-9065-1.

[CK98] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Comm. Math. Phys.* **199**, no. 1, (1998), 203–242. arXiv:hep-th/9808042. doi:10.1007/s002200050499.

[Gub04] M. Gubinelli. Controlling rough paths. *Journal of Functional Analysis* **216**, no. 1, (2004), 86–140. doi:10.1016/j.jfa.2004.01.002.

[Gub10] M. Gubinelli. Ramification of rough paths. *Journal of Differential Equations* **248**, no. 4, (2010), 693–721. doi:10.1016/j.jde.2009.11.015.

[Hai13] M. Hairer. Solving the KPZ equation. *Ann. Math. (2)* **178**, no. 2, (2013), 559–664. arXiv:1109.6811. doi:10.4007/annals.2013.178.2.4.

[Hai14] M. Hairer. A theory of regularity structures. *Invent. Math.* **198**, no. 2, (2014), 269–504. arXiv:1303.5113. doi:10.1007/s00222-014-0505-4.

[Hep69] K. Hepp. On the equivalence of additive and analytic renormalization. *Comm. Math. Phys.* **14**, (1969), 67–69. doi:10.1007/BF01645456.

[Lyo91] T. Lyons. On the nonexistence of path integrals. *Proc. Roy. Soc. London Ser. A* **432**, no. 1885, (1991), 281–290.

[Lyo98] T. J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana* **14**, no. 2, (1998), 215–310. doi:10.4171/RMI/240.

[Zim69] W. Zimmermann. Convergence of Bogoliubov's method of renormalization in momentum space. *Comm. Math. Phys.* **15**, (1969), 208–234. doi:10.1007/BF01645676.

Yvain Bruned [yvain.bruned@ed.ac.uk] holds a doctorate in Mathematics from the University Paris 6 and is a lecturer at the University of Edinburgh. His main interests concern stochastic partial differential equations, Regularity Structures, Rough Paths and Hopf algebras.

Martin Hairer [m.hairer@imperial.ac.uk] obtained his PhD in mathematical physics from the University of Geneva and currently holds a chair for stochastic analysis at Imperial College London. His main mathematical interests lie in stochastic dynamic, stochastic partial differential equations, and the ergodic theory of stochastic processes.

Lorenzo Zambotti [zambotti@lpsm.paris] obtained his PhD from Scuola Normale Superiore in Pisa. He is currently head of the Laboratoire de Probabilités, Statistique et Modélisation in Paris, father of two children (Nicolò and Lidia) and supervisor of four PhD students (Jean-David, Florian, Lucas and David). In the spare time, he studies stochastic partial differential equations and stochastic analysis.