Chapter 15
Application to the KPZ equation

We show how the theory of regularity structures can be used to build a robust solution theory for the KPZ equation. We also give a very short survey of the original approach to the same problem using controlled rough paths and we discuss how the two approaches are linked.

15.1 Formulation of the main result

Let us now briefly explain how the theory of regularity structures can be used to make sense of solutions to very singular semilinear stochastic PDEs. We will keep the discussion in this chapter at a very informal level without attempting to make mathematically precise statements. The interested reader may find more details in [Hai13, Hai14b].

For definiteness, we focus on the case of the KPZ equation [KPZ86], which is formally given by

$$\partial_t h = \partial_x^2 h + (\partial_x h)^2 + \xi - C \;, \tag{15.1}$$

where ξ denotes space-time white noise, the spatial variable takes values in the one-dimensional torus $\mathbb{T}$, i.e. in the interval $[0, 2\pi]$ endowed with periodic boundary conditions, and C is a fixed constant. The problem with such an equation is that even the solution to the linear part of the equation, namely

$$\partial_t \Psi = \partial_x^2 \Psi + \xi \;,$$

is not differentiable as a function of the spatial variable. As a matter of fact, as already noted in Section 12.3, for any fixed time t, Ψ has the regularity of Brownian motion as a function of the spatial variable x. As a consequence, the only way of possibly giving meaning to (15.1) is to "renormalise" the equation by subtracting from its right-hand side an "infinite constant", which counteracts the divergence of the term $(\partial_x h)^2$.

P. K. Friz, M. Hairer, *A Course on Rough Paths*, Universitext,
https://doi.org/10.1007/978-3-030-41556-3_15

This has usually been interpreted in the following way. Assuming for a moment that ξ is a smooth function, a simple consequence of the change of variables formula shows that if we define $h = \log Z$, then Z satisfies the PDE

$$\partial_t Z = \partial_x^2 Z + Z\,\xi\;.$$

The only ill-posed product appearing in this equation now is the product of the solution Z with white noise ξ. As long as Z takes values in L^2, this product can be given a meaning as a classical Itô integral, so that the equation for Z can be interpreted as the Itô equation

$$dZ = \partial_x^2 Z\,dt + Z\,dW\;, \tag{15.2}$$

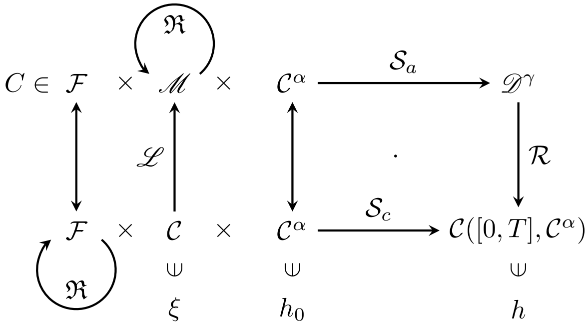
were W is an L^2-cylindrical Wiener process. It is well known [DPZ92] that this equation has a unique (mild) solution and we can then go backwards and *define* the solution to the KPZ equation as $h = \log Z$. The expert reader will have noticed that this argument appears to be flawed: since (15.2) is interpreted as an Itô equation, we should really use Itô's formula to find out what equation h satisfies. If one does this a bit more carefully, one notices that the Itô correction term appearing in this way is indeed an infinite constant! This is the case in the following sense. If W_ε is a Wiener process with spatial covariance given by $x \mapsto \varepsilon^{-1}\varrho(\varepsilon^{-1}x)$ for some smooth compactly supported function ϱ integrating to 1 and Z_ε solves (15.2) with W replaced by W_ε, then $h_\varepsilon = \log Z_\varepsilon$ solves

$$dh = \partial_x^2 h\,dt + (\partial_x h)^2\,dt + dW_\varepsilon - \varepsilon^{-1}C_\varrho\,dt\;, \tag{15.3}$$

for some constant C_ϱ depending on ϱ. Since Z_ε converges to a strictly positive limit Z, this shows that the sequence of functions h_ε solving (15.3) converges to a limit h. This limit is called the Hopf–Cole solution to the KPZ equation [Hop50, Col51, BG97, Qua11].

This notion of solution is of course not very satisfactory since it relies on a nonlinear transformation and provides no direct interpretation of the term $(\partial_x h)^2$ appearing in the right-hand side of (15.1). Furthermore, many natural growth models lead to equations that structurally "look like" (15.1), rather than (15.2). Since perturbations are usually rather badly behaved under exponentiation and since there is no really good approximation theory for (15.2) either (for example it had been an open problem for some time whether space-time regularisations of the noise lead to the same notion of solution), one would like to have a robust solution theory for (15.1) directly.

Such a robust solution theory is precisely what the theory of regularity structures provides. More precisely, it provides spaces $\mathscr{M}$ (a suitable space of "admissible models") and $\mathscr{D}^\gamma$, maps $\mathcal{S}_a$ (an abstract "solution map"), $\mathcal{R}$ (the reconstruction operator) and $\mathscr{L}$ (a "canonical lift map"), as well as a finite-dimensional group $\mathfrak{R}$ acting both on $\mathbf{R}$ and $\mathscr{M}$ such that the following diagram commutes:



(15.4)

Here, $\mathcal{S}_c$ denotes the classical solution map $\mathcal{S}_c(C,\xi,h_0)$ which provides the solution (up to some fixed final time T) to the equation

$$\partial_t h = \partial_x^2 h + (\partial_x h)^2 + \xi - C\;, \quad h(0,x) = h_0(x)\;, \tag{15.5}$$

for regular instances of the noise ξ. The space $\mathcal{F}$ of "formal right-hand sides" is in this case just a copy of **R** which holds the value of the constant C appearing in (15.5). The diagram commutes in the sense that if $M \in \mathfrak{R}$, then

$$\mathcal{S}_c(M(C),\xi,h_0) = \mathcal{R}\mathcal{S}_a(C, M(\mathscr{L}(\xi)), h_0)\;,$$

where we identify M with its respective actions on **R** and $\mathscr{M}$. A full justification of these considerations for a very large class of systems of SPDEs is beyond the scope of this text. The construction of $\mathfrak{R}$ in full generality and its action on the space of admissible models was obtained in [BHZ19]. Its adjoint action on a suitable space of equations $\mathcal{F}$ as well as the commutativity of the above diagram were then obtained in [BCCH17]. Important additional features of this picture are the following:

- If ξ_ε denotes a "natural" regularisation of space-time white noise, then there exists a sequence M_ε of elements in $\mathfrak{R}$ such that $M_\varepsilon \mathscr{L}(\xi_\varepsilon)$ converges to a limiting random element $(\Pi,\Gamma) \in \mathscr{M}$. This element can also be characterised directly without resorting to specific approximation procedures and $\mathcal{R}\mathcal{S}_a(0,(\Pi,\Gamma),h_0)$ coincides almost surely with the Hopf–Cole solution to the KPZ equation. The fact that an analogous statement "always" holds for subcritical equations was shown in the work [CH16].
- The maps $\mathcal{S}_a$ and $\mathcal{R}$ are both continuous, unlike the map $\mathcal{S}_c$ which is discontinuous in its second argument for any topology for which ξ_ε converges to ξ.
- As an abstract group, the "renormalisation group" $\mathfrak{R}$ is simply equal to $(\mathbf{R}^3,+)$. However, it is possible to extend the picture to deal with much larger classes of approximations, which has the effect of increasing both $\mathfrak{R}$ and the space $\mathcal{F}$ of possible right-hand sides. See for example [HQ18] for a proof of convergence to KPZ for a much larger class of interface growth models.

Remark 15.1. An important condition for the convergence result in [CH16] to hold is that T does *not* contain any symbol τ with $\deg\tau \le -\frac{d}{2}$ and such that τ contains more than one noise as a subsymbol. This in particular explains why fractional Brownian motion B^H with Hurst parameter H can only be lifted to a rough path when $H > \frac{1}{4}$ even though SDEs driven by fractional Brownian motion are "subcritical" for every $H > 0$. Indeed, for $H = \frac{1}{4}$, the natural degree of the symbol $\dot{\mathbb{W}}$ of Section 13.2.2 (which would be represented by in the graphical notation used earlier and contains two instances of the noise) would be $(2H-1)^- = -\frac{1}{2}^- < -\frac{d}{2}$.

An example of statement that can be proved from these considerations (see [Hai13, Hai14b, HQ18]) is the following.

Theorem 15.2. *Consider the sequence of equations*

$$\partial_t h_\varepsilon = \partial_x^2 h_\varepsilon + (\partial_x h_\varepsilon)^2 + \xi_\varepsilon - C_\varepsilon \;, \tag{15.6}$$

where $\xi_\varepsilon = \delta_\varepsilon * \xi$ with $\delta_\varepsilon(t,x) = \varepsilon^{-3}\varrho(\varepsilon^{-2}t, \varepsilon^{-1}x)$, for some smooth compactly supported function ϱ with $\int \varrho = 1$, and ξ denotes space-time white noise. Then, there exists a (diverging) choice of constants C_ε such that the sequence h_ε converges in probability to a limiting process h.

Furthermore, one can ensure that the limiting process h does not depend on the choice of mollifier ϱ and that it coincides with the Hopf–Cole solution to the KPZ equation.

Remark 15.3. It is important to note that although the limiting process is independent of the choice of mollifier ϱ, the constant C_ε *does* very much depend on this choice, as we already alluded to earlier.

Remark 15.4. Regarding the initial condition, one can take $h_0 \in \mathcal{C}^\beta$ for any fixed $\beta > 0$. Unfortunately, this result does not cover the case of "infinite wedge" initial conditions, see for example [Cor12].

The aim of this section is to sketch how the theory of regularity structures can be used to obtain this kind of convergence results and how (15.4) is constructed. First of all, we note that while our solution h will be a Hölder continuous space-time function (or rather an element of $\mathscr{D}^\gamma$ for some regularity structure with a model over $\mathbf{R}^2$), the "time" direction has a different scaling behaviour from the three "space" directions. As a consequence, it turns out to be effective to slightly change our definition of "localised test functions" by setting

$$\varphi_{(s,x)}^\lambda(t,y) = \lambda^{-3}\varphi\big(\lambda^{-2}(t-s), \lambda^{-1}(y-x)\big) \;.$$

Accordingly, the "effective dimension" of our space-time is actually 3, rather than 2. The theory presented in Chapter 13 extends *mutatis mutandis* to this setting. (Note however that when considering the degree of a regular monomial, powers of the time variable should now be counted double.) Note also that with this way of measuring regularity, space-time white noise belongs to $\mathcal{C}^{-\alpha}$ for every $\alpha > \frac{3}{2}$. This is because of the bound

$$\big(\mathbf{E}\langle \xi, \varphi_x^\lambda\rangle^2\big)^{1/2} = \|\varphi_x^\lambda\|_{L^2} \approx \lambda^{-\frac{3}{2}} \;,$$

combined with an argument somewhat similar to the proof of Kolmogorov's continuity lemma.

15.2 Construction of the associated regularity structure

Our first step is to build a regularity structure that is sufficiently large to allow to reformulate (15.1) as a fixed point in $\mathscr{D}^\gamma$ for some $\gamma > 0$. Denoting by $\mathcal{G}$ the heat kernel (i.e. the Green's function of the operator $\partial_t - \partial_x^2$), we can rewrite the solution to (15.1) with initial condition h_0 as

$$h = \mathcal{G} * \big((\partial_x h)^2 + \xi\big) + \mathcal{G}h_0 \;, \tag{15.7}$$

where $*$ denotes space-time convolution and where we denote by $\mathcal{G}h_0$ the harmonic extension of h_0. (That is the solution to the heat equation with initial condition h_0.)

Remark 15.5. We view (15.7) as an equation on the whole space by considering its periodic extension.

In order to have a chance of fitting this into the framework described above, we first decompose the heat kernel $\mathcal{G}$ as in Exercise 14.5 as

$$\mathcal{G} = K + \hat{K} \;,$$

where the kernel K satisfies all of the assumptions of Section 14.4 (with $\beta = 2$) and the remainder $\hat{K}$ is smooth. If we consider any regularity structure containing the usual Taylor polynomials and equipped with an admissible model, is straightforward to associate to $\hat{K}$ an operator $\hat{\mathcal{K}}\colon \mathscr{D}^\gamma \to \mathscr{D}^\infty$ via

$$\big(\hat{\mathcal{K}} f\big)(z) = \sum_k \frac{X^k}{k!}\big(D^k \hat{K} * \mathcal{R}f\big)(z) \;,$$

where z denotes a space-time point and k runs over all possible 2-dimensional multiindices. Similarly, the harmonic extension of h_0 can be lifted to an element in $\mathscr{D}^\infty$ which we denote again by $\mathcal{G}h_0$ by considering its Taylor expansion around every space-time point. At this stage, we note that we actually cheated a little: while $\mathcal{G}h_0$ is smooth in $\{(t,x) \,:\, t > 0, x \in \mathbb{T}\}$ and vanishes when $t < 0$, it is of course singular on the time-0 hyperplane $\{(0,x) \,:\, x \in \mathbb{T}\}$. This problem can be cured by introducing weighted versions of the spaces $\mathscr{D}^\gamma$ allowing for singularities on a given hyperplane. A precise definition of these *singular model* spaces and their behaviour under multiplication and the action of the integral operator $\mathcal{K}$ can be found in [Hai14b]; but see Exercise 4.12 for the (singular, controlled) rough path analogue. For the purpose of the informal discussion given here, we will simply ignore this problem.

This suggests that the "abstract" formulation of (15.1) should be given by

$$H = \mathcal{K}\big((\partial H)^2 + \Xi\big) + \hat{\mathcal{K}}\big((\partial H)^2 + \Xi\big) + \mathcal{G}h_0 \,, \tag{15.8}$$

where it still remains to be seen how to define an "abstract differentiation operator" ∂ realising the spatial derivative ∂_x as in Section 14.1. In view of (14.11), this equation is of the type

$$H = \mathcal{I}\big((\partial H)^2 + \Xi\big) + (\ldots) \,, \tag{15.9}$$

where the terms $(\ldots)$ consist of functions that take values in the subspace $\bar{T}$ of T spanned by regular Taylor polynomials in the time variable X_0 and the space variable X_1. (As previously, X denotes the collection of both.) In order to build a regularity structure in which (15.9) can be formulated, it is then natural to start with the structure $\bar{T}$ given by these abstract polynomials (again with the parabolic scaling which causes the abstract "time" variable to have degree 2 rather than 1), and to then add a symbol Ξ to it which we postulate to have degree $-\frac{3}{2}^-$, where we denote by α^- an exponent strictly smaller than, but arbitrarily close to, the value α. As a consequence of our definitions, it will also turn out that the symbol ∂ is always immediately followed by the symbol $\mathcal{I}$, so that it makes sense to introduce the shorthand $\mathcal{I}' = \partial\mathcal{I}$. This is also suggestive of the fact that $\mathcal{I}'$ can itself be considered an abstract integration map, associated to the kernel $K' = \partial_x K$. Comparing this to Remark 14.24, we see that we could alternatively view $\mathcal{I}'$ as the operator $\mathcal{I}_{(0,1)}$.

Remark 15.6. In order to avoid a proliferation of inconsequential terms, we impose from the start the identity $\mathcal{I}'(\mathbf{1}) = 0$ in T (we can do this by Remark 15.6). We could also set $\mathcal{I}(\mathbf{1}) = 0$ by choosing K appropriately, but this is irrelevant anyway in view of Remark 15.8 below.

We then simply add to T all of the formal expressions that an application of the right-hand side of (15.9) can generate for the description of H, ∂H, and $(\partial H)^2$. The degree of a given expression is furthermore completely determined by the rules $\deg \mathcal{I}\tau = \deg \tau + 2$, $\deg \partial\tau = \deg \tau - 1$ and $\deg \tau\bar{\tau} = \deg \tau + \deg \bar{\tau}$. For example, it follows from (15.9) that the symbol $\mathcal{I}(\Xi)$ is required for the description of H, so that $\mathcal{I}'(\Xi)$ is required for the description of ∂H. This then implies that $\mathcal{I}'(\Xi)^2$ is required for the description of the right-hand side of (15.9), which in turn implies that $\mathcal{I}(\mathcal{I}'(\Xi)^2)$ is also required for the description of H, etc. This "Picard iteration" yields the (formal) expansion, writing z for a generic space-time point,[1]

$$\begin{aligned} H(z) = h(z)\,\mathbf{1} + \mathcal{I}(\Xi) + \mathcal{I}(\mathcal{I}'(\Xi)^2) + h'(z)\,X_1 \\ + 2\mathcal{I}(\mathcal{I}'(\Xi)\mathcal{I}'(\mathcal{I}'(\Xi)^2)) + 2h'(z)\mathcal{I}(\mathcal{I}'(\Xi)) + \ldots \end{aligned}$$

where h and h' are to be considered as independent functions (similar to a controlled rough path). In particular, h may not be differentiable at all.

Remark 15.7. Here we made a distinction between $\mathcal{I}(\Xi)$, interpreted as the linear map $\mathcal{I}$ applied to the symbol Ξ, and the symbol $\mathcal{I}(\Xi)$. Since the map $\mathcal{I}$ is then

[1] Note that h' is treated as an independent function (similar to the Gubinelli derivative of a controlled path); we do not even expect h to be differentiable!

defined by $\mathcal{I}(\Xi) := \mathcal{I}(\Xi)$, this distinction is somewhat moot and will be blurred in the sequel. Similarly, the abstract (spatial) differentiation operator ∂ acts on suitable symbols as $\partial(\mathcal{I}(\ldots)) := \mathcal{I}'(\ldots)$, plus of course $\partial(X_0^{k_0}X_1^{k_1}) := k_1 X_0^{k_0}X_1^{k_1-1}$, for every multi-index (k_0, k_1).

More formally, denote by $\mathcal{U}$ the collection of those formal expressions that are required to describe H. This is then defined as the smallest collection containing X^k for all multiindices $k \geq 0$, $\mathcal{I}(\Xi)$, and such that

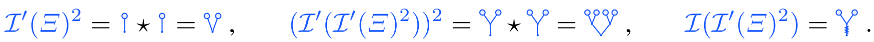
$$\tau_1, \tau_2 \in \mathcal{U} \implies \mathcal{I}(\partial\tau_1\partial\tau_2) \in \mathcal{U}\,.$$

We then set

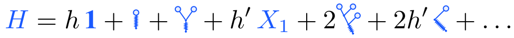
$$\mathcal{W} = \mathcal{U} \cup \{\Xi\} \cup \{\partial\tau_1\partial\tau_2 \,:\, \tau_i \in \mathcal{U}\}\,, \tag{15.10}$$

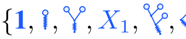
and define T as the set of all linear combinations of elements in a finite subset $\mathcal{W}_0 \subset \mathcal{W}$, sufficiently large to allow close the fixed pointed problem (15.8). Remark that this defines (implicitly!) a multiplication between some (but not all) of the symbols, notably $\partial\tau_1 \star \partial\tau_2 := \partial\tau_1\partial\tau_2$ so that we can safely omit $\star$ in the sequel. Naturally, T_α consists of those linear combinations that only involve elements in $\mathcal{W}_0$ of degree α. (Already $\mathcal{W}$ contains only finitely many elements of degree less than α, which reflects subcriticality of the problem.)

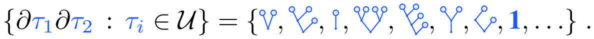
In order to simplify expressions later, we use again a shorthand graphical notation for elements of $\mathcal{W}$ as we already did in Section 14.5. Similarly to before, Ξ is represented a small circle, while the integration map $\mathcal{I}$ is represented by a downfacing wavy line and $\mathcal{I}' = \partial\mathcal{I}$ is represented by a downfacing plain line. For example, we write



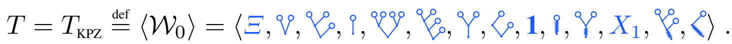
Symbols containing factors of X have no particular graphical representation, so we will for example write $X_i\mathcal{I}'(\Xi)^2 = X_i$. With this notation,



described with symbols in $\mathcal{U} = \{\mathbf{1},$ $, \ldots\}$, here spelled out up to degree $\frac{3}{2}$ (which will turn out to be "enough", cf. Remark 15.8 below). For the "right-hand side" of the equation we need to include Ξ and, spelling out symbols up to degree 0 which is the minimum required to be able to apply the reconstruction operator to it,



As it turns out, provided that we also include the noise itself, the 14 symbols encountered so far already generate a sufficiently large structure space, given by

 (15.11)

Here we ordered symbols by increasing order of degree. In fact, if τ is a tree with l circles, m plain lines and k wavy lines, then $\deg\tau = n \times \frac{3}{2}^{-} + m + 2k$. Note that deg $X_1 = 1$ for the abstract space variable, whereas due to parabolic scaling the abstract time variable has deg $X_0 = 2$ and does not show up here.

Note that at this stage, we have not defined a regularity structure yet, as we have not described a structure group G acting on T. However, similarly to what was done in (14.24), it is already natural to consider "representations" of the existing structure, which are linear maps $\boldsymbol{\Pi}$ from T into some suitable space of functions / distributions respecting a form of admissibility condition. For the sake of the present discussion, we assume that all objects are smooth. Given a (smooth) realisation of a "driving noise" ξ, we can then define its canonical lift by setting

$$\big(\boldsymbol{\Pi}\Xi\big)(x) = \xi(x)\,, \qquad \big(\boldsymbol{\Pi}X^k\big)(x) = x^k\,, \tag{15.12}$$

and then recursively by

$$\boldsymbol{\Pi}\tau\bar{\tau} = \boldsymbol{\Pi}\tau \cdot \boldsymbol{\Pi}\bar{\tau}\,, \qquad \boldsymbol{\Pi}\mathcal{I}\tau = K * \boldsymbol{\Pi}\tau\,. \tag{15.13}$$

In general, we say that a linear map $\boldsymbol{\Pi}\colon T \to \mathcal{C}(\mathbf{R}^d)$ is *admissible* if one has the relations

$$\boldsymbol{\Pi}\mathcal{I}\tau = K * \boldsymbol{\Pi}\tau\,, \qquad \boldsymbol{\Pi}\mathbf{1} = 1\,, \qquad \boldsymbol{\Pi}X^k\tau = (\bullet)^k\boldsymbol{\Pi}\tau\,. \tag{15.14}$$

(And similarly with $\mathcal{I}$ replaced by $\mathcal{I}'$ and K replaced by $\partial_x K$ in the case of KPZ...)

Such a map $\boldsymbol{\Pi}$ is clearly not a model since it is a single linear map rather than a family of such maps and the admissibility condition (14.8) is replaced by the more "natural" identity $\boldsymbol{\Pi}\mathcal{I}\tau = K * \boldsymbol{\Pi}\tau$. We will see in the next section how to construct the structure group G and how to use its construction to assign in a unique way a model to the linear map $\boldsymbol{\Pi}$.

Remark 15.8 (Where to truncate?). The (14-dimensional) space T_{KPZ} is indeed sufficient to treat the KPZ equation. Indeed, once in possession of an admissible model, thanks to Theorem 14.5, the fixed point problem (15.8) can be solved in $\mathscr{D}^\gamma$ as soon as γ is a little bit greater than $3/2$. This is why we only need to keep track of terms describing the abstract KPZ solution up to degree $3/2$. Regarding the terms required to describe the right-hand side of the fixed point problem, we need to go up to degree 0, which guarantees that the reconstruction operator (and therefore also the integration operator $\mathcal{K}$) is well-defined. This is similar to $T = T_{<1/2}$, as in Definition 13.4, being sufficient to treat rough / stochastic integration (and then SDEs) in a Brownian rough path / model context. Indeed, in that context (Proposition 13.21) consider $Y \in \mathscr{D}_0^{2\alpha}$ (now for α to be determined!) and abstract Brownian noise $\dot{W} \in \mathscr{D}^\infty_{-1/2^-}$. Then $f(Y)$, composition with a nice function f, is also in $\mathscr{D}_0^{2\alpha}$ and the product is in $\mathscr{D}^{2\alpha-1/2^-}$. We needed this exponent to be positive to have a well-defined rough integration which in turn allows to formulate a fixed point problem, so that we need $2\alpha \geq 1/2$. By definition of $\mathscr{D}^{2\alpha}$, this means that we need Y to take values in $T_{<1/2}$ which is of course what we did by working in $\langle \dot{W}, \dot{\mathbb{W}}, \mathbf{1}, W\rangle$, ignoring all symbols of higher degree.

15.3 The structure group and positive renormalisation

Recall that the purpose of the group G is to provide a class of linear maps $\Gamma\colon T \to T$ arising as possible candidates for the action of "reexpanding" a "Taylor series" around a different point. In our case, in view of (14.8) and Definition 14.3, the coefficients of these reexpansions will naturally be some polynomials in x and in the expressions appearing in (14.9). This suggests that we should define a space T^+ whose basis vectors consist of formal expressions of the type

$$X^k \prod_{i=1}^{N} \mathcal{J}_{\ell_i}(\tau_i) \;, \tag{15.15}$$

where N is an arbitrary but finite number, the τ_i are canonical basis elements in $\mathcal{W}$ defined in (15.10), and the ℓ_i are d-dimensional multiindices satisfying $|\ell_i| < \deg \tau_i + 2$. (The last bound is a reflection of the restriction of the summands in (14.9) with $\beta = 2$.) The space T^+, which also contains the empty product $\mathbf{1}$, is endowed with a natural commutative product, written as $\cdot$ or (usually) omitted. ($T^+, \cdot, \mathbf{1}$) is nothing but the free commutative algebra over the symbols $\{X_i, \mathcal{J}_\ell(\tau)\}$ with $i \in \{1,\ldots,d\}$ and $\tau \in \mathcal{W}$ with $\deg \mathcal{J}_\ell(\tau) := \deg \tau + 2 - |\ell| > 0$.)

Remark 15.9. While the canonical basis of T^+ is related to that of T, it should be viewed as a completely disjoint space. We emphasise this by using the notation $\mathcal{J}_\ell$ rather than $\mathcal{I}_\ell$.

The space T^+ also has a natural graded structure $T^+ = \bigoplus T^+_\alpha$ similarly to before by setting

$$\deg \mathcal{J}_\ell(\tau) = \deg \tau + 2 - |\ell| \;, \qquad \deg X^k = |k| \;,$$

and by postulating that the degree of a product is the sum of the degrees of its factors. Unlike in the case of T however, elements of T^+ all have strictly positive degree, except for the empty product $\mathbf{1}$ which we postulate to have degree 0.

Still inspired by (14.8), as well as by the multiplicativity constraint given by Definition 14.3, we consider the following construction. We define a linear map, sometimes called *coaction*, $\Delta^+ : T \to T \otimes T^+$ in the following way. For the basic elements Ξ, $\mathbf{1}$ and X_i ($i \in \{0,1\}$), we set

$$\Delta^+\mathbf{1} = \mathbf{1} \otimes \mathbf{1} \;, \quad \Delta^+\Xi = \Xi \otimes \mathbf{1} \;, \quad \Delta^+ X_i = X_i \otimes \mathbf{1} + \mathbf{1} \otimes X_i \;.$$

We then extend this recursively to all of T by imposing the following identities

$$\begin{aligned}
\Delta^+(\tau\bar\tau) &= \Delta^+\tau \cdot \Delta^+\bar\tau \;, \\
\Delta^+\mathcal{I}(\tau) &= (\mathcal{I} \otimes \mathrm{Id})\Delta^+\tau + \sum_\ell \frac{X^\ell}{\ell!} \otimes \mathcal{J}_\ell(\tau) \;, \\
\Delta^+\mathcal{I}'(\tau) &= (\mathcal{I}' \otimes \mathrm{Id})\Delta^+\tau + \sum_{\ell,m} \frac{X^\ell}{\ell!} \otimes \mathcal{J}_{\ell+(0,1)}(\tau) \;.
\end{aligned}$$

Here, we extend $\tau \mapsto \mathcal{J}_k(\tau)$ to a linear map $\mathcal{J}_k\colon T \to T^+$ by setting $\mathcal{J}_k(\tau) = 0$ for those basis vectors $\tau \in \mathcal{W}$ for which $\deg \tau \leq |k| - 2$. This in particular shows that the sums appearing in the above expressions are actually finite.

Let now G_+ denote the set of characters on T^+, i.e. all linear maps $g\colon T^+ \to \mathbf{R}$ with the property that $g(\sigma\bar{\sigma}) = g(\sigma)g(\bar{\sigma})$ for any two elements σ and $\bar{\sigma}$ in T^+. Then, to any such map, we can associate a linear map $\Gamma_g\colon T \to T$ by

$$\Gamma_g \tau = (\mathrm{Id} \otimes g)\Delta^+ \tau\ . \tag{15.16}$$

In principle, this definition makes sense for every $g \in (T^+)^*$. However, as already seen in (14.21) it turns out that the set of such maps with $g \in G_+$ forms a group, which we take as our structure group G by setting again

$$G \stackrel{\text{def}}{=} \{\Gamma_g : g \in G_+\}\ . \tag{15.17}$$

Remark 15.10. A less explicit way to define G is to simply take it as the set of all linear maps that are 'allowed' in the sense that they are upper triangular with the identity on the diagonal as imposed by (13.5), commute with derivatives as in Definition 14.1, are multiplicative with respect to the product as in Definition 14.3, and satisfy (14.7). See for example [Hai16].

Example 15.11 (KPZ structure group). Running through this procedure, and restricting to $T = T_{\mathrm{KPZ}}$ reveals G as a 7-dimensional (non-commutative) matrix group, canonically realised as a subgroup of the invertible maps $T \to T$, themselves representable as 16×16-matrix. Full details are left for Exercise 15.1.

Example 15.12 (KPZ). Recall $T = \langle \Xi, \RSV, \RSW, \RI, \RSVV, \RSWW, \RY, \RW, \mathbf{1}, \ldots \rangle$ in the case of KPZ. Then T^+ is linearly spanned by the symbol $\mathbf{1}$ and polynomials in the commuting symbols as (partially!) listed in

$$\{\mathcal{J}'(\RSW), \mathcal{J}'(\RI), \ldots, \mathcal{J}(\Xi), \mathcal{J}(\RSV), X_1, \mathcal{J}(\RSW), \mathcal{J}(\RI), \ldots\}$$

with (non-negative) degrees $\{\frac{1}{2}^{\equiv}, \frac{1}{2}^{-}, \ldots, 1^{=}, 1, \frac{3}{2}^{\equiv}, \frac{3}{2}^{-}, \ldots\}$ and shorthands $\mathcal{J} = \mathcal{J}_{(0,0)}, \mathcal{J}' = \mathcal{J}_{(0,1)}$. We note that all symbols here can be represented by *elementary* trees,[2] where $\mathcal{J}(\tau)$ (resp. $\mathcal{J}'(\tau)$) is represented by attaching a *single* downfacing wavy (resp. plain) line to the root of τ. For instance

$$3 \cdot \mathbf{1} - \mathcal{J}(\Xi) + 2 \cdot \mathcal{J}'(\RI) \cdot \mathcal{J}'(\RSVV) \in T^+$$

but the symbol $\mathcal{J}'(\Xi)$ (which would be of negative homogeneity) is not an element of T^+.

Before we show that G does indeed form a group (actually a subgroup of the invertible maps from T to T), we show how to use it to turn an admissible linear

[2] With some goodwill this even includes X-factors, which then appear as polynomial decorations of the trees.

map $\boldsymbol{\Pi}\colon T \to \mathcal{C}^\infty(\mathbf{R}^d)$ (in the sense of (15.14)) into a model (Π, Γ). Consider the recursion

$$f_x(\mathcal{J}_\ell(\tau)) = -\sum_{|k+\ell|<|\tau|+2} \frac{(-x)^k}{k!} \int D^{\ell+k} K(x-y)\,\bigl(\Pi_x \tau\bigr)(dy)\;,$$
$$\Pi_x \tau = (\boldsymbol{\Pi} \otimes f_x)\Delta^+ \tau\;, \tag{15.18}$$

where we furthermore impose that the f_x are characters, namely that they extend to all of $(T^+)^*$ in a multiplicative fashion, $f_x(\sigma\bar\sigma) = f_x(\sigma) f_x(\bar\sigma)$. We leave it as a simple exercise to verify that these two identities are sufficient to define the f_x and the Π_x uniquely.

Remark 15.13. The correspondence $\boldsymbol{\Pi} \Leftrightarrow (\Pi, \Gamma)$ can also be inverted and the two notions of admissibility are consistent, so that these are two completely equivalent ways of looking at admissible models for our regularity structure. Indeed, it suffices to set $\boldsymbol{\Pi}\tau = \mathcal{R}H_\tau$, where the elements $H_\tau \in \mathscr{D}^\infty$ (i.e. one can make sure that $H_\tau \in \mathscr{D}^\gamma$ for any fixed γ) are given by $H_{X^k}(x) = (X + x)^k$, $H_\Xi(x) = \Xi$, and then recursively by

$$H_{\mathcal{I}(\tau)} = \mathcal{K} H_\tau\;, \qquad H_{\tau\bar\tau} = H_\tau \cdot H_{\bar\tau}\;.$$

In particular, this correspondence does not at all rely on the fact that the model was built by lifting a smooth function. Note that this is strongly reminiscent of the construction given in Exercise 13.11. See also Exercise 15.3.

If we now define elements $F_x \in G$ by

$$F_x \tau \stackrel{\text{def}}{=} \Gamma_{f_x} = (\mathrm{Id} \otimes f_x)\Delta^+ \tau\;, \tag{15.19}$$

and then set (an expression for F_x^{-1} is given below)

$$\Gamma_{xy} = F_x^{-1} F_y\;, \tag{15.20}$$

it follows immediately from (15.18) that the Π_x and the maps Γ_{xy} do indeed satisfy the desired algebraic relation $\Pi_x \Gamma_{xy} = \Pi_y$. We also note that the coefficients of the linear maps Γ_{xy} are expressed as polynomials of the numbers $f_x(\mathcal{J}_{\ell_i}(\tau_i))$ and $f_y(\mathcal{J}_{\ell_i}(\tau_i))$ for suitable expressions τ_i and multiindices ℓ_i. Note that the linear maps $F_x : T \to T$ perform a kind of "recentering" of $\boldsymbol{\Pi}$ around x in the sense that (15.18) guarantees that, at least when $\boldsymbol{\Pi}$ is sufficiently smooth, $\Pi_x \mathcal{I}(\tau)$ vanishes at the order determined by the degree of τ. As a matter of fact, one could even have taken this as the defining property of the maps F_x (together with the fact that they are of the form (15.19) for some multiplicative functional f_x). We will see in Section 15.5 below that the renormalisation procedure required to give a meaning to singular SPDEs like the KPZ equation can equally be interpreted as a type of recentering procedure, but this time in "probability space". This also explains the terminology "positive renormalisation" which is sometimes encountered for the maps F_x.

We now argue that G as defined above actually forms a group, so that in particular the maps F_x are invertible. To this end, define a linear map $\Delta^+ \colon T^+ \to T^+ \otimes T^+$, very similarly to the previously defined map $\Delta^+ \colon T \to T \otimes T^+$, by

$$\Delta^+ \mathbf{1} = \mathbf{1} \otimes \mathbf{1} \;, \qquad \Delta^+ X = X \otimes \mathbf{1} + \mathbf{1} \otimes X \;,$$

extended recursively to all of T^+ by imposing the identities, for all multiindices k,

$$\begin{aligned} \Delta^+(\sigma\bar\sigma) &= (\Delta^+\sigma)(\Delta^+\bar\sigma) \;, \\ \Delta^+ \mathcal{J}_k(\tau) &= (\mathcal{J}_k \otimes \mathrm{Id})\Delta^+\tau + \sum_{\ell \in \mathbf{N}^2} \frac{X^\ell}{\ell!} \otimes \mathcal{J}_{\ell+k}(\tau) \;. \end{aligned} \tag{15.21}$$

It can be verified that Δ^+ is *coassociatve* in the sense

$$(\Delta^+ \otimes \mathrm{Id})\Delta^+ = (\mathrm{Id} \otimes \Delta^+)\Delta^+ \;. \tag{15.22}$$

This and the multiplicative property make Δ^+ a *coproduct* and T^+ a (connected, graded) *coalgebra*. From general principles there exists a unique linear map $\mathcal{A}^+ \colon T^+ \to T^+$, called *antipode*, so that $(T^+, \cdot, \Delta^+, \mathcal{A}^+)$ is a *Hopf algebra*. Moreover, our notational overload is justified by the fact that (15.22) also holds when both sides of the identity are interpreted as linear maps $T \to T \otimes T^+ \otimes T^+$.

We then define a product $\circ$ on the space of linear functionals $f \colon T^+ \to \mathbf{R}$ by

$$(f \circ g)(\sigma) = (f \otimes g)\Delta^+\sigma \;, \tag{15.23}$$

noting that coassociativity of Δ^+ implies associativity of $\circ$. Restricted to multiplicative elements, i.e. to G_+, the definition of the antipode implies that G_+ is indeed a group with $f^{-1} = f\mathcal{A}^+$, that is $f^{-1} \circ f = f \circ f^{-1} = e$, where $e \colon T^+ \to \mathbf{R}$ maps every basis vector of the form (15.15) to zero, except for $e(\mathbf{1}) = 1$. This is a general construction for Hopf algebras and G_+ is known as the *character group* of T^+. The product $\circ$ in this context is usually called the *convolution product*. Indeed, the first identity in (15.21), valid by definition for every coproduct in a Hopf algebra, ensures that if f and g belong to G_+, then $f \circ g \in G_+$. (Spelled out, this says if $f, g \in (T^+)^*$ are both multiplicative in the sense that $f(\sigma\bar\sigma) = f(\sigma)f(\bar\sigma)$, then $f \circ g$ is again multiplicative.)

Since, by definition, $\Gamma_f = (\mathrm{Id} \otimes f)\Delta$ we can rewrite (15.19) as $F_x = \Gamma_{f_x}$, and the intertwining identity (15.22) entails that

$$\Gamma_{f \circ g} = \Gamma_f \Gamma_g \;.$$

Also, the element e is neutral in the sense that Γ_e is the identity operator, and as a consequence $\Gamma_{f^{-1}} = \Gamma_f^{-1}$ whenever $f \in G_+$. In particular then,

$$F_x^{-1} = \Gamma_{f_x^{-1}} = \Gamma_{f_x \mathcal{A}^+}$$

and we can fully spell out (15.20) as

$$\Gamma_{xy} = \Gamma_{f_x \mathcal{A}^+ \circ f_y} = (\mathrm{Id} \otimes \gamma_{x,y})\Delta^+ \;, \quad \gamma_{x,y} \stackrel{\text{def}}{=} f_x \mathcal{A}^+ \circ f_y = (f_x \mathcal{A}^+ \otimes f_y)\Delta^+ \;.$$

The fact that Δ^+ preserves degree (as can be seen by induction from its definition) and that elements of T^+ all have strictly positive degree, except for $\mathbf{1}$ leads to the conclusion that, for every $\Gamma \in G$ and every $\tau \in T$, $\Gamma\tau$ is indeed of the form (13.5). The multiplicativity property of Δ^+ furthermore guarantees that the constraint mentioned in Definition 14.3 does hold. This justifies our definition of structure group G associated to T as the set of all multiplicative linear functionals on T^+, acting on T via (15.16), as given in (15.17), for G has group structure induced from G_+.

Returning to the relation between Π_x and $\boldsymbol{\Pi}$, we showed actually more, namely that the knowledge of $\boldsymbol{\Pi}$ and the knowledge of (Π, Γ) are equivalent. Indeed, on the one hand one has $\boldsymbol{\Pi} = \Pi_x F_x^{-1}$ and the map F_x can be recovered from Π_x by (15.18) and (15.19). On the other hand however, one also has of course $\Pi_x = \boldsymbol{\Pi} F_x$ and, if we equip T with an adequate recursive structure, then we have already seen that the coefficients f_x are uniquely determined by $\boldsymbol{\Pi}$.

Furthermore, the correspondence $(\Pi, \Gamma) \leftrightarrow \boldsymbol{\Pi}$ outlined above works for *any* admissible model and does not at all rely on the fact that it was built by lifting a continuous function. In particular, it does *not* rely on the fact that Π_x and $\boldsymbol{\Pi}$ are multiplicative. In the general case, the first identity in (15.13) may then of course fail to be true, even if $\boldsymbol{\Pi}\tau$ happens to be a continuous function for every $\tau \in T$. The only reason why our definition of an admissible model does not simply consist of the single map $\boldsymbol{\Pi}$ is that there seems to be no simple way of describing the topology given by Definition 13.5 in terms of $\boldsymbol{\Pi}$.

15.4 Reconstruction for canonical lifts

Recall that, given any sufficiently regular function ξ (say a continuous space-time function), there is a canonical way of lifting ξ to an admissible model $\mathscr{L}\xi = (\Pi, \Gamma)$ for T by imposing (15.12) and (15.13), and then turning $\boldsymbol{\Pi}$ into a model as described in the previous paragraph. With such a model $\mathscr{L}\xi$ at hand, it follows from (15.13) and (13.26) that the associated reconstruction operator satisfies the properties

$$\mathcal{R}\mathcal{K}f = K * \mathcal{R}f \;, \qquad \mathcal{R}(fg) = \mathcal{R}f \cdot \mathcal{R}g \;,$$

as long as all the functions to which $\mathcal{R}$ is applied belong to $\mathscr{D}^\gamma$ for some $\gamma > 0$. As a consequence, applying the reconstruction operator $\mathcal{R}$ to both sides of (15.8), we see that if H solves (15.8) then, provided that the model $(\Pi, \Gamma) = \mathscr{L}\xi$ was built as above starting from any *continuous* realisation ξ of the driving noise, the function $h = \mathcal{R}H$ solves the equation (15.1).

At this stage, the situation is as follows. For any *continuous* realisation ξ of the driving noise, we have factorised the solution map $(h_0, \xi) \mapsto h$ associated to (15.1) into maps

$$(h_0, \xi) \mapsto (h_0, \mathscr{L}\xi) \mapsto H \mapsto h = \mathcal{R}H \;,$$

where the middle arrow corresponds to the solution to (15.8) in some weighted $\mathscr{D}^\gamma$-space. The advantage of such a factorisation is that the last two arrows yield *continuous* maps, even in topologies sufficiently weak to be able to describe driving noise having the lack of regularity of space-time white noise. The only arrow that isn't continuous in such a weak topology is the first one. At this stage, it should be believable that a similar construction can be performed for a very large class of semilinear stochastic PDEs, provided that certain scaling properties are satisfied. This is indeed the case and large parts of this programme have been carried out in [Hai14b].

Given this construction, one is lead naturally to the following question: given a sequence ξ_ε of "natural" regularisations of space-time white noise, for example as in (15.6), do the lifts $\mathscr{L}\xi_\varepsilon$ converge in probably in a suitable space of admissible models? Unfortunately, unlike in the theory of rough paths where this is very often the case (see Section 10), the answer to this question in the context of SPDEs is often an emphatic **no**. Indeed, if it were the case for the KPZ equation, then one could have been able to choose the constant C_ε to be independent of ε in (15.6), which is certainly not the case.

15.5 Renormalisation of the KPZ equation

One way of circumventing the fact that $\mathscr{L}\xi_\varepsilon$ does not converge to a limiting model as $\varepsilon \to 0$ is to consider instead a sequence of *renormalised* models. The main idea is to exploit the fact that our definition of an admissible model does *not* impose the multiplicative identity

$$\boldsymbol{\Pi}\tau\bar{\tau} = \boldsymbol{\Pi}\tau \cdot \boldsymbol{\Pi}\bar{\tau} ,$$

used in (15.13) for the canonical lift, even in situations where ξ itself happens to be a continuous function. One question that then imposes itself is: what are the natural ways of "deforming" the usual product which still lead to lifts to an admissible model? It turns out that the regularity structure whose construction was sketched above comes equipped with a natural *finite-dimensional* group of continuous transformations $\mathfrak{R}$ on its space of admissible models (henceforth called the "renormalisation group"), which essentially amounts to the space of all natural deformations of the product. It then turns out that even though the canonical lift $\mathscr{L}\xi_\varepsilon$ does not converge, it is possible to find a sequence M_ε of elements in $\mathfrak{R}$ such that the sequence $M_\varepsilon\mathscr{L}\xi_\varepsilon$ converges to a limiting model $(\hat{\Pi}, \hat{\Gamma})$. Unfortunately, the elements M_ε do *not* preserve the image of $\mathscr{L}$ in the space of admissible models. As a consequence, when solving the fixed point map (15.8) with respect to the model $M_\varepsilon\mathscr{L}\xi_\varepsilon$ and inserting the solution into the reconstruction operator, it is not clear *a priori* that the resulting function (or distribution) can again be interpreted as the solution to some modified PDE. It turns out that in the present setting this is again the case and the modified equation is precisely given by (15.6), where C_ε is some linear combination of the constants appearing in the description of M_ε.

There are now three questions that remain to be answered:

1. How does one construct the renormalisation group $\mathfrak{R}$?
2. How does one derive the new equation obtained when renormalising a model?
3. What is the right choice of M_ε ensuring that the renormalised models converge?

As already pointed out at the start of this chapter, these questions have now been answered in full generality in the series of articles [Hai14b, BHZ19, CH16, BCCH17]. The aim of this section is to illlustrate how the machinery developed there applies to the particular case of the KPZ equation and go give a feeling for how the main steps of the construction generalise to other settings.

15.5.1 The renormalisation group

How does all this help with the identification of a natural class of deformations for the usual product? Throughout this section, we will only consider models constructed from a single map $\boldsymbol{\Pi}$ by the recursive procedure given in (15.18), combined with (15.20). At this point, we crucially note that if $\boldsymbol{\Pi}\colon T \to \mathcal{C}^\infty(\mathbf{R}^d)$ is an arbitrary admissible linear map (in the sense that $\boldsymbol{\Pi}\mathcal{I}\tau = K * \boldsymbol{\Pi}\tau$ as before), then there is no reason in general for (15.18) and (15.20) to define a model. The reason is that while these definitions do guarantee that $\Pi_x\mathcal{I}\tau$ satisfies the first bound in (13.13), there is no reason in general for $\big(\Pi_x\tau\big)(y)$ to vanish at the right order as $y \to x$ for an arbitrary symbol τ that is not obtained by applying the integration map to some other symbol. It is however the case that these bounds hold whenever $\boldsymbol{\Pi}$ is obtained as the canonical lift of a smooth function, as can easily be seen from the multiplicativity property of the canonical lift.

This suggests to define a space $\mathscr{M}_\infty$ consisting of those admissible maps $\boldsymbol{\Pi}\colon T \to \mathcal{C}^\infty(\mathbf{R}^d)$ which do generate a model by the above procedure. By Remark 15.13, there is a canonical bijection between $\mathscr{M}_\infty$ and the set of all smooth admissible models, so we henceforth also call an element $\boldsymbol{\Pi} \in \mathscr{M}_\infty$ simply a model (or an admissible model). Note that even though the space of linear maps $T \to \mathcal{C}^\infty(\mathbf{R}^d)$ is linear, the space $\mathscr{M}_\infty$ is far from being a linear space.

At this stage, we would like to introduce probability into the game. For this, note first that we have a natural action S of the group of translations $(\mathbf{R}^d, +)$ onto T by setting $S_h X^k = (X + h)^k$, $S_h \Xi = \Xi$, and then recursively by

$$S_h\mathcal{I}\tau = \mathcal{I}S_h\tau\;, \qquad S_h\tau\bar\tau = S_h\tau S_h\bar\tau\;.$$

We then note that if ξ happens to be a stationary stochastic process and $\boldsymbol{\Pi} = \mathscr{L}\xi$ is its canonical lift as a random model, then $\boldsymbol{\Pi}$ is a stationary stochastic process in the generalised sense that

$$\big(\boldsymbol{\Pi}\tau\big)(\cdot + h) \stackrel{\text{law}}{=} \big(\boldsymbol{\Pi}S_h\tau\big)(\cdot)\;.$$

In order to define the renormalisation group $\mathfrak{R}$, it is then natural to consider only transformations of the space of admissible models that preserve this property. Since we are not in general allowed to multiply components of $\boldsymbol{\Pi}$ and we do not want to "pull arbitrary functions out of a hat", the only remaining operation is to form linear combinations. It is therefore natural to look for linear maps $M\colon T \to T$ which furthermore preserve $\mathscr{M}_\infty$ in the sense that if, given $\boldsymbol{\Pi} \in \mathscr{M}_\infty$, we define $\boldsymbol{\Pi}^M$ by

$$\boldsymbol{\Pi}^M \tau = \boldsymbol{\Pi} M \tau\,, \tag{15.24}$$

one would like to have again $\boldsymbol{\Pi}^M \in \mathscr{M}_\infty$. It is clear that in order to guarantee this, M needs to commute with the integration operators $\mathcal{I}$ and $\mathcal{I}'$, but this alone is by no means sufficient.

It turns out that the construction of a natural family of operators with the required properties goes in a way that is strongly reminiscent of the construction of the structure group, but with many aspects of the construction "reversed". A natural starting point of the construction is given by the set $\mathcal{W}_- \subset \mathcal{W}$ consisting of the canonical basis vectors of *strictly negative* degree of our regularity structure T which furthermore have the property that they can be built from products and integrations applied to Ξ, i.e. do not involve any X^k for $k > 0$. We then define T^- similarly to T^+ as the free unital algebra generated by $\mathcal{W}_-$, i.e.[3]

$$T^- \stackrel{\text{def}}{=} \mathrm{Alg}\Big(\Big\{$$  $$\Big\}\Big)\,,$$

the algebra given by all polynomials with real coefficients and indeterminates in $\mathcal{W}_-$; the unit is denoted by $\mathbf{1}$ (or, equivalently, as the empty forest $\emptyset$). The reason why $\mathcal{W}_-$ is expected to play a major role is that, by combing Exercise 13.11 with admissibility and multiplicativity of the action of Γ, $\boldsymbol{\Pi}\tau$ for $\deg \tau > 0$ is uniquely determined by the knowledge of $\boldsymbol{\Pi}\tau$ for all symbols τ with $\deg \tau \le 0$.

By analogy with the BPHZ renormalisation procedure in quantum field theory [BP57, Hep69, Zim69], it is natural to look for renormalisation maps that consist in "contracting subtrees of negative degree". In order to formalise such an operation, we take more seriously the interpretation of the canonical basis elements of T as "trees". More precisely, we consider labelled trees $\tau = (V, E, \varrho, \mathfrak{n}, \mathfrak{e})$, where V is a finite vertex set, $E \subset V \times V$ is an edge set, $\varrho \in V$ is a root, $\mathfrak{n}\colon V \to \mathbf{N}^d$ is a "polynomial label" and $\mathfrak{e}\colon E \to \{\Xi, \mathcal{I}, \mathcal{I}'\}$ is an "edge label". As usual, we identify labelled trees if they can be related by a tree isomorphism preserving the root and labels. The way this correspondence works is as follows. The symbol X^k is represented as the (unique) tree with a sole vertex $V = \{\varrho\}$ and polynomial label $\mathfrak{n}(\varrho) = k$. The symbol Ξ is represented by the tree with two vertices $V = \{\varrho, \bullet\}$, one (oriented) edge $E = \{e\} = \{(\bullet, \varrho)\}$, and labels $\mathfrak{n} = 0$, $\mathfrak{e}(e) = \Xi$. Integration is then performed by adding an edge of the corresponding type to the root, i.e. we have for example

[3] As in the case of rough volatility, cf. 14.26, we colour basis elements of T^- differently to distinguish them from those of T and / or T^+. Elements in T^- are naturally represented as (unordered) *forests*.

$$\mathcal{I}(V, E, \varrho, \mathfrak{n}, \mathfrak{e}) = (V \sqcup \{\bar{\varrho}\}, E \sqcup \{(\varrho, \bar{\varrho})\}, \bar{\varrho}, \mathcal{I}\mathfrak{n}, \mathcal{I}\mathfrak{e}) ,$$

where $\mathcal{I}\mathfrak{n}(\bar{\varrho}) = 0$ and otherwise agrees with $\mathfrak{n}$, while $\mathcal{I}\mathfrak{e}((\varrho, \bar{\varrho})) = \mathcal{I}$ and again otherwise agrees with $\mathfrak{e}$. Multiplication is obtained by joining roots:

$$(V, E, \varrho, \mathfrak{n}, \mathfrak{e}) \cdot (\bar{V}, \bar{E}, \bar{\varrho}, \bar{\mathfrak{n}}, \bar{\mathfrak{e}}) = ((V \sqcup \bar{V})/\{\varrho, \bar{\varrho}\}, E \sqcup \bar{E}, \{\varrho, \bar{\varrho}\}, \mathfrak{n} \sqcup \bar{\mathfrak{n}}, \mathfrak{e} \sqcup \bar{\mathfrak{e}}) ,$$

where $(\mathfrak{n} \sqcup \bar{\mathfrak{n}})(\{\varrho, \bar{\varrho}\}) = \mathfrak{n}(\varrho) + \bar{\mathfrak{n}}(\bar{\varrho})$.

Remark 15.14. This is nothing but a formalisation of the graphical notation already used earlier. The notation used in (15.11) for example suggests that one could equivalently have viewed the noise as part of a "vertex label" and this is the viewpoint taken for example in [BCCH17]. It appears however that viewing noises as edges, as for example in [BHZ19], usually yields a more consistent formalism. This is especially the case in situations where one would like to "attach" additional information to noises as done in [CCHS20, Sec. 5].

In a similar way, elements of T^- can be interpreted as elements $A = (V, E, \varrho, \mathfrak{e})$ as above, except that there is no "polynomial label" $\mathfrak{n}$ and (V, E) is allowed to be a forest, with ϱ denoting the set of its roots, one per connected component. In particular, the empty forest $V = \emptyset$ is allowed, which wasn't the case for T.

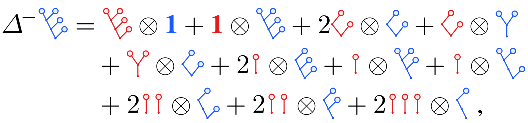
Given $A = (\bar{V}, \bar{E}, \bar{\varrho}, \bar{\mathfrak{e}}) \in T^-$ and $\tau = (V, E, \varrho, \mathfrak{n}, \mathfrak{e}) \in T$, we say that $A \subset \tau$ if one has an injective map $\iota \colon \bar{V} \sqcup \bar{E} \to V \sqcup E$ preserving connectivity and edge labels. Note that the injectivity of ι implies in particular that the different connected components of A are vertex-disjoint in τ. In such a situation, we then write $\mathcal{R}_A\tau$ for the tree obtained by contracting the connected components of A in τ, i.e. the vertex set of $\mathcal{R}_A\tau$ consists of $V/\!\sim$ where $v \sim \bar{v}$ if v and $\bar{v}$ are equal or belong to the image of the same connected component of A, while the edge set of $\mathcal{R}_A\tau$ equals $E \setminus \iota\bar{E}$.

We then define an operator $\Delta^- : T \to T^- \otimes T$ by

$$\Delta^-\tau = \sum_{A \subset \tau} \mathcal{Q}_- A \otimes \mathcal{R}_A\tau , \tag{15.25}$$

where $\mathcal{Q}_- A = A$ if every connected component of A has negative degree and $\mathcal{Q}_- A = 0$ otherwise. Note again the graphical interpretation of extracting possibly empty collections of subtrees of negative degree.

Example 15.15. For the regularity structure associated to the KPZ equation, we have for example[4]



(15.26)

where we used red symbols to denote elements of T^- just as in Section 14.5. In most situations it is natural to only consider characters of T^- that vanish on planted trees,

[4] Mind that  $\equiv$  $\subset$  in three distinct ways which explains the terms 2 $\otimes$  +  $\otimes$ .

i.e. trees with only one edge incident to the root,[5] in which case this simplifies to

Note also that there is for example no term  appearing in (15.26); indeed  fails to have negative degree, hence is not an element of T^- and killed by $\mathcal{Q}_-$.

Remark 15.16. Since $\mathcal{I}'(\mathbf{1}) = 0$ by Remark 15.6, there is no term such as  appearing in the right-hand side of (15.26).

Remark 15.17. While the present construction is sufficient for KPZ, in full generality, one should also allow polynomial decorations for elements in T^- in which case the expression for Δ^- involves additional combinatorial factors, similarly to the definition of Δ^+.

Our motivation for the definition of Δ^- is as follows. Assigning a number to each $\tau \in \mathcal{W}_-$ is equivalent to choosing an algebra morphism $g\colon T^- \to \mathbf{R}$. If we ignore for a moment the labels $\mathfrak{n}$ and $\mathfrak{e}$, an operation of the type $M_g\colon T \to T$ with

$$M_g \tau = (g \otimes \mathrm{Id})\Delta^- \tau \;, \tag{15.27}$$

then corresponds to iterating over all ways of contracting subtrees of negative degree contained in τ and replacing them by the corresponding constant assigned to it by g. This corresponds to replacing a kernel of possibly several variables by a multiple of a Dirac delta function forcing all arguments to collapse.

Similarly to before, one can also define an operator $\Delta^-\colon T^- \to T^- \otimes T^-$ by setting

$$\Delta^- B = \sum_{A \subset B} \mathcal{Q}_- A \otimes \mathcal{Q}_- \mathcal{R}_A B \;,$$

where the notions of inclusion $A \subset B$ and the contraction $\mathcal{R}_A B$ are defined in complete analogy to above.

This yields an algebraic structure very similar to the one given by T and T^+. We will however not describe it in any more detail here, but refer instead to [BHZ19] for additional details. In particular, T^-, with forest product and coproduct Δ^-, admits an antipode $\mathcal{A}^-$ turning it into a commutative Hopf algebra. Its characters then form a group with product analogous to (15.23) and inverse given by $g \mapsto g\mathcal{A}^-$, acting on T by (15.27).

Definition 15.18. The *renormalisation group* $\mathfrak{R}$ for our regularity structure T is defined as the character group of T^-.

Remark 15.19. The original definition of the "renormalisation group" given in [Hai14b] (and in the first edition of this book) is slightly more general. In the situation of the regularity structure built for a two-component KPZ equation, i.e.

[5] In essence, extracting negative trees will help to renormalise otherwise ill-posed products. A single edge incident to the root corresponds to convolution with a (compactly supported) kernel, which is always well-posed.

exactly the same as discussed here, except that there are two "noises" Ξ_1 and Ξ_2 and every occurrence of Ξ can be replaced by either of them, the old definition would for example include the map M that swaps the two noises in a consistent way. (Consistency is in the sense that $M\mathcal{I}'(\Xi_2)\mathcal{I}'(\mathcal{I}'(\Xi_1)^2) = \mathcal{I}'(\Xi_1)\mathcal{I}'(\mathcal{I}'(\Xi_2)^2)$ for example.) This is not an operation that is described by a character of T^-. The advantage of the present definition is that it is much more explicit. Furthermore, it follows from the analytical results of [CH16] that it is sufficiently large to serve the purpose of renormalising divergent models.

Example 15.20. Continuing the above example, we have

Note that we have not considered the simplification of removing planted trees. Instead, the analogues of the remaining terms appearing in (15.26) are killed by the projection $\mathcal{Q}_-$. We also note that this expression is symmetric in the two factors T^- which is the case for all the symbols appearing in the analysis of the KPZ equation. This implies that the KPZ renormalisation group $\mathfrak{R}$ is abelian. (In general though, the presence of "overlapping divergencies" can cause $\mathfrak{R}$ to be non-abelian.)

One of the main results of [BHZ19] is a generalisation of the following statement, which shows that the action of the renormalisation group plays nice with our notion of admissible model.

Theorem 15.21. *Let $g \in \mathfrak{R}$ and define $M_g = (g \otimes \mathrm{Id})\Delta^-$ as in (15.27). Then, for any $\boldsymbol{\Pi} \in \mathscr{M}_\infty$, one has $\boldsymbol{\Pi}^g \stackrel{\text{def}}{=} \boldsymbol{\Pi} M_g \in \mathscr{M}_\infty$. Furthermore, one has*

$$\Pi_x^g = \Pi_x M_g\;, \qquad \Gamma_{xy}^g = M_g^{-1}\Gamma_{xy}M_g\;. \tag{15.28}$$

Proof. We sketch the proof. Recall that Δ^- has been defined (with notational overload) as map from $T \to T^- \otimes T$ and $T^- \to T^- \otimes T^-$; we now also define $\Delta^- : T^+ \to T^- \otimes T^+$ as multiplicative linear map, determined by

$$\Delta^- X_i = \mathbf{1} \otimes X_i, \qquad \Delta^- \mathcal{J}_\ell(\tau) = (\mathrm{Id} \otimes \mathcal{J}_\ell(\cdot))\Delta^-\tau\;.$$

In the special case of KPZ one can check by hand that, thanks in particular to the fact that $\mathcal{I}'(\mathbf{1}) = 0$ by Remark 15.6 (which correctly suggests that we should also impose $\mathcal{J}'(\mathbf{1}) = 0$),

(i) On T one has the *cointeraction formula*

$$M_{13}(\Delta^- \otimes \Delta^-)\Delta^+ = (\mathrm{Id} \otimes \Delta^+)\Delta^-\;, \tag{15.29}$$

where $M_{13} : T^- \otimes T \otimes T^- \otimes T^+ \to T^- \otimes T \otimes T^+$ is the map that multiplies the first and third factor (in T^-), and the same holds also on T^+.

(ii) The actions of $\mathfrak{R}$ onto T and T^+ given by M_g do not decrease the degree. (For the relevant set of characters g, this is seen explicitly in Exercise 15.2.)

Recall the correspondence $\boldsymbol{\Pi} \Leftrightarrow (\Pi, \Gamma)$ given in Remark 15.13. With the special properties (i)-(ii) it is straightforward to verify that, for $g \in \mathfrak{R}$ arbitrary, $\boldsymbol{\Pi}^g = \boldsymbol{\Pi} M_g$ defines a model $\boldsymbol{\Pi}^g \Leftrightarrow (\Pi^g, \Gamma^g)$ with

$$\Pi_x^g = \Pi_x M_g = (g \otimes \Pi_x)\Delta^- \;, \quad \Gamma_{xy}^g = (\mathrm{Id} \otimes \gamma_{x,y}^g)\Delta^+ \;, \quad \gamma_{xy}^g \stackrel{\text{def}}{=} (g \otimes \gamma_{xy})\Delta^- \;.$$

(The second identity in (15.28) then follows from the formula for γ_{xy}^g, combined with the cointeraction formula.) To show all this, first write $f_x = f_x^{\boldsymbol{\Pi}}$ for f_x obtained from $\boldsymbol{\Pi}$ as in (15.18). One shows recursively that

$$f_x^{\boldsymbol{\Pi}^g} = f_x^{\boldsymbol{\Pi}} M_g \;.$$

One then uses (i), on T, to show that the required identity for Π_z^g holds. Finally, one uses (i), on T^+ to show that if one views $M_g = (g \otimes \mathrm{Id})\Delta^-$ as acting on T^+, then its action distributes over the product in the character group defined in (15.23) in the sense that $(M_g f) \circ (M_g \bar{f}) = M_g(f \circ \bar{f})$, which then implies the required identity for γ_{xy}^g. The fact that the action of M_g does not decrease degrees guarantees that (Π^g, Γ^g) is again a model (since (Π, Γ) is). $\square$

Remark 15.22. In general (i.e. in the case of similar regularity structures set up for different examples of subcritical semilinear SPDEs), the cointeraction property (15.29) may fail. It turns out however that it can still be rescued by working in a suitably extended regularity structure, see [Hai16, BHZ19].

One important feature of this theorem is that the last statement provides quantitative bounds on the map $\boldsymbol{\Pi} \mapsto \boldsymbol{\Pi}^g$ which show that it can be extended to a continuous action of $\mathfrak{R}$ onto the space $\mathscr{M}$ of all admissible models. A crucial property of $\mathfrak{R}$ is that it is sufficiently large to allow us to "recenter" models in a natural way.

Definition 15.23. Let ξ be a (smooth) stationary stochastic process and let $\boldsymbol{\Pi}$ be its canonical lift. Then, there exists a unique character $g^{\text{BPHZ}} \in \mathfrak{R}$ such that $\boldsymbol{\Pi}^{\text{BPHZ}} = \boldsymbol{\Pi} M_{g^{\text{BPHZ}}}$ satisfies $\mathbf{E}(\boldsymbol{\Pi}^{\text{BPHZ}}\tau)(0) = 0$ for every canonical basis vector $\tau \in T$ with $\deg \tau < 0$. We call $\boldsymbol{\Pi}^{\text{BPHZ}}$ the *BPHZ lift* of ξ.

Remark 15.24. This is named after Bogoliubow, Parasiuk, Hepp and Zimmermann [BP57, Hep69, Zim69] who introduced an analogous renormalisation procedure in the context of perturbative quantum field theory in the sixties.

Remark 15.25. Note also that while the BPHZ lift of a noise ξ is "canonical", it does depend on the choice of kernel K for our notion of admissibility. In particular, different truncations of the heat kernel will in general lead to different values for the BPHZ renormalisation constants.

A beautiful property of the BPHZ lift is that it is much more stable than the canonical lift. Indeed, it was shown in [CH16] that one can introduce a natural measure of the "size" $N(\xi)$ of a stationary noise ξ which is such that for any sequence ξ_n such that $\sup_n N(\xi_n) < \infty$ and $\xi_n \to \xi$ in probability as random distributions, the corresponding BPHZ lifts $\boldsymbol{\Pi}_n^{\text{BPHZ}}$ converge to a limiting model $\boldsymbol{\Pi}^{\text{BPHZ}}$. This limiting model is furthermore independent of the choice of approximating sequence.

15.5.2 The renormalised equations

As introduced, the renormalisation group $\mathfrak{R}$ for KPZ is a Lie group of dimension 8, equal to the number of symbols ([tree], [tree], [tree], [tree], [tree], [tree], [tree], [tree]) used to generate T^-. As already hinted in Example 15.15 above, we will not need to renormalise planted trees, nor the noise symbol itself, nor symbols with three leaves (cubic in Gaussian noise, hence of zero mean, so that the BPHZ condition is trivially satisfied). We thus define a character g on T^- by specifying

$$g([\text{tree}]) = C_0, \qquad g([\text{tree}]) = C_1, \qquad g([\text{tree}]) = C_2, \qquad g([\text{tree}]) = C_3\,, \tag{15.30}$$

and set to vanish on the remaining symbols which require no renormalisation. The resulting renormalisation maps $M : T \to T$ is then given by $M := (g \otimes \mathrm{Id})\Delta^-$. (It turns out that we only need a three-parameter subgroup of $\mathfrak{R}$ to renormalise the equation, but in order to explain the procedure we prefer to work with the larger 4-dimensional subgroup of $\mathfrak{R}$.) It is now rather straightforward to show the following:

Proposition 15.26. *Let $M := (g \otimes \mathrm{Id})\Delta^-$ with g as specified in (15.30) and let $(\Pi^M, \Gamma^M) = M\mathscr{L}\xi$, where $\mathscr{L}\xi$ is the canonical lift of some smooth function ξ. Let furthermore H be the solution to (15.8) with respect to the model (Π^M, Γ^M). Then, writing $\mathcal{R}^M$ for the reconstruction operator associated to this renormalised model, the function $h(t,x) = \big(\mathcal{R}^M H\big)(t,x)$ solves the equation*

$$\partial_t h = \partial_x^2 h + (\partial_x h)^2 - 4C_0\, \partial_x h + \xi - (C_1 + C_2 + 4C_3)\,.$$

Proof. By Theorem 14.5, it turns out that (15.8) can be solved in $\mathscr{D}^\gamma$ as soon as γ is a little bit greater than $3/2$. Therefore, we only need to keep track of its solution H up to terms of degree $3/2$. By repeatedly applying the identity (15.9), we see that the solution $H \in \mathscr{D}^\gamma$ for γ close enough to $3/2$ is necessarily of the form

$$H = h\,\mathbf{1} + [\text{tree}] + [\text{tree}] + h'\,X_1 + 2[\text{tree}] + 2h'[\text{tree}]\,,$$

for some real-valued functions h and h'. (Note that h' is treated as an independent function here, we certainly do not suggest that the function h is differentiable! Our notation is only by analogy with the classical Taylor expansion.) As an immediate consequence, ∂H is given by

$$\partial H = [\text{tree}] + [\text{tree}] + h'\,\mathbf{1} + 2[\text{tree}] + 2h'[\text{tree}]\,, \tag{15.31}$$

as an element of $\mathscr{D}^\gamma$ for γ sufficiently close to $1/2$. Similarly, the right-hand side of the equation is given up to order 0 by

$$(\partial H)^2 + \Xi = \Xi + [\text{tree}] + 2[\text{tree}] + 2h'[\text{tree}] + [\text{tree}] + 4[\text{tree}] + 2h'[\text{tree}] + 4h'[\text{tree}] + (h')^2\,\mathbf{1}\,. \tag{15.32}$$

It follows from the definition of M that one then has the identity

$$M\partial H = \partial H - 4C_0[\text{tree}]\,,$$

so that, as an element of $\mathscr{D}^\gamma$ with very small (but positive) γ, one has the identity

$$(M\partial H)^2 = (\partial H)^2 - 8C_0$$  .

As a consequence, after neglecting all terms of strictly positive order, one has the identity (writing c instead of $c\mathbf{1}$ for real constants c)

$M\big((\partial H)^2 + \Xi\big) = (\partial H)^2 + \Xi - C_0\big(4$  $+ 4$  $+ 8$  $+ 4h'\,\mathbf{1}\big) - C_1 - C_2 - 4C_3$

$$= (M\partial H)^2 + \Xi - 4C_0\, M\partial H - (C_1 + C_2 + 4C_3)\,.$$

Combining this with the fact that M and ∂ commute, the claim now follows at once. □

Remark 15.27. It turns out that, thanks to the symmetry $x \mapsto -x$ enjoyed by our problem, the corresponding model can be renormalised by a map M as above, but with $C_0 = 0$. The reason why we considered the general case here is twofold. First, it shows that it is possible to obtain renormalised equations that differ from the original equation in a more complicated way than just by the addition of a large constant. Second, if one tries to approximate the KPZ equation by a microscopic model which is not symmetric under space inversion, then the constant C_0 plays a non-trivial role, see for example [HS17].

15.5.3 Convergence of the renormalised models

It remains to argue why one expects to be able to find constants C_i^ε such that the sequence of renormalised models $M^\varepsilon \mathscr{L}\xi_\varepsilon$ with $M^\varepsilon = \exp(\sum_{i=1}^3 C_i^\varepsilon L_i)$ converges to a limiting model. Instead of considering the actual sequence of models, we only consider the sequence of stationary processes $\hat{\boldsymbol{\Pi}}^\varepsilon \tau := \boldsymbol{\Pi}^\varepsilon M^\varepsilon \tau$, where $\boldsymbol{\Pi}^\varepsilon$ is associated to $(\Pi^\varepsilon, \Gamma^\varepsilon) = \mathscr{L}\xi_\varepsilon$ as in Section 15.5.1.

Remark 15.28. It is important to note that we do *not* attempt here to give a full proof that the renormalised model converges to a limit in the correct topology for the space of admissible models. We only aim to argue that it is *plausible* that $\hat{\boldsymbol{\Pi}}^\varepsilon$ converges to a limit in *some* topology. A full proof of convergence (but in a slightly different setting) can be found in [Hai13], see also [Hai14b, Section 10] and [CH16] for most general statements.

Since there are general arguments available to deal with all the expressions τ of positive degree as well as expressions of the type $\mathcal{I}'(\tau)$ and Ξ itself, we restrict ourselves to those that remain. Inspecting (15.11), we see that they are given by

For this part, some elementary notions from the theory of Wiener chaos expansions are required, but we'll try to hide this as much as possible. At a formal level, one has the identity

$$\boldsymbol{\Pi}^{\varepsilon}\,⫯ = K' * \xi_\varepsilon = K'_\varepsilon * \xi\;,$$

where the kernel K'_ε is given by $K'_\varepsilon = K' * \delta_\varepsilon$. This shows that, at least formally, one has

$$\big(\boldsymbol{\Pi}^{\varepsilon}\vee\big)(z) = \big(K' * \xi_\varepsilon\big)(z)^2 = \iint K'_\varepsilon(z-z_1)K'_\varepsilon(z-z_2)\,\xi(z_1)\xi(z_2)\,dz_1\,dz_2\;.$$

Similar but more complicated expressions can be found for any formal expression τ. This naturally leads to the study of random variables of the type

$$I_k(f) = \int\cdots\int f(z_1,\ldots,z_k)\,\xi(z_1)\cdots\xi(z_k)\,dz_1\cdots dz_k\;. \tag{15.33}$$

Ideally, one would hope to have an Itô isometry of the type $\mathbf{E} I_k(f)I_k(g) = \langle f^{\mathrm{sym}}, g^{\mathrm{sym}}\rangle$, where $\langle\cdot,\cdot\rangle$ denotes the L^2-scalar product and f^{sym} denotes the symmetrisation of f. This is unfortunately *not* the case. Instead, one should replace the products in (15.33) by *Wick products*, which are formally generated by all possible *contractions* of the type

$$\xi(z_i)\xi(z_j) \mapsto \xi(z_i)\diamond\xi(z_j) + \delta(z_i - z_j)\;.$$

If we then set

$$\hat I_k(f) = \int\cdots\int f(z_1,\ldots,z_k)\,\xi(z_1)\diamond\cdots\diamond\xi(z_k)\,dz_1\cdots dz_k\;,$$

One has indeed

$$\mathbf{E}\hat I_k(f)\hat I_k(g) = \langle f^{\mathrm{sym}}, g^{\mathrm{sym}}\rangle\;.$$

Furthermore, one has equivalence of moments in the sense that, for every $k > 0$ and $p > 0$ there exists a constant $C_{k,p}$ such that

$$\mathbf{E}|\hat I_k(f)|^p \le C_{k,p}\|f^{\mathrm{sym}}\|^p\;.$$

Finally, one has $\mathbf{E}\hat I_k(f)\hat I_\ell(g) = 0$ if $k \neq \ell$. Random variables of the form $\hat I_k(f)$ for some $k \ge 0$ and some square integrable function f are said to belong to the *kth homogeneous Wiener chaos*.

Returning to our problem, we first argue that it should be possible to choose M^ε in such a way that $\hat{\boldsymbol{\Pi}}^{\varepsilon}\vee$ converges to a limit as $\varepsilon \to 0$. The above considerations suggest that one should rewrite $\boldsymbol{\Pi}^{\varepsilon}\vee$ as

$$\begin{aligned}\big(\boldsymbol{\Pi}^{\varepsilon}\vee\big)(z) &= \big(K' * \xi_\varepsilon\big)(z)^2 \\ &= \iint K'_\varepsilon(z-z_1)K'_\varepsilon(z-z_2)\,\xi(z_1)\diamond\xi(z_2)\,dz_1\,dz_2 + C_\varepsilon^{(1)}\;,\end{aligned} \tag{15.34}$$

where the constant $C_\varepsilon^{(1)}$ is given by the contraction

$$C_\varepsilon^{(1)} = \RSVc \stackrel{\text{def}}{=} \int \bigl(K'_\varepsilon(z)\bigr)^2 \, dz \; .$$

Note now that K'_ε is an ε-approximation of the kernel K' which has the same singular behaviour as the derivative of the heat kernel. In terms of the parabolic distance, the singularity of the derivative of the heat kernel scales like $K(z) \sim |z|^{-2}$ for $z \to 0$. (Recall that we consider the parabolic distance $|(t,x)| = \sqrt{|t|} + |x|$, so that this is consistent with the fact that the derivative of the heat kernel is bounded by t^{-1}.) This suggests that one has $\bigl(K'_\varepsilon(z)\bigr)^2 \sim |z|^{-4}$ for $|z| \gg \varepsilon$. Since parabolic space-time has scaling dimension 3 (time counts double!), this is a non-integrable singularity. As a matter of fact, there is a whole power of z missing to make it borderline integrable, which suggests that one has

$$C_\varepsilon^{(1)} \sim \frac{1}{\varepsilon} \; .$$

This already shows that one should not expect $\boldsymbol{\Pi}^\varepsilon \RSV$ to converge to a limit as $\varepsilon \to 0$. However, it turns out that the first term in (15.34) converges to a distribution-valued stationary space-time process, so that one would like to somehow get rid of this diverging constant $C_\varepsilon^{(1)}$. This is exactly where the renormalisation map M^ε (in particular the factor $\exp(-C_1 L_1)$) enters into play. Following the above definitions, we see that one has

$$\bigl(\hat{\boldsymbol{\Pi}}^\varepsilon \RSV\bigr)(z) = \bigl(\boldsymbol{\Pi}^\varepsilon M \RSV\bigr)(z) = \bigl(\boldsymbol{\Pi}^\varepsilon \RSV\bigr)(z) - C_1 \; .$$

This suggests that if we make the choice $C_1 = C_\varepsilon^{(1)}$, then $\hat{\boldsymbol{\Pi}}^\varepsilon \RSV$ does indeed converge to a non-trivial limit as $\varepsilon \to 0$. This limit is a distribution given, at least formally, by

$$\bigl(\boldsymbol{\Pi}^\varepsilon \RSV\bigr)(\psi) = \iint \psi(z) K'(z - z_1) K'(z - z_2) \, dz \, \xi(z_1) \diamond \xi(z_2) \, dz_1 \, dz_2 \; .$$

Using again the scaling properties of the kernel K', it is not too difficult to show that this yields indeed a random variable belonging to the second homogeneous Wiener chaos for every choice of smooth test function ψ.

The case $\tau = \RSW$ is treated in a somewhat similar way. This time one has

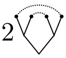
$$\begin{aligned}\bigl(\boldsymbol{\Pi}^\varepsilon \RSW\bigr)(z) &= \bigl(K' * \xi_\varepsilon\bigr)(z) \bigl(K' * K' * \xi_\varepsilon\bigr)(z) \\ &= \iint K'_\varepsilon(z - z_1)(K * K'_\varepsilon)(z - z_2) \, \xi(z_1) \diamond \xi(z_2) \, dz_1 \, dz_2 + C_\varepsilon^{(0)} \; ,\end{aligned}$$

where the constant $C_\varepsilon^{(0)}$ is given by the contraction

$$C_\varepsilon^{(0)} = \RSWc \stackrel{\text{def}}{=} \int K'_\varepsilon(z) \bigl(K' * K'_\varepsilon\bigr)(z) \, dz \; .$$

This time however K'_ε is an odd function (in the spatial variable) and $K' * K'_\varepsilon$ is an even function, so that $C_\varepsilon^{(0)}$ vanishes for every $\varepsilon > 0$. This is why we can set $C_0 = 0$ and no renormalisation is required for $\RSW$.

Turning to our list of terms of negative degree, it remains to consider , , and . It turns out that the latter two are the more difficult ones, so we only discuss these. Let us first argue why we expect to be able to choose the constant C_2 in such a way that $\hat{\boldsymbol{\Pi}}^{\varepsilon}$ converges to a limit. In this case, the "bad" term comes from the part of $\big(\boldsymbol{\Pi}^{\varepsilon}$$\big)(z)$ belonging to the homogeneous chaos of order 0. This is simply a constant, which is given by

$$C_\varepsilon^{(2)} = 2$$  $$\stackrel{\text{def}}{=} 2\int K'(z)K'(\bar z)Q_\varepsilon^2(z-\bar z)\,dz\,d\bar z\;, \tag{15.35}$$

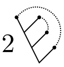
where the kernel Q_ε is given by

$$Q_\varepsilon(z) = \int K_\varepsilon'(\bar z)K_\varepsilon'(\bar z - z)\,d\bar z\;.$$

Remark 15.29. The factor 2 comes from the fact that the contraction (15.35) appears twice, since it is equal to the contraction . In principle, one would think that the contraction  also contributes to $C_\varepsilon^{(2)}$. This term however vanishes due to the fact that the integral of K_ε' vanishes.

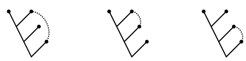
Since K_ε' is an ε-mollification of a kernel with a singularity of order -2 and the scaling dimension of the underlying space is 3, we see that Q_ε behaves like an ε-mollification of a kernel with a singularity of order $-2-2+3=-1$ at the origin. As a consequence, the singularity of the integrand in (15.35) is of order -6, which gives rise to a logarithmic divergence as $\varepsilon \to 0$. This suggests that one should choose $C_2 = C_\varepsilon^{(2)}$ in order to cancel out this diverging term and obtain a non-trivial limit for $\hat{\boldsymbol{\Pi}}^{\varepsilon}$ as $\varepsilon \to 0$. This is indeed the case.

We finally turn to the case $\tau =$ . In this case, there are "bad" terms appearing in the Wiener chaos decomposition of $\boldsymbol{\Pi}^{\varepsilon}$ both in the second and the zeroth Wiener chaos. This time, the constant appearing in the zeroth Wiener chaos is given by

$$C_\varepsilon^{(3)} = 2$$  $$\stackrel{\text{def}}{=} 2\int K'(z)K'(\bar z)Q_\varepsilon(\bar z)Q_\varepsilon(z+\bar z)\,dz\,d\bar z\;,$$

which diverges logarithmically for exactly the same reason as $C_\varepsilon^{(2)}$. Setting $C_2 = C_\varepsilon^{(2)}$, this diverging constant can again be cancelled out. The combinatorial factor 2 arises in essentially the same way as for  and the contribution of the term where the two top nodes are contracted vanishes for the same reason as previously.

It remains to consider the contribution of $\boldsymbol{\Pi}^{\varepsilon}$ to the second Wiener chaos. This contribution consists of three terms, which correspond to the contractions



It turns out that the first one of these terms does not give raise to any singularity. The last two terms can be treated in essentially the same way, so we focus on the last one, which we denote by η^ε. For fixed ε, the distribution (actually smooth function) η^ε is given by

$$\eta^\varepsilon(\psi) = \int \psi(z_0)K'(z_0 - z_1)Q_\varepsilon(z_0 - z_1)K'(z_2 - z_1) \\ \times K'_\varepsilon(z_3 - z_2)K'_\varepsilon(z_4 - z_2)\,\xi(z_3)\diamond\xi(z_4)\,dz\;.$$

The problem with this is that as $\varepsilon \to 0$, the product $\hat{Q}_\varepsilon := K'Q_\varepsilon$ converges to a kernel $\hat{Q} = K'Q$, which has a non-integrable singularity at the origin. In particular, it is not clear *a priori* whether the action of integrating a test function against $\hat{Q}_\varepsilon$ converges to a limiting distribution as $\varepsilon \to 0$. Our saving grace here is that since Q_ε is even and K' is odd, the kernel $\hat{Q}_\varepsilon$ integrates to 0 for every fixed ε.

This is akin to the problem of making sense of the "Cauchy principal value" distribution, which formally corresponds to the integration against $1/x$. For the sake of the argument, let us consider a function $W\colon \mathbf{R} \to \mathbf{R}$ which is compactly supported and smooth everywhere except at the origin, where it diverges like $|W(x)| \sim 1/|x|$. It is then natural to associate to W a "renormalised" distribution $\mathscr{R}W$ given by

$$\big(\mathscr{R}W\big)(\varphi) = \int W(x)\big(\varphi(x) - \varphi(0)\big)\,dx\;.$$

Note that $\mathscr{R}W$ has the property that if $\varphi(0) = 0$, then it simply corresponds to integration against W, which is the standard way of associating a distribution to a function. Furthermore, the above expression is always well-defined, since φ is smooth and therefore the factor $(\varphi(x) - \varphi(0))$ cancels out the singularity of W at the origin. It is also straightforward to verify that if W_ε is a sequence of smooth approximations to W (say one has $W_\varepsilon(x) = W(x)$ for $|x| > \varepsilon$ and $|W_\varepsilon| \lesssim 1/\varepsilon$ otherwise) which has the property that each W_ε integrates to 0, then $W^\varepsilon \to \mathscr{R}W$ in a distributional sense.

In the same way, one can show that $\hat{Q}_\varepsilon$ converges as $\varepsilon \to 0$ to a limiting distribution $\mathscr{R}\hat{Q}$. As a consequence, one can show that η^ε converges to a limiting (random) distribution η given by

$$\eta(\psi) = \int \psi(z_0)\,\mathscr{R}\hat{Q}(z_0-z_1)K'(z_2-z_1)K'(z_3-z_2)K'(z_4-z_2)\,\xi(z_3)\diamond\xi(z_4)\,dz\;.$$

It should be clear from this whole discussion that while the precise values of the constants C_i depend on the details of the mollifier δ_ε, the limiting (random) model $(\hat{\Pi}, \hat{\Gamma})$ obtained in this way is independent of it. Combining this with the continuity of the solution to the fixed point map (15.8) and of the reconstruction operator $\mathcal{R}$ with respect to the underlying model, we see that the statement of Theorem 15.2 follows almost immediately.

15.6 The KPZ equation and rough paths

In the particular case of the KPZ equation, it turns out that is possible to give a robust solution theory by only using "classical" controlled rough path theory, as exposed in the earlier part of this book. This is actually how it was originally treated in [Hai13]. To see how this can be the case, we make the following crucial remarks:

1. First, looking at the expression (15.31) for ∂H, we see that most symbols come with constant coefficients. The only non-constant coefficients that appear are in front of the term $\mathbf{1}$, which is some kind of renormalised value for ∂H, and in front of the term ⟨. This suggests that the problem of finding a solution h to the KPZ equation (or equivalently a solution h' to the corresponding Burgers' equation) can be simplified considerably by considering instead the function v given by
$$v = \partial_x h - \boldsymbol{\Pi}\big(⫯ + Y + 2\,[\text{tree}]\big)\;, \tag{15.36}$$
where $\boldsymbol{\Pi}$ is the operator given by (15.12–15.14).
2. The only symbol τ appearing in ∂H such that $\deg\tau + \deg ⟨ < 0$ is the symbol ⫯. Furthermore, one has
$$\Delta\mathbf{1} = \mathbf{1}\otimes\mathbf{1}\;,\qquad \Delta ⟨ = ⟨\otimes\mathbf{1} + \mathbf{1}\otimes\mathcal{J}'(⫯)\;,$$
$$\Delta ⫯ = ⫯\otimes\mathbf{1}\;,\qquad \Delta ◇ = ◇\otimes\mathbf{1} + ⫯\otimes\mathcal{J}'(⫯)\;.$$
It then follows from this and the definition (15.16) of the structure group G that the space $\langle ⫯, ◇, \mathbf{1}, ⟨\rangle \subset T$ is invariant under the action of G. Furthermore, its action on this subspace is completely described by one real number corresponding to $\mathcal{J}'(⫯)$. Finally, viewing this subspace as a regularity structure in its own right, we see that it is nothing but the regularity structure of Section 13.3.2, provided that we make the identifications ⫯ $\sim \dot W$, ⟨ $\sim W$, and ◇ $\sim \dot{\mathbb{W}}$.
3. One has the identities
$$\Delta[\text{tree}] = [\text{tree}]\otimes\mathbf{1} + ⫯\otimes\mathcal{J}'([\text{tree}])\;,\qquad \Delta[\text{tree}'] = [\text{tree}']\otimes\mathbf{1} + ⫯\otimes\mathcal{J}'([\text{tree}])\;,$$
so that the pair of symbols $\{[\text{tree}], [\text{tree}']\}$ could also have played the role of $\{W, \dot{\mathbb{W}}\}$ in the previous remark.

Let now ξ be a *smooth* function and let h be given by the solution to the unrenormalised KPZ equation (15.1). Defining $\boldsymbol{\Pi}$ by $\boldsymbol{\Pi}\Xi = \xi$ and then recursively as in (15.13), and defining v by (15.36), we then obtain for v the equation
$$\partial_t v = \partial_x^2 v + \partial_x\big(v\,\boldsymbol{\Pi}⫯ + 4\,\boldsymbol{\Pi}[\text{tree}']\big) + R\;, \tag{15.37}$$
where the "remainder" R belongs to $\mathcal{C}^\alpha$ for every $\alpha < -1$. Similarly to before, it also turns out that if we replace $\boldsymbol{\Pi}$ bi $\hat{\boldsymbol{\Pi}} = \boldsymbol{\Pi}^M$ defined as in (15.24) (with $C_0 = 0$) and h as the solution to the renormalised KPZ equation (15.6) with $C_\varepsilon = C_1 + C_2 + 4C_3$, then v also satisfies (15.37), but with $\boldsymbol{\Pi}$ replaced by the renormalised model $\hat{\boldsymbol{\Pi}}$.

We are now in the following situation. As a consequence of (15.31) we can *guess* that for any fixed time t, the solution v should be controlled by the function $\hat{\boldsymbol{\Pi}}$[tree], which we can interpret as one component (say W^1) of some rough path $(W, \mathbb{W})$. Note that here the spatial variable plays the role of time! The time variables merely plays the role of a parameter, so we really have a family of rough paths indexed by time. Furthermore, $\hat{\boldsymbol{\Pi}}$[tree] can be interpreted as the distributional derivative of another component (say W^0) of the rough path W. Finally, the function $\hat{\boldsymbol{\Pi}}$[tree] can be interpreted as a third component W^2 of W.

As a consequence of the second and third remarks above, the two distributions $\hat{\boldsymbol{\Pi}}$[tree] and $\hat{\boldsymbol{\Pi}}$[tree] can then be interpreted as the distributional derivatives of the "iterated integrals" $\mathbb{W}^{1,0}$ and $\mathbb{W}^{2,1}$. It follows automatically from these algebraic relations combined with the analytic bounds (13.13) that $\mathbb{W}^{1,0}$ and $\mathbb{W}^{2,1}$ then satisfy the required estimates (2.3). Our model does not provide any values for $\mathbb{W}^{1,2}$, but these turn out not to be required. Assuming that v is indeed controlled by $X_1 = \hat{\boldsymbol{\Pi}}$[tree], it is then possible to give meaning to the term $v\,\boldsymbol{\Pi}$[tree] appearing in (15.37) by using "classical" rough integration.

As a consequence, we then see that the right-hand side of (15.37) is of the form $\partial_x^2 Y$, for some function Y controlled by W^0. One of the main technical results of [Hai13] guarantees that if Z solves

$$\partial_t Z = \partial_x^2 Z + \partial_x^2 Y ,$$

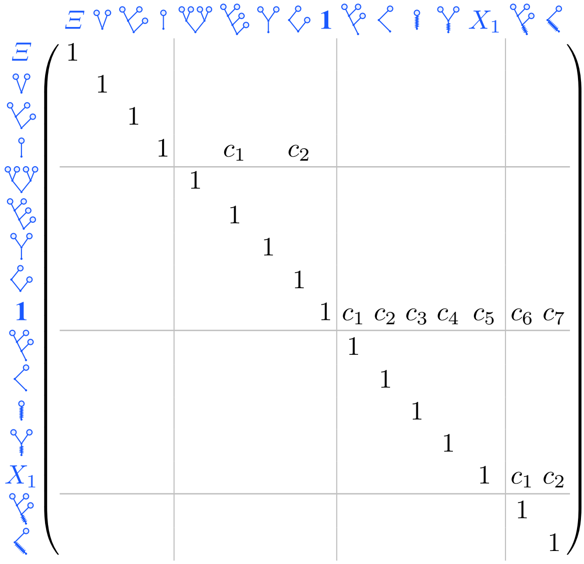
and Y is controlled by W^0, then Z is necessarily controlled by $W^1 = \hat{\Pi}$[tree]. This "closes the loop" and allows to set up a fixed point equation for v that is stable as a function of the underlying model $\hat{\boldsymbol{\Pi}}$ and therefore also allows to deal with the limiting case of the KPZ equation driven by space-time white noise.

15.7 Exercises

Exercise 15.1 (KPZ Structure Group) *Consider the* 16*-dimensional KPZ regularity structure with* $T = T_{\mathrm{KPZ}}$ *given by*

$$T = \langle\, \Xi, [tree], [tree], [tree], [tree], [tree], [tree], [tree], \mathbf{1}, [tree], [tree], [tree], [tree], X_1, [tree], [tree] \,\rangle\,.$$

Show that the structure group G *is a* 7*-dimensional (non-commutative) Lie group, an element* $\Gamma \in G \subset \mathcal{L}(T,T)$ *of which has the upper triangular matrix representation*

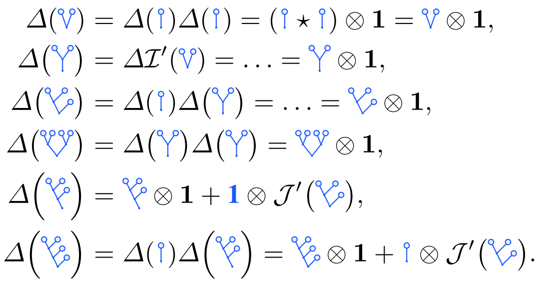


where empty entries mean zeros. Note that the upper-triangular form reflects the fact that $\Gamma - \mathrm{Id}$ *is only allowed to produce lower order terms. (Remark: It is immediate from this representation that* $\langle$[tree], [tree], $\mathbf{1}$, [tree]$\rangle$ *and* $\langle$[tree], [tree], $\mathbf{1}$, [tree]$\rangle$ *are indeed sectors, with "rough path" index set* $\{-\frac{1}{2}^-, 0^-, 0, \frac{1}{2}^-\}$*, and action of the structure group exactly as in the rough path case (13.12) (with "h" replaced by* c_1 *and* c_2*, respectively.)*

Solution. We first derive the coaction on all the symbols, and here prefer to write Δ for the coaction and keep Δ^+ for the coproduct on T^+. By definition of the coaction, $\Delta(\Xi) = \Xi \otimes \mathbf{1}$ and

$$\Delta(\mathcal{I}'(\Xi)) = \mathcal{I}'(\Xi) \otimes \mathbf{1} + \sum_{k \in \mathbf{N}^2} \frac{X^k}{k!} \otimes \mathcal{J}'_k(\Xi) = \mathcal{I}'(\Xi) \otimes \mathbf{1}\,,$$

since $\deg \mathcal{J}'_k(\Xi) = \deg \mathcal{J}_{k+(0,1)}(\Xi) = \deg \Xi + 1 - |k| < 0$ so that $\mathcal{J}'_k(\Xi) = 0$. Similarly, write Δ instead of Δ^+ for better readability,



Note the interpretation of *cutting off positive branches*: $\deg \mathcal{J}'(\text{tree}) = 1+3(-\frac{3}{2}^-)+4 = \frac{1}{2}^- > 0$, and also $\deg \mathcal{J}'(\text{tree}) = \frac{1}{2}^-$ as seen in

$$\begin{aligned}\Delta(\text{tree}) &= \text{tree} \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{J}'(\text{tree}),\\ \Delta(\text{tree}) &= \Delta(\text{tree})\Delta(\text{tree}) = \text{tree} \otimes \mathbf{1} + \text{tree} \otimes \mathcal{J}'(\text{tree}).\end{aligned}$$

To deal with $\text{tree} = \mathcal{I}(\Xi)$, note $\deg \mathcal{J}(\Xi) > 0$, $\deg \mathcal{J}'(\Xi) < 0$ so that the latter term does not figure (same reasoning for $\text{tree} = \mathcal{I}(\text{tree})$), and obtain

$$\begin{aligned}\Delta(\text{tree}) &= \text{tree} \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{J}(\Xi),\\ \Delta(\text{tree}) &= \text{tree} \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{J}(\text{tree}).\end{aligned}$$

By definition, $\Delta X_1 = X_1 \otimes \mathbf{1} + \mathbf{1} \otimes X_1$. Next consider tree and tree. In view of $|\mathcal{J}(\text{tree})|$ and $|\mathcal{J}'(\text{tree})| > 0$ we have (same reasoning for tree),

$$\begin{aligned}\Delta(\text{tree}) &= \text{tree} \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{J}(\text{tree}) + X_1 \otimes \mathcal{J}'(\text{tree}),\\ \Delta(\text{tree}) &= \text{tree} \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{J}(\text{tree}) + X_1 \otimes \mathcal{J}'(\text{tree}),\end{aligned}$$

Inspecting the above reveals that we need $\mathbf{1}$ and then the following 7 "positive" symbols (also viewable as trees) in T^+,

$$\mathcal{J}'(\text{tree}), \mathcal{J}'(\text{tree}), \mathcal{J}(\Xi), \mathcal{J}(\text{tree}), X_1, \mathcal{J}(\text{tree}), \mathcal{J}(\text{tree}), \tag{15.38}$$

of resp. homogeneities $\frac{1}{2}^-, \frac{1}{2}^-, \frac{1}{2}^-, 1^-, 1, \frac{3}{2}^-, \frac{3}{2}^-$. On the other hand, T^+ was introduced abstractly as free commutative algebra generated by all of the above symbols (with unit element 1). Even upon truncation, say $T^+ = T^+_{<3/2}$ with abusive notation, this leaves us with $10+4+1 = 15$ generating symbols,

$$\mathcal{J}'(\text{tree}), \mathcal{J}'(\text{tree}), \ldots, \mathcal{J}'(\text{tree}); \mathcal{J}(\Xi), \ldots, \mathcal{J}(\text{tree}); X_1 \tag{15.39}$$

(of which only 7 are needed). Of course, T^+ also contains (free) products such as $\mathcal{J}'(\text{tree})\mathcal{J}'(\text{tree})$, $X_1\mathcal{J}'(\text{tree})$, $\mathcal{J}'(\text{tree})\mathcal{J}(\text{tree},)$ (all of degree $< 3/2$), however by working in T these did not appear as "right-hand side"-image of Δ above.

Consider now a *character* of the algebra T^+; that is, an element $g \in (T^+)^*$, so that $g(1) = 1$ and $g(\sigma\bar{\sigma}) = g(\sigma)g(\bar{\sigma})$. (Actually, in view of the truncation we impose this only for $\sigma, \bar{\sigma}$ with $\deg(\sigma\bar{\sigma}) = \deg \sigma + \deg \bar{\sigma} < 3/2$.) Such g is obviously determined by its value on each of the 15 basis symbols listed in (15.39). Now T^+ can be given a Hopf structure, with coproduct Δ^+ and antipode, so that the set of characters forms the group G^+, with product given by

$$(f \circ g)(\sigma) = (f \otimes g)\Delta^+\sigma = \sum_{(\sigma)} \langle f, \sigma' \rangle \langle g, \sigma'' \rangle;$$

inverses are given in terms of the antipode. One thus sees that G^+ is a 15-dimensional (Lie) group. However, only a 7-dimensiorial subgroup is needed, for we only care about the 7 values arising from (15.38), which we call

$$c_1 = g\big(\mathcal{J}'(\tau)\big), \ldots, c_7 = g(\mathcal{J}(\tau)).$$

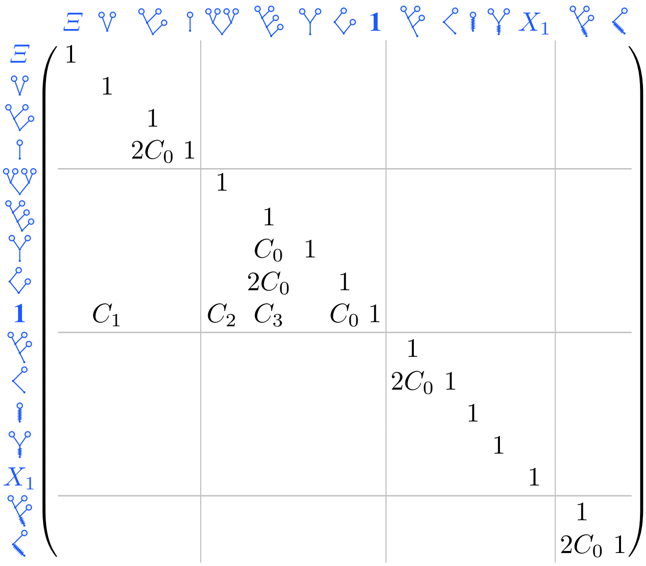
It then follows from $\Gamma_g := (\mathrm{Id} \otimes g)\Delta$ that $\Gamma_g : T \to T$ acts as identity on all symbols other than

$$\begin{aligned}
\Gamma_g\big(\tau\big) &= \tau + g(\mathcal{J}'(\tau))\mathbf{1} \equiv \tau + c_1\mathbf{1},\\
\Gamma_g\big(\tau\big) &= \tau + g(\mathcal{J}'(\tau))\tau = \tau + c_1\tau,\\
\Gamma_g(\tau) &= \tau + g(\mathcal{J}'(\tau))\mathbf{1} = \tau + c_2\mathbf{1},\\
\Gamma_g(\tau) &= \tau + c_2\tau,\\
\Gamma_g(\tau) &= \tau + g(\mathcal{J}(\Xi))1 = \tau + c_3\mathbf{1},\\
\Gamma_g(\tau) &= \tau + g(\mathcal{J}(\tau))\mathbf{1} = \tau + c_4\mathbf{1},\\
\Gamma_g(X_1) &= X_1 + c_5\mathbf{1},\\
\Gamma_g(\tau) &= \tau + c_6\mathbf{1} + c_1X_1,\\
\Gamma_g(\tau) &= \tau + c_7\mathbf{1} + c_2X_1.
\end{aligned}$$

The matrix representation of Γ_g is then immediate.

Exercise 15.2 (KPZ Renormalisation Group) *Consider again the* 16*-dimensional KPZ regularity structure with structure space $T = T_{\mathrm{KPZ}}$. The renormalisation group was given as subgroup $\mathfrak{R} \subset \mathcal{L}(T,T)$, given by $M_g\tau = (g \otimes \mathrm{Id})\Delta^-\tau$, where g ranges over the characters of T^-. Consider more specifically $M = M_g$ with g as specified in (15.30), i.e. $g(\tau) = C_0, g(\tau) = C_1, g(\tau) = C_2, g(\tau) = C_3$ and set to vanish on the remaining symbols.*

Show that this gives a subgroup of $\mathfrak{R}$ which is a 4*-dimensional (commutative) Lie group, an element $M \in \mathfrak{R} \subset \mathcal{L}(T,T)$ of which has lower triangular matrix representation*



Exercise 15.3 *Show that the two procedures for recovering* $\boldsymbol{\Pi}$ *from the knowledge of* (Π, Γ) *outlined in Remark 15.13 and on page 301 are equivalent.*

15.8 Comments

The original proof [Hai13] of well-posedness of the KPZ equation without using the Cole–Hopf transform did not use regularity structures but instead viewed the solution at any fixed time as a spatial rough path controlled by the solution to the linearised equation, in the spirit of Section 12.3. An alternative approach using paracontrolled distributions as developed in [GIP15] was used in [GP17] to obtain a number of additional properties of the solutions, including a clean variational formulation.

Given that the KPZ equation is expected to enjoy a form of "universality", a very natural question is that of showing that "most" classes of interface fluctuation models converge to in in the weakly asymmetric regime. The first result in this direction was obtained by Bertini–Giacomin [BG97], but this relied crucially on a microscopic version of the Hopf–Cole transform to show that the transformed particle system converges to the multiplicative stochastic heat equation. A first more general result was obtained by Jara–Conçalves [GJ14] who showed that the large scale fluctuations of a large number of particle systems solve the KPZ equation in a relatively weak sense. It has been an open problem for quite some time now whether such a weak notion of solution characterises solutions to the KPZ uniquely. Major progress in this direction was obtained by Gubinelli–Perkowski [GP18] who showed that this is indeed the case at stationarity under an additional structural assumption on the generator of the particle system that can be verified for a number of systems of interest.

On the other hand, a large class of interface fluctuation models that fall outside of this approach is given by solutions to an equation of the type

$$\partial_t h_\varepsilon = \partial_x^2 h_\varepsilon + \sqrt{\varepsilon} F(\partial_x h_\varepsilon) + \eta(t,x) , \tag{15.40}$$

where η is a (smooth) space-time random field with sufficiently good mixing properties, $F\colon \mathbf{R} \to \mathbf{R}$ is an even function growing at infinity, and $\varepsilon > 0$ is a parameter controlling the asymmetry of the problem. Under rather weak assumptions on η and F one then expects to be able to find constants C_ε such that $\varepsilon^{-1/2} h_\varepsilon(\varepsilon^{-2}t, \varepsilon^{-1}x) - C_\varepsilon t$ converges to solutions to the KPZ equation. This was shown to be indeed the case in various special cases of increasing generality in [HS17, HQ18, HX19, FG19]. (The last reference treats a different class of models but its proofs could be adapted to the setting of (15.40).)

There is a natural generalisation of the KPZ equation going in a completely different direction. Indeed, given a Riemannian manifold $(\mathcal{M}, g)$ (where g denotes the metric tensor), we can ask ourselves what the natural "stochastic heat equation with values in $\mathcal{M}$" looks like. A moment's thought suggests that it should be given, in local coordinates, by an equation of the form

$$\partial_t u^\alpha = \partial_x^2 u^\alpha + \Gamma^\alpha_{\beta\gamma}(u)\, \partial_x u^\beta\, \partial_x u^\gamma + \sigma_i^\alpha(u)\, \xi_i , \tag{15.41}$$

where the ξ_i are i.i.d. space-time white noises, $\Gamma^\alpha_{\beta\gamma}$ are the Christoffel symbols for $\mathcal{M}$, the σ_i are any finite collection of vector fields such that

$$\sigma_i^\alpha \sigma_i^\beta = g , \tag{15.42}$$

and summation over repeated indices is implied. By combining the results of [CH16, BHZ19, BCCH17], it is not difficult to see that there are natural notions of solution to (15.41), but these are of course only well-defined modulo an element of the renormalisation group $\mathfrak{R}$. It turns out that in this case, even after taking into account simplifications due to the symmetry $x \leftrightarrow -x$ and the fact that the noises are i.i.d. Gaussian, the relevant subgroup of $\mathfrak{R}$ is generically (namely for large enough dimension of $\mathcal{M}$) of dimension 54.

This is a good example illustrating the role played by symmetries. In this particular case, there are two additional symmetries one would like to exploit. On the one hand, one would like to enforce equivariance under the group of diffeomorphisms of $\mathcal{M}$. In other words, solutions to (15.41) should be independent of the specific coordinate system used to write (15.41). This is akin to the property of solutions to regular SDEs written in *Stratonovich form* (or indeed those of RDEs driven by a geometric rough path). On the other hand, the derivation of (15.41) implicitly makes use of Itô's isometry to guarantee that, at least in law, its solutions do not depend on the specific choice of the vector fields satisfying (15.42). This in turn is akin to the property of solutions to SDEs written in *Itô form*. It turns out – and this is the main result of [BGHZ19] – that in this context it is possible to find solution theories that do satisfy both properties *simultaneously*! In fact there still exists a two-parameter family of them, but if we restrict ourselves to (15.1) (i.e. with Γ and σ related to the same metric g), then it reduces to a one-parameter family and the corresponding correction term (analogous to the Itô-Stratonovich correction term allowing to switch between solution theories for SDEs) is given by a multiple of the gradient of the scalar curvature of $\mathcal{M}$. This sheds new light on observations that had previously been made in a closely related context both in the physics [Che72, Um74] and in the mathematics [Dar84, IM85, AD99] literatures.