Regularity structures and the dynamical Φ^4_3 model

Martin Hairer

ABSTRACT. We give a concise overview of the theory of regularity structures as first exposed in [**Hai14**]. In order to allow to focus on the conceptual aspects of the theory, many proofs are omitted and statements are simplified. In order to provide both motivation and focus, we concentrate on the study of solutions to the stochastic quantisation equations for the Euclidean Φ^4_3 quantum field theory which can be obtained with the help of this theory. In particular, we sketch the proofs of how one can show that this model arises quite naturally as an idealised limiting object for several classes of smooth models.

Contents

1. Introduction

The purpose of these notes is to give a short informal introduction to the main concepts of the theory of regularity structures, focusing as an example on the construction and approximation of the dynamical Φ^4_3 model. They expand and complement the notes [**Hai15**] which focus more on the general theory and the ideas of the proofs of the main abstract results. Here, we instead focus mainly on the construction of the dynamical Φ^4_3 model, as well as on the way in which the abstract theory can be used to obtain a number of rather non-obvious approximation results for this model.

The theory of "regularity structures", introduced in [**Hai14**], unifies various flavours of the theory of (controlled) rough paths (including Gubinelli's theory of controlled rough paths [**Gub04**], as well as his branched rough paths [**Gub10**]), as well as the usual Taylor expansions. While it has its conceptual roots in Lyons's theory of rough paths [**Lyo98**], its main advantage is that it is no longer tied to the one-dimensionality of the time parameter, which makes it also suitable for the description of solutions to stochastic *partial* differential equations, rather than just stochastic ordinary differential equations. This broader scope requires a theory that admits more flexibility than the theory of rough paths, so we will see that the underlying algebraic structure (which in the case of the theory of rough paths is always given by the tensor algebra endowed with the concatenation and shuffle products) is problem-dependent and enters as a parameter of the theory. While the exposition of these notes aims to be mostly self-contained, none of the proofs will be given in detail, instead we will only sketch the main arguments.

The main achievement of the theory of regularity structures is that it allows to give a meaning, as well as a robust approximation theory, to ill-posed stochastic PDEs that arise naturally when trying to describe the crossover regime between two universality classes for various system of statistical mechanics. One example of such an equation is the KPZ equation arising as a natural model for one-dimensional interface motion in the crossover regime between the Edwards-Wilkinson and the KPZ universality classes [**KPZ86, BG97, Qua11, Hai13**]:

$$(\text{KPZ}) \qquad\qquad \partial_t h = \partial_x^2 h + (\partial_x h)^2 + \xi - C \ .$$

Another example is the dynamical Φ_3^4 model arising for example in the stochastic quantisation of Euclidean quantum field theory [**PW81, JLM85, AR91, DPD03, Hai14**]. However, it also arises as a description of the crossover regime for the dynamic of phase coexistence models near their critical point between the mean-field theory and the "Wilson-Fisher" renormalisation fixed point [**GLP99, MW14**]. This model can be written formally as

$$(\Phi^4) \qquad\qquad \partial_t \Phi = \Delta \Phi + C\Phi - \Phi^3 + \xi \ .$$

In both of these examples, ξ denotes space-time white noise, namely the generalised Gaussian random field such that

$$(1.1) \qquad\qquad \mathbf{E}\xi(t,x)\xi(s,y) = \delta(t-s)\delta(y-x) \ .$$

Furthermore, C is a constant (which will actually turn out to be infinite in some sense!), and we consider these equations on bounded tori. In the case of the dynamical Φ_3^4 model, the spatial variable has dimension 3, while it has dimension 1 in the case of the KPZ equation. Why are these equations problematic? As one can guess from (1.1), typical realisations of the noise ξ are not functions, but rather irregular space-time distributions. As a matter of fact, it follows immediately from (1.1) that the law of ξ is invariant under

the substitution $\xi(t, x) \mapsto \lambda^{\frac{d}{2}+1}\xi(\lambda^2 t, \lambda x)$, which correctly suggests that its samples typically belong to the parabolic $\mathcal{C}^\alpha$ spaces only for $\alpha < -1 - \frac{d}{2}$. (We write $\mathcal{C}^\alpha$ as a shorthand for the parabolic Besov space $B^\alpha_{\infty,\infty,\text{loc}}$, which coincides with the usual parabolic $\mathcal{C}^\alpha$ spaces for positive non-integer values of α. It is in some sense the largest space of distributions that is invariant under the scaling $\varphi(\cdot) \mapsto \lambda^{-\alpha}\varphi(\lambda^{-1}\cdot)$, see for example [**BP08**].)

As a consequence, even the solution u to the simplest parabolic stochastic PDE, the stochastic heat equation $\partial_t u = \Delta u + \xi$, is quite irregular. As a consequence of the above, combined with classical parabolic Schauder estimates, one can show that if the spatial variable has dimension d, then the solution u belongs to $\mathcal{C}^\alpha$ (again in the parabolic sense) for every $\alpha < \frac{2-d}{2}$, but *not* for $\alpha = \frac{2-d}{2}$. In particular, one expects solutions h to (KPZ) to be of class "just about not" Hölder-$\frac{1}{2}$. This begs the following question: if typical solutions h are nowhere differentiable, what does the nonlinearity $(\partial_x h)^2$ mean? Similarly, we see that one needs to take $\alpha < 0$ in dimensions 2 or higher, which means that in these dimensions the solutions to the stochastic heat equation are no longer functions themselves, but only make sense as space-time distributions. As a consequence, it is not clear at all what the meaning of the nonlinearity Φ^3 is in the dynamical Φ_3^4 model.

This is the type of question that will be addressed in this article. It is hopeless to try to build a consistent theory allowing to multiply any two Schwartz distributions, as already pointed out by Schwartz himself [**Sch54**]. It is possible to extend the space of Schwartz distributions to some larger class of objects that do form an algebra (and such that the product of a Schwartz distribution with a smooth function has its usual meaning), and this has been explored by Colombeau [**Col83**]. However, while this frameworks does allow one to treat stochastic PDEs [**AHR01**], it is not clear at all what the meaning of the resulting solutions are. In fact, there are very strong hints that solutions built using such a theory are not "physically relevant": in the case of the sine-Gordon model they fail to "see" the Kosterlitz-Thouless transition [**FS81, Fal12**], while the theory of regularity structures does break down precisely at that transition [**HS14**], as expected for a theory designed to treat "subcritical" models.

Instead, the direction pursued here is to exploit as much as possible a priori information on the model at hand. Instead of building a general theory allowing to multiply any two distributions, we are only interested in multiplying those distributions that could potentially arise in the right hand side of the equation under consideration. The theory of regularity structures provides the tools and techniques to design such "purpose-built" spaces in a systematic way, as well as the objects required to encode the various renormalisation procedures arising in these problems. While a full exposition of the theory is well beyond the scope of this short introduction, we aim to give a concise overview to most of its concepts. In most cases, we will only state results in a rather informal way and give some ideas as to

how the proofs work, focusing on conceptual rather than technical issues. For precise statements and complete proofs of most of the results exposed here, we refer to the articles [**Hai14, HS15, HQ15, HX15**]. The type of well-posedness results that can be proven using the techniques surveyed in this article include the following.

THEOREM 1.1. *Let* $\xi_\varepsilon = \varrho_\varepsilon * \xi$ *denote the convolution of space-time white noise with a compactly supported smooth mollifier* ϱ *that is scaled by* ε *in the spatial direction(s) and by* ε^2 *in the time direction. Denote by* Φ_ε *the solution to*

$$(1.2) \qquad \partial_t \Phi_\varepsilon = \Delta \Phi_\varepsilon + C_\varepsilon \Phi_\varepsilon - \Phi_\varepsilon^3 + \xi_\varepsilon \ ,$$

on the three-dimensional torus. Then, there exist choices of constants C_ε *diverging as* $\varepsilon \to 0$, *as well as a process* Φ *such that* $\Phi_\varepsilon \to \Phi$ *in probability. Furthermore, if* C_ε *is suitable chosen, then* Φ *does not depend on the mollifier* ϱ.

REMARK 1.2. Very similar results have recently been obtained by slightly different techniques. In [**CC13**], the authors use the theory of paracontrolled distributions introduced in [**GIP12**]. This theory is a kind of Fourier space analogue to the theory of regularity structures, but seems for the moment restricted to first-order expansions. In [**Kup15**] on the other hand, the author uses a variant of Wilson's renormalisation group ideas.

REMARK 1.3. We made an abuse of notation, since the space-time white noise appearing in the equation for h_ε is on $\mathbf{R} \times \mathbf{T}^1$, while the one appearing in the equation for Φ_ε is on $\mathbf{R} \times \mathbf{T}^3$. Similarly, the mollifier ϱ_ε is of course different for the two equations.

REMARK 1.4. In both cases, convergence is in probability in $\mathcal{C}^\alpha$ for $0 < \alpha < \frac{1}{2}$ in the case of the KPZ equation and $-\frac{2}{3} < \alpha < -\frac{1}{2}$ in the case of the Φ_3^4 model. This requires the corresponding initial conditions to also belong to the relevant $\mathcal{C}^\alpha$ spaces.

It is also possible to show that various natural approximation schemes converge to the dynamical Φ_3^4 model. Take for example a function $\theta \mapsto V_\theta$ taking values in the space of even polynomials on $\mathbf{R}$ of some fixed degree $2m$ and consider the equation

$$(1.3) \qquad \partial_t \Phi_{\varepsilon,\theta} = \Delta \Phi_{\varepsilon,\theta} - \varepsilon^{-3/2} V_\theta'(\sqrt{\varepsilon}\Phi_{\varepsilon,\theta}) + \xi_\varepsilon \ ,$$

with ξ_ε as above. This particular scaling arises naturally when rescaling a weakly nonlinear model, see [**HX15**] for more details. Denote by $\mu = \mathcal{N}(0, C)$ the centred Gaussian measure with covariance

$$(1.4) \qquad C = \| \varrho \star P \|_{L^2}^2 \ ,$$

where P denotes the 3-dimensional heat kernel and ϱ is the mollifier appearing in the definition of ξ_ε. Note that since $d = 3$, this quantity is indeed finite. (In dimensions $d \leq 2$ the heat kernel fails to be square integrable at large

scales.) We then define the "effective potential" $\langle V_\theta \rangle = \mu \star V_\theta$ and assume that $\theta \mapsto \langle V_\theta \rangle$ exhibits a pitchfork bifurcation at the origin at $\theta = 0$. We also normalise the solution in such a way that $\langle V_0 \rangle^{(4)}(0) = 6$, which guarantees that $\langle V_0 \rangle'(u) = u^3 + \mathcal{O}(u^5)$ for $u \ll 1$. The main result of [**HX15**], which builds on the analogous results obtained in [**HQ15**] for the KPZ equation then reads

THEOREM 1.5. *In the above setting, there exist values $a > 0$ and $b \in \mathbf{R}$, such that if one sets $\theta = a\varepsilon \log \varepsilon + b\varepsilon$, then the solution to* (1.3) *converges as $\varepsilon \to 0$ to the process Φ built in Theorem 1.1.*

REMARK 1.6. Concerning the initial condition, one needs to consider a sequence $\Phi_0^{(\varepsilon)}$ of smooth initial conditions converging to $\Phi_0 \in \mathcal{C}^\alpha$ with $-\frac{1}{2} - \frac{1}{4m} < \alpha < -\frac{1}{2}$ in a suitable sense. This is because if we chose a fixed initial condition in $\mathcal{C}^{-1/2}$ say, (1.3) may fail to even possess local solutions for fixed $\varepsilon > 0$.

We can also consider approximations of the type given in Theorem 1.1, but with ξ_ε a non-Gaussian approximation to white noise. In this case, one typically has to add additional counterterms in order to obtain the same limit. Consider for example an approximation of the type

$$(1.5) \qquad \xi_\varepsilon(t, x) = \varepsilon^{-5/2} \eta(t/\varepsilon^2, x/\varepsilon) \,,$$

for a stationary process η on $\mathbf{R} \times \mathbf{R}^3$ which admits moments of all orders and has finite dependence in the sense that σ_K and $\sigma_{\bar{K}}$ are independent whenever

$$\inf_{\substack{z \in K \\ \bar{z} \in \bar{K}}} |z - \bar{z}| \geq 1 \,.$$

Here, we wrote σ_K for the σ-algebra generated by all evaluation maps $\eta(z)$ for $z \in K$.[1] It is then possible to show that there exist suitable choices of constants $C_\varepsilon^{(i)}$ such that solutions to

$$(1.6) \qquad \partial_t \Phi_\varepsilon = \Delta \Phi_\varepsilon + C_\varepsilon^{(1)} + C_\varepsilon^{(2)} \Phi_\varepsilon - \Phi_\varepsilon^3 + \xi_\varepsilon \,,$$

converge as $\varepsilon \to 0$ to the process Φ built in Theorem 1.1. Details of the proof can be found in [**HS15**] for the case of the KPZ equation. It is shown in [**HS15**, Remark 6.6] that these constants are of the type

$$(1.7) \qquad C_\varepsilon^{(1)} = \frac{C^{(1,1)}}{\varepsilon^{3/2}} + \frac{C^{(1,2)}}{\varepsilon^{1/2}} \,, \qquad C_\varepsilon^{(2)} = \frac{C^{(2,1)}}{\varepsilon} + c \log \varepsilon + C^{(2,2)} \,,$$

where the $C^{(i,j)}$ depend on the details of the process η, but c is universal. We will provide a short sketch of the argument showing how these constants appear in Section 5.1 below.

Another very nice way in which the dynamical Φ^4 model appears is as the crossover regime for the Glauber dynamic in an Ising-Kac model [**MW14**].

---

[1]Since we always consider equations on bounded tori, we actually need to consider a suitably periodised version of ξ_ε, see [**HS15**, Assumption 2.1] for more details.

Here, one considers an Ising model, but with the usual nearest-neighbour Hamiltonian replaced by a Hamiltonian of the type

$$H = \sum_{x,y} K_\gamma(|x-y|)\sigma_x\sigma_x \;,$$

where $K_\gamma(r) = \gamma^{-d}K(\gamma r)$ for some small value γ, and K is a positive, smooth, compactly supported function. On is then interested in the simultaneous limit $\gamma \to 0$ and $N \to \infty$ (where N is the side length of a discrete torus on which σ lives) at (or sufficiently near) the critical temperature. In dimension $d = 2$, it was shown in [**MW14**] that by suitably choosing N and the inverse temperature as a function of γ and considering the Glauber dynamic on suitable long time scales, one recovers the dynamical Φ_2^4 model in the limit. A similar result is of course conjectured to hold for $d = 3$. We will not consider any discrete approximation of this type in these notes but refer instead to [**CW15**] for a recent review of the techniques involved in the proof of such a result.

One of the main insights used in the proof of the type of results mentioned in this introduction is that while solutions to equations like (KPZ) or (Φ^4) may appear to be very rough (indeed, they are not even functions in the case of (Φ^4)), they can actually be considered to be smooth, provided that one looks at them in the right way. To understand what we mean by this, it is worth revisiting the very notion of "regularity". The way we usually measure regularity is by asking how well a given function can be approximated by polynomials. More precisely, we say that a function $F\colon \mathbf{R}^d \to \mathbf{R}$ is of class $\mathcal{C}^\gamma$ with $\gamma > 0$ if, for every point $x \in \mathbf{R}^d$ we can find a polynomial P_x of degree $\lfloor\gamma\rfloor$ such that $|F(y) - P_x(y)| \lesssim |y-x|^\gamma$ for y close to x. The idea now is to consider spaces of "regular" functions / distributions, but where "regularity" is measured by local proximity not to polynomials, but to linear combinations of some other "basis functions" that are specific to the problem at hand. In particular, the objects that play the role of polynomials are allowed to be random themselves, and they are also allowed to be distributions rather than just functions. One might now hope (and this is indeed the case) that if we are given a consistent product rule on these basic objects and if we have two distributions that are sufficiently "smooth" in the sense that they are locally described by linear combinations of these objects (a "local specification"), then there exists a unique distribution, which we interpret as the product, whose local specification is given by the products of the local specifications of each factor. This then allows to formulate equations like (KPZ) or (Φ^4) as fixed point problems in some of these spaces of "smooth" functions / distributions, and to build a local solution theory with many nice properties very similar to those one would have for the corresponding deterministic problems with smooth inputs.

The remainder of these notes is organised as follows. First, in Section 2, we provide a fresh look at the definitions of the classical spaces $\mathcal{C}^\gamma$ and we describe a natural generalisation, as well as a situation where such a

generalisation is useful. In Section 3, we then give the general definition of a regularity structure and of the associated analogues to the spaces $\mathcal{C}^\gamma$. We also formulate the reconstruction theorem which is fundamental to the theory and we give a simple application. We then provide some operational tools in Sections 3.2 and 3.3. Finally, we show in Sections 4 and 4.3 how to apply the theory to stochastic PDEs in general and the dynamical Φ_3^4 model in particular.

2. Another look at smooth functions

Before we turn to a more precise description of how this idea of replacing Taylor polynomials by purpose-built functions / distributions is implemented in practice, let us first have a slightly different look at the definition of the classical Hölder spaces. For the sake of this discussion, let us fix some exponent $\gamma \in (1,2)$. The space $\mathcal{C}^\gamma(S^1)$ then consists of those functions f that are continuously differentiable and such that their derivative f' satisfies $|f'(t) - f'(s)| \lesssim |t-s|^{\gamma-1}$. A natural seminorm on $\mathcal{C}^\gamma$ is then given by

$$(2.1) \qquad \|f\|_\gamma = \sup_{s \neq t} \frac{|f'(t) - f'(s)|}{|t-s|^{\gamma-1}} \,,$$

where f' denotes the derivative of f. This definition is more complicated than it may appear at first sight since it is given in terms of f', rather than the function f itself, and f' has to be computed first. Instead, we would like to consider an element of $\mathcal{C}^\gamma$ as a *pair* of functions (f, f'), endowed with the norm

$$(2.2) \quad \|(f,f')\|_\gamma = \sup_{s \neq t} \frac{|f'(t) - f'(s)|}{|t-s|^{\gamma-1}} \vee \sup_{s \neq t} \frac{|f(t) - f(s) - f'(s)(t-s)|}{|t-s|^\gamma} \,,$$

where $a \vee b$ denotes the maximum of a and b. It is not difficult to verify that if f' is the derivative of f, then as a consequence of the identity

$$f(t) - f(s) - f'(s)(t-s) = \int_s^t (f'(r) - f'(s))\, dr \,,$$

the seminorm (2.2) is indeed equivalent to (2.1). Furthermore, the finiteness of the second part of the expression (2.2) *forces* f to be differentiable with derivative f', since otherwise that supremum would be infinite. In this point of view, we therefore do not have to impose *a priori* that f' is the derivative of f, this simply turns out to be a side-effect of having $\|(f,f')\|_\gamma < \infty$. Furthermore, viewing the pair (f, f') as our primary data has the advantage that we do not need to compute any additional data (the derivative) from it in order to express its norm. At this stage, one may worry that the notation quickly becomes cumbersome when considering higher degrees of regularity.

Indeed, for $\gamma \in (2,3)$ we could come up with a similar norm, but this time we would have to consider triples (f, f', f''), etc. Instead, we prefer to rewrite an element of $\mathcal{C}^\gamma$ as a single function, but taking values in $\mathbf{R} \oplus \mathbf{R} \approx \mathbf{R}^2$, with the first component equal to f and the second component equal to f'. In order to be able to easily distinguish these components, we denote by $\mathbf{1}$ the basis vector corresponding to the first component and by X the basis vector for the second component, so that our pair (f, f') can be written as

$$F(t) = f(t)\mathbf{1} + f'(t)X \ .$$

At this stage, this is nothing but a change of notation. However, this notation already suggests a very natural product rule for elements in $\mathcal{C}^\gamma$: we *postulate* that $\mathbf{1} \cdot \mathbf{1} = \mathbf{1}$, $\mathbf{1} \cdot X = X \cdot \mathbf{1} = X$, and $X \cdot X = 0$, and we define $(F \cdot G)(t) = F(t) \cdot G(t)$. With this definition, one has

$$(F \cdot G)(t) = f(t)g(t)\mathbf{1} + (f'(t)g(t) + f(t)g'(t))X \ ,$$

i.e. the component multiplying X is "automatically" given by what the Leibniz rule suggests should be its correct expression.

Each of the basis vectors $\tau \in \{\mathbf{1}, X\}$ also comes with a natural "degree" (or "homogeneity") $|\tau|$ given by $|\mathbf{1}| = 0$ and $|X| = 1$. There is an analytical meaning to this homogeneity: the vector $\mathbf{1}$ also quite naturally represents the constant function 1 and the vector X represents the monomial of degree 1 in a Taylor expansion around some base point. In this way, we can view F as a function taking values in the space of Taylor polynomials of degree 1. This suggests the introduction of a family of linear maps $\{\Pi_s\}_{s \in S^1}$ which associate to each basis vector the corresponding Taylor monomial:

$$(2.3) \qquad\qquad (\Pi_s \mathbf{1})(t) = 1 \ , \qquad (\Pi_s X)(t) = t - s \ .$$

With this notation, given $F \in \mathcal{C}^\gamma$, the function $t \mapsto (\Pi_s F(s))(t)$ is nothing but its first-order Taylor expansion at s. Furthermore, the degree of a vector $\tau \in \{\mathbf{1}, X\}$ is precisely the order at which the map $t \mapsto (\Pi_s \tau)(t)$ vanishes near $t = s$. The maps Π_s yield an easy way to recover the "actual" real-valued function described by the vector-valued function F: setting

$$(2.4) \qquad\qquad (\mathcal{R}F)(t) \stackrel{\text{def}}{=} (\Pi_t F(t))(t) \ ,$$

we see that in this case $\mathcal{R}F$ is nothing but the component of F multiplying $\mathbf{1}$, which is indeed the real-valued function f represented by F. One crucial property of the maps Π_s is that varying s can also be achieved by composing them with an adequate linear transformation. Indeed, setting

$$(2.5) \qquad\qquad \Gamma_{st}\mathbf{1} = \mathbf{1} \ , \qquad \Gamma_{st}X = X + (s - t)\mathbf{1} \ ,$$

it is straightforward to verify that one has the identities

$$(2.6) \qquad\qquad \Pi_t = \Pi_s \Gamma_{st} \ , \qquad \Gamma_{st}\Gamma_{tu} = \Gamma_{su} \ .$$

With this notation, and making a slight abuse of notation by also using $\{\mathbf{1}, X\}$ for the dual basis, the norm (2.2) can be rewritten in a more concise

and much more natural way as

$$\|F\|_\gamma = \sup_{\tau \in \{\mathbf{1}, X\}} \frac{|\langle \tau, F(t) - \Gamma_{ts}F(s)\rangle|}{|t - s|^{\gamma - |\tau|}} \ . \tag{2.7}$$

It is now immediate to generalise these definitions to the case of $\gamma \in \mathbf{R}_+ \setminus \mathbf{N}$ by introducing additional basis vectors X^k, postulating that $X^0 = \mathbf{1}$ and $X^k \cdot X^\ell = X^{k+\ell}$, and extending the definition of the maps Γ_{st} to these additional vectors by imposing that it satisfies the multiplicative property

$$\Gamma_{st}(\tau\bar\tau) = (\Gamma_{st}\tau)(\Gamma_{st}\bar\tau) \ .$$

One can then verify that (2.6) still holds provided that one sets $\big(\Pi_s X^k\big)(t) = (t - s)^k$. This is a consequence of the fact that first changing the base point s and then multiplying two monomials is the same as first multiplying them and then changing the base point.

Note now that the definition of the seminorm (2.7) is extremely robust and does not refer anymore to any of the details that make a Taylor polynomial a polynomial. For example, for $\alpha \in (0, 1)$ we could fix an α-Hölder continuous (in the usual sense) function W and, instead of (2.3), we could set

$$\big(\Pi_s \mathbf{1}\big)(t) = 1 \ , \qquad \big(\Pi_s X\big)(t) = W(t) - W(s) \ .$$

Since $\big(\Pi_s X\big)(t)$ now only vanishes at order α near $t = s$, this strongly suggests that we should set $|X| = \alpha$. The identity (2.6) is still satisfied if we change the definition of Γ_{st} into

$$\Gamma_{st}\mathbf{1} = \mathbf{1} \ , \qquad \Gamma_{st}X = X + \big(W(s) - W(t)\big)\mathbf{1} \ . \tag{2.8}$$

With these definitions, the norm (2.7) is still very natural, but its meaning now is that, setting $F(t) = f(t)\mathbf{1} + f'(t)X$ as before, the function f' is Hölder continuous of order $\gamma - \alpha$ and the function f is such that

$$f(t) = f(s) + f'(s)\big(W(t) - W(s)\big) + \mathcal{O}(|t - s|^\gamma) \ . \tag{2.9}$$

In particular, if $\gamma > \alpha$, it no longer implies means that f is of class $\mathcal{C}^\gamma$, only that its increments "look like" some multiple of those of W, up to a remainder of order γ. Furthermore, f' is now no longer equal to the derivative of f.

REMARK 2.1. Consider the interesting case $0 < \alpha < \gamma < 1$ for the situation just described. Then it is no longer necessarily true that the "derivative" f' is uniquely determined by f if all we know is that $\|F\|_\gamma < \infty$. If for example we take $W(t) = \cos(t)$ which is certainly α-Hölder continuous, although it is of course much more than that, then the identity (2.9) simply forces f to be γ-Hölder continuous and puts no restriction whatsoever on f'. If, on the other hand, W happens to be "nowhere γ-Hölder" in the sense that for every $s \in S^1$ one can find a sequence $t_n \to s$ such that $|W(t_n) - W(s)|/|t_n - s|^\gamma \to \infty$, then (2.9) determines f' uniquely. A quantitative version of this statement, together with an application, can be found in [**HP13**].

The functions of class "$\mathcal{C}^\gamma$ with respect to some function W" just described in (2.9) play a prominent role in the theory of controlled rough paths [**Lyo98, Gub04, LCL07**], so let us see how these definitions are useful there. The setting is the following: we want to provide a robust solution theory for a controlled differential equation of the type

$$(2.10) \qquad\qquad dY = f(Y)\, dW(t)\,,$$

where $W \in \mathcal{C}^\alpha$ is a rather rough function (say a typical sample path for an m-dimensional Brownian motion). In general, we allow for W to take values in $\mathbf{R}^m$ and Y in $\mathbf{R}^n$ for arbitrary integers n and m, so we cannot solve (2.10) by simply setting $Y(t) = Z(W(t))$ with $\bar{Y}$ the solution to the ODE $\dot{Z} = f(Z)$. It is a classical result by Young [**You36**] that the Riemann-Stieltjes integral $(Y, W) \mapsto \int_0^\cdot Y\, dW$ makes sense as a continuous map from $\mathcal{C}^\alpha \times \mathcal{C}^\alpha$ into $\mathcal{C}^\alpha$ if and only if $\alpha > \frac{1}{2}$. As a consequence, "naïve" approaches to a pathwise solution to (2.10) with $\int_0^\cdot f(Y)\, dW$ interpreted in Young's sense are bound to fail if W has the regularity of Brownian motion, since the modulus of continuity $\omega(h)$ of the latter behaves no better than $\omega(h) = \sqrt{2h \log 1/h}$. A fortiori, it will fail for example if W is a typical sample path of fractional Brownian motion with Hurst parameter $H < \frac{1}{2}$.

In order to break through this barrier, the main idea is to exploit the a priori "guess" that solutions to (2.10) should "look like W at small scales", i.e. in order to try to define an integral $\int_0^\cdot Z\, dW$ for any two elements Z and W of some function space, we only consider those Z's that could plausibly arise as $Z = f(Y)$ for Y the solution to (2.10). More precisely, one would naturally expect the solution Y to satisfy

$$(2.11) \qquad\qquad Y_t = Y_s + Y_s' W_{s,t} + \mathcal{O}(|t - s|^{2\alpha})\,,$$

for some function Y', where we wrote $W_{s,t}$ as a shorthand for the increment $W_t - W_s$. As a matter of fact, one would expect to have such an expansion with the specific choice $Y' = f(Y)$. As a consequence, one would also expect a similar bound to (2.11) for $Z = f(Y)$ with $Z' = f'(Y)$. In other words, one would make the *a priori* guess that solutions to (2.10) satisfy $\|Y\mathbf{1} + Y'X\|_{2\alpha} < \infty$, provided that we define $\|\cdot\|_{2\alpha}$ as in (2.7) with $|X| = \alpha$ and Γ_{st} as in (2.8) (with the caveat that in the multidimensional case one should now introduce symbols X_i for $i \in \{1, \ldots, m\}$, interpret (2.8) as a "vector" identity, take Y' matrix-valued, etc).

It now remains to provide a coherent construction of the integral $\int Z\, dW$ for such functions (or rather pairs of functions (Z, Z')). The main idea is to then simply *postulate* the values of the integrals

$$(2.12) \qquad\qquad \mathbb{W}_{s,t} =: \int_s^t W_{s,r} \otimes dW_r\,,$$

instead of trying to compute them from W (which is doomed to failure). For a two-parameter function $\mathbb{W}$ to be a "reasonable" candidate for the right

hand side of (2.12), one would like it to satisfy Chen's relations

$$(2.13) \qquad \mathbb{W}_{s,t} - \mathbb{W}_{s,u} - \mathbb{W}_{u,t} = W_{s,u} \otimes W_{u,t} \,,$$

since these follow from the requirements that $\int_s^t c\,dW_r = cW_{s,t}$ for any $c, s, t \in \mathbf{R}$ and that the sum of two integrals with the same integrand over adjacent intervals equals the integral over the union of these intervals. By simple scaling, it is also natural to impose the analytic bound

$$(2.14) \qquad |\mathbb{W}_{s,t}| \lesssim |t - s|^{2\alpha} \,.$$

One can then exploit this additional data to give a coherent definition of expressions of the type $\int Z\,dW$, provided that the path W is "enhanced" with its iterated integrals $\mathbb{W}$ and Z is a "controlled path" of the type (2.11). Indeed, it suffices to set

$$\int_0^t Z_s\,dW_s \overset{\text{def}}{=} \lim_{|\mathcal{P}| \to 0} \sum_{[s,u] \in \mathcal{P}} \left(Z_s\,W_{s,u} + Z'_s\,\mathbb{W}_{s,u} \right) \,,$$

where $\mathcal{P}$ denotes a partition of $[0, t]$ (interpreted as a collection of closed intervals) and $|\mathcal{P}|$ denotes the length of the longest interval in $\mathcal{P}$. It is a fact that the analytical bounds given above, together with (2.13), guarantee that this limit exists as soon as $\alpha > \frac{1}{3}$ and that, if we set

$$Y_t = \int_0^t Z_s\,dW_s \mathbf{1} + Z_t X \,,$$

then one has again $\|Y\|_{2\alpha} < \infty$, thus allowing to formulate (2.10) as a well-behaved fixed point problem. See for example [**Gub04**] or the lecture notes [**FH14**] for a more detailed exposition.

3. The basic theory of regularity structures

Let us now set up in more detail a general framework in which one can define "Hölder-type" spaces as above. Our first ingredient is a vector space T that contains the coefficients of our "Taylor-like" expansion at each point. In the previous example, this space was given by the linear span of $\mathbf{1}$ and X. In general, it is natural to postulate that T is an arbitrary graded vector space $T = \bigoplus_{\alpha \in A} T_\alpha$, for some set A of possible "homogeneities". For example, in the case of the usual Taylor expansion, it is natural to take for A the set of natural numbers and to have T_ℓ contain the coefficients corresponding to the derivatives of order ℓ. In the case of controlled rough paths however, we have already seen that it is natural to take $A = \{0, \alpha\}$, to have again T_0 contain the value of the function Y at any time s, and to have T_α contain its "Gubinelli derivative" Y'_s. This reflects the fact that in that case the "monomial" $t \mapsto (\Pi_s X)(t) = W_{s,t}$ only vanishes at order α near $t = s$, while the usual monomials $t \mapsto (t - s)^\ell$ vanish at integer order ℓ. In general, we only assume that each T_α is a real Banach space, although in many examples of interest these spaces will be finite-dimensional.

This however isn't the full algebraic structure describing Taylor-like expansions. Indeed, we have already seen that a crucial characteristic of Taylor expansions is that an expansion around some point x_0 can be re-expanded around any other point x_1, namely simply by writing

$$(3.1) \qquad (x - x_0)^m = \sum_{k+\ell=m} \binom{m}{k} (x_1 - x_0)^k \cdot (x - x_1)^\ell .$$

(In the case when $x \in \mathbf{R}^d$, k, ℓ and m denote multi-indices and $k! = k_1! \ldots k_d!$.) In general, we have seen in both of our examples that there are linear maps Γ_{st} transforming the coefficients of an expansion around t into the coefficients of the same "polynomial", but this time expanded around s.

What is a natural abstraction of this fact? In view of the above examples, it is natural to impose that any such "reexpansion map" Γ_{st} has the property that if $\tau \in T_\alpha$, then $\Gamma_{st}\tau - \tau \in \bigoplus_{\beta<\alpha} T_\beta =: T_{<\alpha}$. In other words, when reexpanding a homogeneous monomial around a different point, the leading order coefficient remains the same, but lower order monomials may appear. Furthermore, one should be able to compose reexpansions, since taking an expansion around t, reexpanding it around s and then reexpanding the result around a third point r should be the same as reexpanding the first expansion around r. In other words, it seems natural that one has the identity $\Gamma_{rs}\Gamma_{st} = \Gamma_{rt}$, which is indeed the case for the examples we have seen so far. These considerations can be summarised in the following definition of an algebraic structure which we call a *regularity structure*:

DEFINITION 3.1. Let $A \subset \mathbf{R}$ be bounded from below and without accumulation point, and let $T = \bigoplus_{\alpha \in A} T_\alpha$ be a vector space graded by A such that each T_α is a Banach space. Let furthermore G be a group of continuous operators on T such that, for every $\alpha \in A$, every $\Gamma \in G$, and every $\tau \in T_\alpha$, one has $\Gamma\tau - \tau \in T_{<\alpha}$. The triple $\mathscr{T} = (A, T, G)$ is called a *regularity structure* with *model space* T and *structure group* G.

REMARK 3.2. In principle, the set A can be infinite. By analogy with the polynomials, it is then natural to consider T as the set of all formal series of the form $\sum_{\alpha \in A} \tau_\alpha$, where only finitely many of the τ_α's are non-zero, endowed with the topology of term-wise convergence. This also dovetails nicely with the particular form of elements in G. In practice however we will only ever work with finite subsets of A so that the precise topology on T does not matter.

A regularity structure as given by Definition 3.1 is just a kind of algebraic "skeleton": it does not require any underlying configuration space and it contains no information as to which actual functions / distributions its elements are supposed to describe. It only becomes useful for our purpose when endowed with a *model*, which is the analytical "flesh" associating to any $\tau \in T$ and $x_0 \in \mathbf{R}^d$, the actual "Taylor polynomial based at x_0" represented by τ. In order to link the algebraic description to the corresponding

analytical objects, we want elements $\tau \in T_\alpha$ to represent functions (or possibly distributions!) that "vanish at order α" around the given point x_0.

Since we would like to allow elements in T to represent distributions and not just functions, we cannot evaluate them at points and thus need a suitable notion of "vanishing at order α". We achieve this by controlling the size of our distributions when tested against test functions that are localised in a small region around the given point x_0. Given a test function φ on $\mathbf{R}^d$, we write φ_x^λ as a shorthand for

$$\varphi_x^\lambda(y) = \lambda^{-d} \varphi\big(\lambda^{-1}(y - x)\big) \ .$$

Given $r > 0$, we also denote by $\mathcal{B}_r$ the set of all smooth functions $\varphi \colon \mathbf{R}^d \to \mathbf{R}$ such that $\|\varphi\|_{\mathcal{C}^r} \le 1$, and that are furthermore supported in the unit ball around the origin. With this notation, and writing furthermore $\mathcal{S}'(\mathbf{R}^d)$ for the space of distributions (not necessarily tempered) on $\mathbf{R}^d$, our definition of a model for a given regularity structure $\mathscr{T}$ is as follows.

DEFINITION 3.3. Given a regularity structure $\mathscr{T}$ and an integer $d \ge 1$, a *model* for $\mathscr{T}$ on $\mathbf{R}^d$ consists of maps

$$\Pi \colon \mathbf{R}^d \to \mathcal{L}\big(T, \mathcal{S}'(\mathbf{R}^d)\big) \qquad \Gamma \colon \mathbf{R}^d \times \mathbf{R}^d \to G$$
$$x \mapsto \Pi_x \qquad\qquad\qquad (x, y) \mapsto \Gamma_{xy}$$

such that $\Gamma_{xy}\Gamma_{yz} = \Gamma_{xz}$ and $\Pi_x \Gamma_{xy} = \Pi_y$. Furthermore, given $r > |\inf A|$, for any compact set $\mathfrak{K} \subset \mathbf{R}^d$ and constant $\gamma > 0$, we assume that there exists a constant C such that the bounds

$$(3.2) \qquad \big|(\Pi_x \tau)(\varphi_x^\lambda)\big| \le C \lambda^{|\tau|} \|\tau\|_\alpha \ , \qquad \|\Gamma_{xy}\tau\|_\beta \le C|x - y|^{\alpha - \beta} \|\tau\|_\alpha \ ,$$

hold uniformly over $\varphi \in \mathcal{B}_r$, $(x, y) \in \mathfrak{K}$, $\lambda \in (0, 1]$, $\tau \in T_\alpha$ with $\alpha \le \gamma$, and $\beta < \alpha$. We denote by $\mathscr{M}$ the space of all models for a given regularity structure.

REMARK 3.4. Given $\tau \in T$, we wrote $\|\tau\|_\alpha$ for the norm of its component in T_α. In other words, if $\tau = \bigoplus_\alpha \tau_\alpha$ with $\tau_\alpha \in T_\alpha$, then $\|\tau\|_\alpha = \|\tau_\alpha\|_{T_\alpha}$.

REMARK 3.5. The identity $\Pi_x \Gamma_{xy} = \Pi_y$ reflects the fact that Γ_{xy} is the linear map that takes an expansion around y and turns it into an expansion around x. The first bound in (3.2) states what we mean precisely when we say that $\tau \in T_\alpha$ represents a term of order α. The second bound in (3.2) is very natural in view of both (2.5) and (2.8). It states that when expanding a monomial of order α around a new point at distance h from the old one, the coefficient appearing in front of lower-order monomials of order β is of order at most $h^{\alpha - \beta}$.

REMARK 3.6. In many cases of interest, it is natural to scale the different directions of $\mathbf{R}^d$ in a different way. This is the case for example when using the theory of regularity structures to build solution theories for parabolic stochastic PDEs, in which case the time direction "counts double". To deal with such a situation, one can introduce a scaling $\mathfrak{s}$ of $\mathbf{R}^d$, which is just

a collection of d mutually prime strictly positive integers and one defines φ_x^λ in such a way that the ith direction is scaled by $\lambda^{\mathfrak{s}_i}$. In this case, the Euclidean distance between two points should be replaced everywhere by the corresponding scaled distance $|x|_{\mathfrak{s}} = \sum_i |x_i|^{1/\mathfrak{s}_i}$. See also [**Hai14**] for more details.

With these definitions at hand, it is then natural to define an equivalent in this context of the space of γ-Hölder continuous functions in the following way, which is the natural generalisation of (2.7).

DEFINITION 3.7. Given a regularity structure $\mathscr{T}$ equipped with a model (Π, Γ) over $\mathbf{R}^d$, the space $\mathcal{D}^\gamma = \mathcal{D}^\gamma(\mathscr{T}, \Gamma)$ is given by the set of functions $f \colon \mathbf{R}^d \to \bigoplus_{\alpha < \gamma} T_\alpha$ such that, for every compact set $\mathfrak{K}$ and every $\alpha < \gamma$, the exists a constant C with

$$\|f(x) - \Gamma_{xy} f(y)\|_\alpha \leq C |x - y|^{\gamma - \alpha} \tag{3.3}$$

uniformly over $x, y \in \mathfrak{K}$.

REMARK 3.8. Note that, given $\mathscr{T}$ and a model $(\Pi, \Gamma) \in \mathscr{M}$, the corresponding space $\mathcal{D}^\gamma$ is a Fréchet space. However, as we have already seen in Remark 2.1 above, this space does in general depend crucially on the choice of model. We thus have a "total space" $\mathscr{M} \ltimes \mathcal{D}^\gamma$ containing all triples of the form (Π, Γ, F) with $F \in \mathcal{D}^\gamma$ based on the model (Π, Γ). This space is no longer a linear space, but it still comes with a natural topology: the distance between (Π, Γ, f) and $(\bar\Pi, \bar\Gamma, \bar f)$ is given by the smallest constant ϱ such that

$$\|f(x) - \bar f(x) - \Gamma_{xy} f(y) + \bar\Gamma_{xy} \bar f(y)\|_\alpha \leq \varrho |x - y|^{\gamma - \alpha} ,$$
$$\left|(\Pi_x \tau - \bar\Pi_x \tau)(\varphi_x^\lambda)\right| \leq \varrho \lambda^\alpha \|\tau\| ,$$
$$\|\Gamma_{xy} \tau - \bar\Gamma_{xy} \tau\|_\beta \leq \varrho |x - y|^{\alpha - \beta} \|\tau\| ,$$

uniformly for x, y in some compact set.

At this point, we should pause and ask the following question: given $f \in \mathcal{D}^\gamma$ for a given regularity structure and model, what is the actual function / distribution represented by f? Recall that we have seen previously in (2.4) that it should represent $\mathcal{R}f$ given by $(\mathcal{R}f)(x) = \left(\Pi_x f(x)\right)(x)$. However, this definition now no longer makes sense since $\Pi_x f(x)$ is a distribution in general and can therefore not be evaluated at x! The most fundamental result in the theory of regularity structures then states that, given $f \in \mathcal{D}^\gamma$ with $\gamma > 0$, there exists a *unique* distribution $\mathcal{R}f$ on $\mathbf{R}^d$ such that, for every $x \in \mathbf{R}^d$, $\mathcal{R}f$ "looks like $\Pi_x f(x)$ near x". More precisely, one has

THEOREM 3.9. *Let $\mathscr{T}$ be a regularity structure as above and let $(\Pi, \Gamma) \in \mathscr{M}$ be a model for $\mathscr{T}$ on $\mathbf{R}^d$. Then, there exists a unique linear map $\mathcal{R} \colon \mathcal{D}^\gamma \to \mathcal{S}'(\mathbf{R}^d)$ such that*

$$\left|(\mathcal{R}f - \Pi_x f(x))(\varphi_x^\lambda)\right| \lesssim \lambda^\gamma , \tag{3.4}$$

uniformly over $\varphi \in \mathcal{B}_r$ and λ as before, and locally uniformly in x.

We do not provide a proof of this result in these notes, but a concise version can be found in [**Hai15**]. It relies crucially on the fact that our assumptions guarantee that the distributions $\zeta_x = \Pi_x f(x)$ vary slowly with x:

$$(3.5) \quad \big(\zeta_x - \zeta_y\big)(\varphi_x^\lambda) = \Pi_x\big(f(x) - \Gamma_{xy}f(y)\big)(\varphi_x^\lambda) \lesssim \sum_{\alpha < \gamma} \lambda^\alpha\, |x - y|^{\gamma - \alpha} \lesssim \lambda^\gamma \,,$$

for $|y - z| \lesssim \lambda$. In particular, the value of $\zeta_y(\varphi_x^\lambda)$ varies by at most of the order λ^γ as y varies over the support of φ_x^λ. In fact, one can show more generally that given a function $x \mapsto \zeta_x$ which is (locally) uniformly bounded in $\mathcal{C}^\alpha$ for some $\alpha < 0$ and such that (3.5) holds for some $\gamma > 0$, then there exists a unique distribution $\zeta \in \mathcal{C}^\alpha$ such that $\big|(\zeta - \zeta_x)(\varphi_x^\lambda)\big| \lesssim \lambda^\gamma$, locally uniformly in x, see [**Hai14**, Prop. 3.25].

REMARK 3.10. The spaces $\mathcal{D}^\gamma$ are natural generalisations of Hölder spaces. One may wonder whether there are also natural generalisations of other classical function spaces and whether Theorem 3.9 still applies. In [**HL15**], the authors show that approximations to the multiplicative stochastic heat equation converge even when the initial condition is taken to be a Dirac mass. For this, it appears that weighted Hölder-type spaces are not suitable spaces to work with. Instead, the authors work with a generalisation of a scale of inhomogeneous Besov spaces. It seems likely that most classical function spaces admit generalisations to the present context. The rule of thumb regarding the reconstruction theorem is that one typically has existence and uniqueness of the reconstruction operator as soon as the "regularity index" of the corresponding classical function space is positive.

REMARK 3.11. Since the spaces $\mathcal{D}^\gamma$ depend on the choice of model (Π, Γ), the reconstruction operator $\mathcal{R}$ itself also depends on that choice. Since the aim of the theory is to provide robust approximation procedures, we should really view $\mathcal{R}$ as a map from the total space $\mathscr{M} \ltimes \mathcal{D}^\gamma$ into $\mathcal{S}'(\mathbf{R}^d)$. It turns out that even if viewed it in this way, $\mathcal{R}$ is still a continuous map, although it is no longer linear. The fact that this stronger continuity property also holds is crucial when showing that sequences of solutions to mollified equations all converge to the same limiting object.

REMARK 3.12. In the particular case where $\Pi_x \tau$ happens to be a continuous function for every $\tau \in T$ (and every $x \in \mathbf{R}^d$), $\mathcal{R}f$ is also a continuous function and one has the identity

$$(3.6) \qquad\qquad \big(\mathcal{R}f\big)(x) = \big(\Pi_x f(x)\big)(x) \,.$$

This can easily be seen by noting that in this case, if φ integrates to 1, then as $\lambda \to 0$, the quantity $\big(\Pi_x f(x)\big)(\varphi_x^\lambda)$ appearing in (3.4) converges to $\big(\Pi_x f(x)\big)(x)$. Since the right hand side converges to 0, this implies that $\big(\mathcal{R}f\big)(\varphi_x^\lambda)$ converges to $\big(\Pi_x f(x)\big)(x)$, whence (3.6) follows.

3.1. A classical result from harmonic analysis. It is a classical result in harmonic analysis [**BCD11**] that the product extends naturally to $\mathcal{C}^{-\alpha} \times \mathcal{C}^{\beta}$ into $\mathcal{S}'(\mathbf{R}^d)$ if and only if $\beta > \alpha$. The reconstruction theorem yields a straightforward proof of the "if" part of this result:

THEOREM 3.13. *There is a continuous bilinear map* $B\colon \mathcal{C}^{-\alpha} \times \mathcal{C}^{\beta} \to \mathcal{S}'(\mathbf{R}^d)$ *such that* $B(f,g) = fg$ *for any two continuous functions* f *and* g.

PROOF. Assume from now on that $\xi \in \mathcal{C}^{-\alpha}$ for some $\alpha > 0$ and that $f \in \mathcal{C}^{\beta}$ for some $\beta > \alpha$. We then build a regularity structure $\mathscr{T}$ in the following way. For the set A, we take $A = \mathbf{N} \cup (\mathbf{N} - \alpha)$ and for T, we set $T = V \oplus W$, where each one of the spaces V and W is a copy of the polynomial structure in d commuting variables described in Section 2. We also choose the structure group G as in the polynomial structure, acting simultaneously on each of the two instances.

As before, we denote by X^k the canonical basis vectors in V. We also use the suggestive notation "ΞX^k" for the corresponding basis vector in W, but we postulate that $|\Xi X^k| = \alpha + |k|$ rather than $|\Xi X^k| = |k|$ as would usually be the case. Given any distribution $\xi \in \mathcal{C}^{-\alpha}$, we then define a model (Π^ξ, Γ), where Γ is as in the canonical polynomial model, while Π^ξ acts as

$$\left(\Pi_x^\xi X^k\right)(y) = (y - x)^k , \qquad \left(\Pi_x^\xi \Xi X^k\right)(y) = (y - x)^k \xi(y) ,$$

with the obvious abuse of notation in the second expression. It is then straightforward to verify that $\Pi_y = \Pi_x \circ \Gamma_{xy}$ and that the relevant analytical bounds are satisfied, so that this is indeed a model.

Denote now by $\mathcal{R}^\xi$ the reconstruction map associated to the model (Π^ξ, Γ) and, for $f \in \mathcal{C}^\beta$, denote by F the element in $\mathcal{D}^\beta$ given by the local Taylor expansion of f of order β at each point. Note that even though the space $\mathcal{D}^\beta$ does in principle depend on the choice of model, in our situation $F \in \mathcal{D}^\beta$ for any choice of ξ since Γ_{xy} is independent of ξ. It follows immediately from the definitions that the map $x \mapsto \Xi F(x)$ belongs to $\mathcal{D}^{\beta - \alpha}$ so that, provided that $\beta > \alpha$, one can apply the reconstruction operator to it. This suggests that the multiplication operator we are looking for can be defined as

$$B(f, \xi) = \mathcal{R}^\xi\left(\Xi F\right) .$$

By Theorem 3.9, this is a jointly continuous map from $\mathcal{C}^\beta \times \mathcal{C}^{-\alpha}$ into $\mathcal{S}'(\mathbf{R}^d)$, provided that $\beta > \alpha$. If ξ happens to be a smooth function, then it follows immediately from Remark 3.12 that $B(f, \xi) = f(x)\xi(x)$, so that B is indeed the requested continuous extension of the usual product. $\square$

REMARK 3.14. As a consequence of (3.4), it follows that $B\colon \mathcal{C}^{-\alpha} \times \mathcal{C}^\beta \to \mathcal{C}^{-\alpha}$.

3.2. Products. One of the main purposes of the theory presented here is to give a robust way to multiply distributions (or functions with distributions) that goes beyond the barrier illustrated by Theorem 3.13. Provided that our functions / distributions are represented as elements in $\mathcal{D}^\gamma$ for some

model and regularity structure, we can multiply their "Taylor expansions" pointwise, provided that we give ourselves a table of multiplication on T.

It is natural to consider products with the following properties. Here, given a regularity structure, we say that a subspace $V \subset T$ is a *sector* if it is invariant under the action of the structure group G and if it can furthermore be written as $V = \bigoplus_{\alpha \in A} V_\alpha$ with $V_\alpha \subset T_\alpha$.

DEFINITION 3.15. Given a regularity structure (T, A, G) and two sectors $V, \bar{V} \subset T$, a *product* on $(V, \bar{V})$ is a bilinear map $\star \colon V \times \bar{V} \to T$ such that, for any $\tau \in V_\alpha$ and $\bar{\tau} \in \bar{V}_\beta$, one has $\tau \star \bar{\tau} \in T_{\alpha+\beta}$ and such that, for any element $\Gamma \in G$, one has $\Gamma(\tau \star \bar{\tau}) = \Gamma\tau \star \Gamma\bar{\tau}$.

REMARK 3.16. The condition that homogeneities add up under multiplication is very natural bearing in mind the case of the polynomial regularity structure. The second condition is also very natural since it merely states that if one reexpands the product of two "polynomials" around a different point, one should obtain the same result as if one reexpands each factor first and then multiplies them together.

Given such a product, we can ask ourselves when the pointwise product of an element $\mathcal{D}^{\gamma_1}$ with an element in $\mathcal{D}^{\gamma_2}$ again belongs to some $\mathcal{D}^\gamma$. In order to answer this question, we introduce the notation $\mathcal{D}_\alpha^\gamma$ to denote those elements $f \in \mathcal{D}^\gamma$ such that furthermore $f(x) \in T_{\geq\alpha}$ for every x. With this notation at hand, it is not too difficult to verify that one has the following result:

THEOREM 3.17. *Let $f_1 \in \mathcal{D}_{\alpha_1}^{\gamma_1}(V)$, $f_2 \in \mathcal{D}_{\alpha_2}^{\gamma_2}(\bar{V})$, and let $\star$ be a product on $(V, \bar{V})$. Then, the function f given by $f(x) = f_1(x) \star f_2(x)$ belongs to $\mathcal{D}_\alpha^\gamma$ with*

$$(3.7) \qquad \alpha = \alpha_1 + \alpha_2\,, \qquad \gamma = (\gamma_1 + \alpha_2) \wedge (\gamma_2 + \alpha_1)\,.$$

The proof of this result is straightforward and can be found in [**Hai14, Hai15**]. It is clear that the formula (3.7) for γ is optimal in general as can be seen from the following two "reality checks". First, consider the case of the polynomial model and take $f_i \in \mathcal{C}^{\gamma_i}$. In this case, the truncated Taylor series F_i for f_i belong to $\mathcal{D}_0^{\gamma_i}$. It is clear that in this case, the product cannot be expected to have better regularity than $\gamma_1 \wedge \gamma_2$ in general, which is indeed what (3.7) states. The second reality check comes from the example of Section 3.1. In this case, one has $F \in \mathcal{D}_0^\beta$, while the constant function $x \mapsto \Xi$ belongs to $\mathcal{D}_{-\alpha}^\infty$ so that, according to (3.7), one expects their product to belong to $\mathcal{D}_{-\alpha}^{\beta-\alpha}$, which is indeed the case.

REMARK 3.18. In order to obtain robust approximation results, one would like to obtain bounds on the distance between $f_1 \star f_2$ and $\bar{f}_1 \star \bar{f}_2$ in cases where the f_i belong to spaces $\mathcal{D}^{\gamma_i}$ based on some model $(\Pi, \Gamma) \in \mathcal{M}$, while the $\bar{f}_i$ belong to spaces $\mathcal{D}^{\gamma_i}$ based on a different model $(\bar{\Pi}, \bar{\Gamma}) \in \mathcal{M}$. This can also be obtained, but makes the proof slightly more lengthy.

REMARK 3.19. Even if both $\mathcal{R}f_1$ and $\mathcal{R}f_2$ happens to be continuous functions, this does *not* in general imply that $\mathcal{R}(f_1 \star f_2)(x) = (\mathcal{R}f_1)(x)\,(\mathcal{R}f_2)(x)$!

For example, fix $\kappa < 0$ and consider the regularity structure given by $A = (-2\kappa, -\kappa, 0)$, with each T_α being a copy of $\mathbf{R}$ given by $T_{-n\kappa} = \langle \Xi^n \rangle$. We furthermore take for G the trivial group. This regularity structure comes with an obvious product by setting $\Xi^m \star \Xi^n = \Xi^{m+n}$ provided that $m+n \le 2$. Then a perfectly valid model is

$$(3.8) \qquad \left(\Pi_x \Xi^0\right)(y) = 1\,, \quad \left(\Pi_x \Xi^1\right)(y) = 0\,, \quad \left(\Pi_x \Xi^2\right)(y) = c\,,$$

for which, setting $F(x) = f(x)\Xi^0 + g(x)\Xi^1$, one has $(\mathcal{R}F)(x) = f(x)$, but $(\mathcal{R}F^2)(x) = f^2(x) + cg^2(x)$. This flexibility is crucial since the "naive" definition of the product is usually broken by renormalisation, as we will see in Section 4.5 below.

3.3. Schauder estimates and admissible models. One of the reasons why the theory of regularity structures is very successful at providing detailed descriptions of the small-scale features of solutions to semilinear (S)PDEs is that it comes with very sharp Schauder estimates. Recall that the classical Schauder estimates state that if $K \colon \mathbf{R}^d \to \mathbf{R}$ is a kernel that is smooth everywhere, except for a singularity at the origin that is (approximately) homogeneous of degree $\beta - d$ for some $\beta > 0$, then the operator $f \mapsto K * f$ maps $\mathcal{C}^\alpha$ into $\mathcal{C}^{\alpha+\beta}$ for every $\alpha \in \mathbf{R}$, except for those values for which $\alpha + \beta \in \mathbf{N}$. (See for example [**Sim97**].)

It turns out that similar Schauder estimates hold in the context of general regularity structures in the sense that it is in general possible to build an operator $\mathcal{K} \colon \mathcal{D}^\gamma \to \mathcal{D}^{\gamma+\beta}$ with the property that $\mathcal{R}\mathcal{K}f = K * \mathcal{R}f$. Of course, such a statement can only be true if our regularity structure contains not only the objects necessary to describe $\mathcal{R}f$ up to order γ, but also those required to describe $K * \mathcal{R}f$ up to order $\gamma + \beta$. What are these objects? At this stage, it might be useful to reflect on the effect of the convolution of a singular function (or distribution) with K.

Let us assume for a moment that f is also smooth everywhere, except at some point x_0. It is then straightforward to convince ourselves that $K * f$ is also smooth everywhere, except at x_0. Indeed, for any $\delta > 0$, we can write $K = K_\delta + K_\delta^c$, where K_δ is supported in a ball of radius δ around 0 and K_δ^c is a smooth function. Similarly, we can decompose f as $f = f_\delta + f_\delta^c$, where f_δ is supported in a δ-ball around x_0 and f_δ^c is smooth. Since the convolution of a smooth function with an arbitrary distribution is smooth, it follows that the only non-smooth component of $K * f$ is given by $K_\delta * f_\delta$, which is supported in a ball of radius 2δ around x_0. Since δ was arbitrary, the statement follows. By linearity, this strongly suggests that the local structure of the singularities of $K * f$ can be described completely by only using knowledge on the local structure of the singularities of f. It also suggests that the "singular part" of the operator $\mathcal{K}$ should be local, with the non-local parts of $\mathcal{K}$ only contributing to the "regular part".

This discussion suggests that we certainly need the following ingredients to build an operator $\mathcal{K}$ with the desired properties:

- The canonical polynomial structure should be part of our regularity structure in order to be able to describe the "regular parts".
- We should be given an "abstract integration operator" $\mathcal{I}$ on T which describes how the "singular parts" of $\mathcal{R}f$ transform under convolution by K.
- We should restrict ourselves to models which are "compatible" with the action of $\mathcal{I}$ in the sense that the behaviour of $\Pi_x \mathcal{I}\tau$ should relate in a suitable way to the behaviour of $K * \Pi_x \tau$ near x.

One way to implement these ingredients is to assume first that our model space T contains abstract polynomials in the following sense.

ASSUMPTION 3.20. *There exists a sector $\bar{T} \subset T$ isomorphic to the space of abstract polynomials in d commuting variables. In other words, $\bar{T}_\alpha \neq 0$ if and only if $\alpha \in \mathbf{N}$, and one can find basis vectors X^k of $T_{|k|}$ such that every element $\Gamma \in G$ acts on $\bar{T}$ by $\Gamma X^k = (X - h)^k$ for some $h \in \mathbf{R}^d$.*

Furthermore, we assume that there exists an abstract integration operator $\mathcal{I}$ with the following properties.

ASSUMPTION 3.21. *There exists a linear map $\mathcal{I}: T \to T$ such that $\mathcal{I}T_\alpha \subset T_{\alpha+\beta}$, such that $\mathcal{I}\bar{T} = 0$, and such that, for every $\Gamma \in G$ and $\tau \in T$, one has*

$$(3.9) \qquad \Gamma\mathcal{I}\tau - \mathcal{I}\Gamma\tau \in \bar{T} .$$

Finally, we want to consider models that are compatible with this structure for a given kernel K. For this, we first make precise what we mean exactly when we said that K is approximately homogeneous of degree $\beta - d$.

ASSUMPTION 3.22. *One can write $K = \sum_{n \geq 0} K_n$ where each of the kernels $K_n: \mathbf{R}^d \to \mathbf{R}$ is smooth and compactly supported in a ball of radius 2^{-n} around the origin. Furthermore, we assume that for every multiindex k, one has a constant C such that the bound*

$$(3.10) \qquad \sup_x |D^k K_n(x)| \leq C2^{n(d-\beta+|k|)} ,$$

holds uniformly in n. Finally, we assume that $\int K_n(x)P(x)\,dx = 0$ for every polynomial P of degree at most N, for some sufficiently large value of N.

REMARK 3.23. *It turns out that in order to define the operator $\mathcal{K}$ on $\mathcal{D}^\gamma$, we will need K to annihilate polynomials of degree N for some $N \geq \gamma + \beta$.*

REMARK 3.24. *The last assumption may appear to be extremely stringent at first sight. In practice, this turns out not to be a problem at all. Say for example that we want to define an operator that represents convolution with P, the fundamental solution to the heat equation. Then, P can be decomposed into a sum of terms satisfying the bound (3.10) with $\beta = 2$, but it does of course not annihilate generic polynomials and it is not supported in the ball of radius 1.*

However, for any fixed value of $N > 0$, it is straightforward to decompose P as $P = K + \hat{K}$, where the kernel K is compactly supported and satisfies all of the properties mentioned above, and the kernel $\hat{K}$ is smooth. Lifting the convolution with $\hat{K}$ to an operator from $\mathcal{D}^\gamma \to \mathcal{D}^{\gamma+\beta}$ (actually to $\mathcal{D}^{\bar{\gamma}}$ for any $\bar{\gamma} > 0$) is straightforward, so that we have reduced our problem to that of constructing an operator describing the convolution by K.

Given such a kernel K, we can now make precise what we meant earlier when we said that the models under consideration should be compatible with the kernel K.

DEFINITION 3.25. Given a kernel K as in Assumption 3.22 and a regularity structure $\mathcal{T}$ satisfying Assumptions 3.20 and 3.21, we say that a model (Π, Γ) is *admissible* if the identities

$$(3.11) \qquad \left(\Pi_x X^k\right)(y) = (y - x)^k \,, \qquad \Pi_x \mathcal{I}\tau = K * \Pi_x \tau - \Pi_x \mathcal{J}(x)\tau \,,$$

hold for every $\tau \in T$ in the domain of $\mathcal{I}$. Here, $\mathcal{J}(x) \colon T \to \bar{T}$ is the linear map given on homogeneous elements by

$$(3.12) \qquad \mathcal{J}(x)\tau = \sum_{|k| < |\tau| + \beta} \frac{X^k}{k!} \int D^{(k)} K(x - y) \left(\Pi_x \tau\right)(dy) \,.$$

REMARK 3.26. While $K * \xi$ is well-defined for any distribution ξ, it is not so clear *a priori* whether the operator $\mathcal{J}(x)$ given in (3.12) is also well-defined. It turns out that the axioms of a model do ensure that this is the case. The correct way of interpreting (3.12) is by

$$\mathcal{J}(x)\tau = \sum_{|k| < |\tau| + \beta} \sum_{n \geq 0} \frac{X^k}{k!} \left(\Pi_x \tau\right)\left(D^{(k)} K_n(x - \cdot)\right) \,.$$

The scaling properties of the K_n ensure that $2^{(\beta - |k|)n} D^{(k)} K_n(x - \cdot)$ is of the form $c\varphi_x^\lambda$ with $\varphi \in \mathcal{B}_r$ (for arbitrary r), $c \approx 1$ and $\lambda \approx 2^{-n}$. As a consequence of (3.2), one then has

$$\left|\left(\Pi_x \tau\right)\left(D^{(k)} K_n(x - \cdot)\right)\right| \lesssim 2^{(|k| - \beta - |\tau|)n} \,,$$

so that this expression is indeed summable as long as $|k| < |\tau| + \beta$.

REMARK 3.27. The above definition of an admissible model dovetails very nicely with our axioms defining a general model. Indeed, starting from *any* regularity structure $\mathcal{T}$, *any* model (Π, Γ) for $\mathcal{T}$, and a kernel K satisfying Assumption 3.22, it is usually possible to build a larger regularity structure $\hat{\mathcal{T}}$ containing $\mathcal{T}$ (in the "obvious" sense that $T \subset \hat{T}$ and the action of $\hat{G}$ on T is compatible with that of G) and endowed with an abstract integration map $\mathcal{I}$, as well as an admissible model $(\hat{\Pi}, \hat{\Gamma})$ on $\hat{\mathcal{T}}$ which reduces to (Π, Γ) when restricted to T. See [**Hai14**] for more details.

The only exception to this rule arises when the original structure T contains some homogeneous element τ which does not represent a polynomial and which is such that $|\tau| + \beta \in \mathbf{N}$. Since the bounds appearing both in the

definition of a model and in Assumption 3.22 are only upper bounds, it is in practice easy to exclude such a situation by slightly tweaking the definition of either the exponent β or of the original regularity structure $\mathscr{T}$.

With all of these definitions in place, we can finally build the operator $\mathcal{K}\colon \mathcal{D}^\gamma \to \mathcal{D}^{\gamma+\beta}$ announced at the beginning of this section. Recalling the definition of $\mathcal{J}$ from (3.12), we set

$$(3.13) \qquad \big(\mathcal{K}f\big)(x) = \mathcal{I}f(x) + \mathcal{J}(x)f(x) + \big(\mathcal{N}f\big)(x) \,,$$

where the operator $\mathcal{N}$ is given by

$$(3.14) \qquad \big(\mathcal{N}f\big)(x) = \sum_{|k|<\gamma+\beta} \frac{X^k}{k!} \int D^{(k)}K(x-y)\big(\mathcal{R}f - \Pi_x f(x)\big)(dy) \,.$$

Note first that thanks to the reconstruction theorem, it is possible to verify that the right hand side of (3.14) does indeed make sense for every $f \in \mathcal{D}^\gamma$ in virtually the same way as in Remark 3.26. One has:

THEOREM 3.28. *Let K be a kernel satisfying Assumption 3.22, let $\mathscr{T} = (A, T, G)$ be a regularity structure satisfying Assumptions 3.20 and 3.21, and let (Π, Γ) be an admissible model for $\mathscr{T}$. Then, for every $f \in \mathcal{D}^\gamma$ with $\gamma \in (0, N-\beta)$ and $\gamma + \beta \notin \mathbf{N}$, the function $\mathcal{K}f$ defined in (3.13) belongs to $\mathcal{D}^{\gamma+\beta}$ and satisfies $\mathcal{R}\mathcal{K}f = K * \mathcal{R}f$.*

PROOF. The complete proof of this result can be found in [**Hai14**] and will not be given here. Let us simply show that one has indeed $\mathcal{R}\mathcal{K}f = K * \mathcal{R}f$ in the particular case when our model consists of continuous functions so that Remark 3.12 applies. In this case, one has

$$\big(\mathcal{R}\mathcal{K}f\big)(x) = \big(\Pi_x(\mathcal{I}f(x) + \mathcal{J}(x)f(x))\big)(x) + \big(\Pi_x\big(\mathcal{N}f\big)(x)\big)(x) \,.$$

As a consequence of (3.11), the first term appearing in the right hand side of this expression is given by

$$\big(\Pi_x(\mathcal{I}f(x) + \mathcal{J}(x)f(x))\big)(x) = \big(K * \Pi_x f(x)\big)(x) \,.$$

On the other hand, the only term contributing to the second term is the one with $k = 0$ (which is always present since $\gamma > 0$ by assumption) which then yields

$$\big(\Pi_x\big(\mathcal{N}f\big)(x)\big)(x) = \int K(x-y)\big(\mathcal{R}f - \Pi_x f(x)\big)(dy) \,.$$

Adding both of these terms, we see that the expression $\big(K * \Pi_x f(x)\big)(x)$ cancels, leaving us with the desired result. $\qquad\square$

4. Application of the theory to the dynamical Φ_3^4 model

We now sketch how the theory of regularity structures can be used to obtain the kind of convergence result stated in Theorem 1.1. We will only focus on the dynamical Φ_3^4 model and ignore the corresponding result for the KPZ equation. First of all, we note that while our solution Φ will be a space-time distribution (or rather an element of $\mathcal{D}^\gamma$ for some regularity

structure with a model over $\mathbf{R}^4$), 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

$$\varphi^\lambda_{(s,x)}(t,y) = \lambda^{-5}\varphi\big(\lambda^{-2}(t-s), \lambda^{-1}(y-x)\big) \ .$$

Accordingly, the "effective dimension" of our space-time is actually 5, rather than 4. The theory presented above extends *mutatis mutandis* to this setting. (Note in particular that when considering the degree of a regular monomial, powers of the time variable should now be counted double.) With this way of measuring regularity, space-time white noise belongs to $\mathcal{C}^{-\alpha}$ for every $\alpha > \frac{5}{2}$.

4.1. Construction of the associated regularity structure. Our first step is to build a regularity structure that is sufficiently large to allow to reformulate (Φ^4) as a fixed point in $\mathcal{D}^\gamma$ for some $\gamma > 0$. Denoting by P the heat kernel, we can write the solution to (Φ^4) with initial condition Φ_0 as

$$(4.1) \qquad\qquad \Phi = P * \mathbf{1}_{t>0}\big(\xi - \Phi^3\big) + P\Phi_0 \ ,$$

where $*$ denotes space-time convolution and where we denote by $P\Phi_0$ the solution to the heat equation with initial condition Φ_0. Here, $\mathbf{1}_{t>0}$ denotes the indicator function of the set $\{(t,x) : t > 0\}$. In order to have a chance of fitting this into the framework described above, we first decompose the heat kernel P as $P = K + \hat{K}$, where K satisfies all of the assumptions of Section 3.3 with $\beta = 2$ and the remainder $\hat{K}$ is smooth, see Remark 3.24. For any regularity structure containing the usual Taylor polynomials and equipped with an admissible model, is straightforward to associate to $\hat{K}$ an operator $\hat{\mathcal{K}}\colon \mathcal{D}^\gamma \to \mathcal{D}^{\gamma+2}$ via

$$\big(\hat{\mathcal{K}}f\big)(z) = \sum_{|k|<\gamma+2} \frac{X^k}{k!}\big(D^{(k)}\hat{K} * \mathcal{R}f\big)(z) \ ,$$

where z denotes a space-time point. Similarly, the harmonic extension of Φ_0 can be lifted to an element in $\mathcal{D}^\gamma$ (for any fixed $\gamma > 0$) which we denote again by $P\Phi_0$ by considering its Taylor expansion around every space-time point. At this stage, we note that we actually cheated a little: while $P\Phi_0$ is smooth in $\{(t,x) : t > 0, x \in \mathbf{T}^3\}$ and vanishes when $t < 0$, it is of course singular on the time-0 hyperplane $\{(0,x) : x \in \mathbf{T}^3\}$. Similarly, we have the problem that the function $\mathbf{1}_{t>0}$ does not belong to any $\mathcal{D}^\gamma$. Both of these problems can be cured at once by introducing weighted versions of the spaces $\mathcal{D}^\gamma$ allowing for singularities on a given hyperplane. A precise definition of these spaces and their behaviour under multiplication and the action of the integral operator $\mathcal{K}$ can be found in [**Hai14**]. For the purpose of the informal discussion given here, we will simply ignore this problem.

As in Section 3.1, we furthermore introduce a new symbol Ξ which will be used to represent the noise ξ. This suggests that the formulation of (Φ^4)

in a suitable (weighted) space $\mathcal{D}^\gamma$ should be of the form

$$(4.2) \qquad \Phi = \mathcal{P}\mathbf{1}_{t>0}\big(\Xi - \Phi^3\big) + P\Phi_0 \,,$$

where we set $\mathcal{P} = \mathcal{K} + \hat{\mathcal{K}}$. In view of (3.13), for $t > 0$ this equation is of the type

$$(4.3) \qquad \Phi = \mathcal{I}\big(\Xi - \Phi^3\big) + (\ldots) \,,$$

where the terms $(\ldots)$ consist of functions that take values in the subspace $\bar{T} \subset T$ spanned by the regular Taylor monomials X^k. In order to build a regularity structure in which (4.3) can be formulated, it is natural to start with the structure given by abstract polynomials (again with the parabolic scaling which causes the abstract "time" variable to have homogeneity 2 rather than 1), and to add a symbol Ξ to it which we postulate to have homogeneity $-\frac{5}{2}^-$, where we denote by α^- an exponent strictly smaller than, but arbitrarily close to, the value α.

We then simply add to T all of the formal expressions that an application of the right hand side of (4.3) can generate for the description of Φ, Φ^2, and Φ^3. The homogeneity of a given expression is completely determined by the rules $|\mathcal{I}\tau| = |\tau| + 2$ and $|\tau\bar{\tau}| = |\tau| + |\tau|$. More precisely, we consider a collection $\mathcal{U}$ of formal expressions which is the smallest collection containing X^k and $\mathcal{I}(\Xi)$, and such that

$$(4.4) \qquad \tau_1, \tau_2, \tau_3 \in \mathcal{U} \quad \Rightarrow \quad \mathcal{I}(\tau_1\tau_2\tau_3) \in \mathcal{U} \,,$$

where it is understood that $\mathcal{I}(X^k) = 0$ for every multiindex k. We then set

$$(4.5) \qquad \mathcal{W} = \{\Xi\} \cup \{\tau_1\tau_2\tau_3 \,:\, \tau_i \in \mathcal{U}\} \,,$$

and we define our space T as the set of all linear combinations of elements in $\mathcal{W}$. (Note that since $\mathbf{1} \in \mathcal{U}$, one does in particular have $\mathcal{U} \subset \mathcal{W}$.) Naturally, $\mathcal{T}_\alpha$ consists of those linear combinations that only involve elements in $\mathcal{W}$ that are of homogeneity α. It is not too difficult to convince oneself that, for every $\alpha \in \mathbf{R}$, $\mathcal{W}$ contains only finitely many elements of homogeneity less than α, so that each $\mathcal{T}_\alpha$ is finite-dimensional.

In order to simplify expressions later, we will use the following shorthand graphical notation for elements of $\mathcal{W}$. For Ξ, we simply draw a dot. The integration map is then represented by a downfacing line and the multiplication of symbols is obtained by joining them at the root. For example, we have

$$\mathcal{I}(\Xi) = \mathord{\uparrow} \,, \quad \mathcal{I}(\Xi)^3 = \mathord{\vee\!\!\!\vee} \,, \quad \mathcal{I}(\Xi)\mathcal{I}(\mathcal{I}(\Xi)^3) = \mathord{\curlyvee} \,.$$

Symbols containing factors of X have no particular graphical representation, so we will for example write $X_i\mathcal{I}(\Xi)^2 = X_i\mathord{\vee}$. With this notation, the space T is given by

$$T = \langle \Xi, \mathord{\vee\!\!\!\vee}, \mathord{\vee}, \mathord{\Downarrow}, \mathord{\uparrow}, \mathord{\curlyvee}, \mathord{\curlyvee\!\!}, X_i\mathord{\vee}, \mathbf{1}, \mathord{\curlyvee}, \mathord{\curlyvee}, \ldots \rangle \,,$$

where we ordered symbols in increasing order of homogeneity and used $\langle \cdot \rangle$ to denote the linear span. Given any sufficiently regular function ξ (say a

continuous space-time function), there is then a canonical way of lifting ξ to a model $\mathscr{L}(\xi) = (\Pi, \Gamma)$ for T by setting

$$(4.6a) \qquad \left(\Pi_x \Xi\right)(y) = \xi(y) \,, \qquad \left(\Pi_x X^k\right)(y) = (y - x)^k \,,$$

and then recursively by

$$(4.6b) \qquad \left(\Pi_x \tau \bar{\tau}\right)(y) = \left(\Pi_x \tau\right)(y) \cdot \left(\Pi_x \bar{\tau}\right)(y) \,,$$

as well as (3.11). (Note that here we used x and y as notations for generic space-time points in order to keep notations compact.)

4.2. Construction of the structure group. So far, we have only described the vector space T arising in the construction of a regularity structure suitable for the analysis of the dynamical Φ_3^4 model, but we have not yet described the corresponding structure group G. The reason why a non-trivial group G is needed is that in (3.11), as soon as $|\mathcal{I}(\tau)| > 0$, $\Pi_z \mathcal{I}(\tau)$ depends non-trivially on the base point z, so that G is needed in order to be able to enforce the algebraic relation $\Pi_{\bar{z}} = \Pi_z \Gamma_{z\bar{z}}$ of Definition 3.3, which should be interpreted as the action of "reexpanding" a "Taylor series" around a different point.

In our case, in view of (3.11), the coefficients of these reexpansions will naturally be some polynomials in x and in the expressions appearing in (3.12). This suggests that we should define a space T^+ whose basis vectors consist of formal expressions of the type

$$(4.7) \qquad X^k \prod_{i=1}^{N} \mathscr{I}_{\ell_i}(\tau_i) \,,$$

where N is an arbitrary but finite number, the τ_i are basis elements of T different from the Taylor monomials X^k, and the ℓ_i are d-dimensional multiindices satisfying $|\ell_i| < |\tau_i| + 2$. (The last bound is a reflection of the restriction of the summands in (3.12) with $\beta = 2$.) The space T^+ also admits a natural graded structure $T^+ = \bigoplus T_\alpha^+$ by setting

$$|\mathscr{I}_\ell(\tau)| = |\tau| + 2 - |\ell| \,, \qquad |X^k| = |k| \,,$$

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

To any given admissible model (Π, Γ), it is then natural to associate linear maps $f_x \colon T^+ \to \mathbf{R}$ by $f_x(X^k) = (-x)^k$, $f_x(\sigma\bar{\sigma}) = f_x(\sigma)f_x(\bar{\sigma})$, and

$$(4.8) \qquad f_x(\mathscr{I}_\ell \tau) = - \sum_{|k+\ell| < |\tau|+2} \frac{(-x)^k}{k!} \int D^{(\ell+k)} K(x - y) \left(\Pi_x \tau\right)(dy) \,.$$

The minus signs are here purely by convention and serve to make some expressions simpler later on.[2] It then turns out that it is possible to build a linear map $\Delta \colon T \to T \otimes T^+$ such that if we define $F_x \colon T \to T$ by

$$(4.9) \qquad\qquad F_x \tau = (\mathrm{id} \otimes f_x)\Delta\tau \ ,$$

where I denotes the identity operator on T, then these maps are invertible and, given an admissible model (Π, Γ), $\Pi_x F_x^{-1}$ is independent of x. If this is the case, then one can then recover the maps Γ_{xy} by

$$(4.10) \qquad\qquad \Gamma_{xy} = F_x^{-1} \circ F_y \ .$$

The "correct" definition for the map Δ compatible with (3.11) and (4.8) is then given by

$$\Delta \mathbf{1} = \mathbf{1} \otimes \mathbf{1} \ , \qquad \Delta \Xi = \Xi \otimes \mathbf{1} \ , \qquad \Delta X = X \otimes X \ ,$$

and then recursively by

$$\Delta(\tau\bar{\tau}) = (\Delta\tau)(\Delta\bar{\tau}) \ , \qquad \Delta\mathcal{I}(\tau) = (\mathcal{I} \otimes \mathrm{id})\Delta\tau + \sum_k \frac{X^k}{k!} \otimes \mathscr{I}_k(\tau) \ .$$

REMARK 4.1. This definition shows that the convention $\mathcal{I}(X^k) = 0$ is compatible with the convention $\mathscr{I}_\ell(X^k) = 0$.

The identities (4.9) and (4.10) and the definition of f_x suggest that one should take for G the set of all linear maps of the type

$$\Gamma_f \tau = (\mathrm{id} \otimes f)\Delta\tau \ ,$$

where f is a multiplicative linear functional on T^+. Before we show that G does indeed form a group, we argue that if Π_x is defined recursively as above and F_x is given by (4.9), then there exists a single linear map $\mathbf{\Pi} \colon T \to \mathcal{S}'$ such that

$$(4.11) \qquad\qquad \Pi_x \tau = \mathbf{\Pi} F_x \tau \qquad \forall \tau \in T \ .$$

We can simply exhibit $\mathbf{\Pi}$ explicitly. Set

$$(4.12) \qquad\qquad \big(\mathbf{\Pi}\Xi\big)(x) = \xi(x) \ , \qquad \big(\mathbf{\Pi}X^k\big)(x) = x^k \ ,$$

and then recursively

$$(4.13) \qquad\qquad \mathbf{\Pi}\tau\bar{\tau} = \mathbf{\Pi}\tau \cdot \mathbf{\Pi}\bar{\tau} \ , \qquad \mathbf{\Pi}\mathcal{I}\tau = K * \mathbf{\Pi}\tau \ .$$

Note that this is very similar to the definition of $\mathscr{L}(\xi)$, with the notable exception that (3.11) is replaced by the more "natural" identity $\mathbf{\Pi}\mathcal{I}\tau = K * \mathbf{\Pi}\tau$. It is then a simple exercise in binomial identities to show that indeed $\Pi_x X^k = \mathbf{\Pi}F_x X^k$ and that, assuming that (4.11) holds for some τ and that f_x is given by (4.8), one also has $\Pi_x \mathcal{I}(\tau) = \mathbf{\Pi}F_x \mathcal{I}(\tau)$. Finally, since all relevant objects are multiplicative, one can see that if (4.11) holds for symbols τ and $\bar{\tau}$, then it must also hold for their product $\tau\bar{\tau}$.

---

[2]This definition differs from that in [**Hai14**] by a simple change of basis in T^+ and, while the inclusion of the sum over k in (4.8) may not appear very natural here, it leads to more natural expressions for the maps Δ and Δ^+ below.

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

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

and then recursively by
(4.14)

$$\Delta^+(\sigma\bar\sigma) = (\Delta^+\sigma)(\Delta^+\bar\sigma) \,, \qquad \Delta^+ \mathscr{I}_\ell(\tau) = (\mathscr{I}_\ell \otimes \mathrm{id})\Delta\tau + \sum_k \frac{X^k}{k!} \otimes \mathscr{I}_{\ell+k}(\tau) \,,$$

then it can be verified that this map intertwines with Δ via the relations

(4.15) $\qquad (\Delta \otimes I)\Delta = (I \otimes \Delta^+)\Delta \,, \qquad (\Delta^+ \otimes I)\Delta^+ = (I \otimes \Delta^+)\Delta^+ \,.$

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

$$(f \circ g)(\sigma) = (f \otimes g)\Delta^+\sigma \,.$$

If we furthermore denote by Γ_f the operator T associated to any such linear functional as in (4.9), the first identity of (4.15) yields the identity $\Gamma_f\Gamma_g = \Gamma_{f\circ g}$. The first identity of (4.14) furthermore ensures that if f and g are both multiplicative in the sense that $f(\sigma\bar\sigma) = f(\sigma)f(\bar\sigma)$, then $f \circ g$ is again multiplicative. It also turns out that every multiplicative linear functional f admits a unique inverse $f^{-1} = \mathcal{A}f$ for some linear map $\mathcal{A} \colon T \to T$ such that $f^{-1} \circ f = f \circ f^{-1} = e$, where $e \colon T^+ \to \mathbf{R}$ maps every basis vector of the form (4.7) to zero, except for $e(\mathbf{1}) = 1$. The element e is neutral in the sense that Γ_e is the identity operator.

It is now natural to define the structure group G associated to T as the set of all multiplicative linear functionals on T^+, acting on T via (4.9). Furthermore, for any admissible model, one has the identity

$$\Gamma_{xy} = F_x^{-1}F_y = \Gamma_{\gamma_{xy}} \,, \qquad \gamma_{xy} = f_x^{-1} \circ f_y \,.$$

Returning to the relation between Π_x and $\mathbf{\Pi}$, we showed actually more, namely that the knowledge of $\mathbf{\Pi}$ and the knowledge of (Π, Γ) are equivalent. Indeed, on the one hand one has $\mathbf{\Pi} = \Pi_x F_x^{-1}$ and the map F_x can be recovered from Π_x by (4.8) and (4.9). On the other hand however, one also has of course $\Pi_x = \mathbf{\Pi}F_x^{-1}$ and, if we equip T with an adequate recursive structure (denote by $|\tau|_I$ the number of times $\mathcal{I}$ appears in the symbol τ), then it is possible to show that the determination of $F_x^{-1}\tau$ (and therefore of $\Pi_x\tau$) only requires knowledge of $f_x(\mathscr{I}_k(\sigma))$ (and therefore of $\Pi_x\sigma$) for symbols σ with $|\sigma|_I \le |\tau|_I - 1$.

Furthermore, the translation $(\Pi, \Gamma) \leftrightarrow \mathbf{\Pi}$ 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 $\mathbf{\Pi}$ are multiplicative. In the general case, the first identity in (4.13) may then of course fail to be true, even if $\mathbf{\Pi}\tau$ 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 $\mathbf{\Pi}$ is that there seems to be

no simple way of describing the topology given by Definition 3.3 in terms of $\mathbf{\Pi}$.

4.3. Renormalisation of the dynamical Φ^4_3 model. Combining Theorem 3.17 and Theorem 3.28, we see that the map

$$\Phi \mapsto \mathcal{P}\big(\Xi - \Phi^3\big) \,,$$

is a continuous map from $\mathcal{D}^\gamma$ into itself, provided that $\gamma > 1 + 2\kappa$. Actually, one obtains a regularity improvement in the sense that it maps $\mathcal{D}^\gamma$ into $\mathcal{D}^{\bar\gamma}$ for some $\bar\gamma > \gamma$. In way reminiscent of usual parabolic PDE techniques, one can leverage this regularity improvement to build spaces $\mathcal{D}^{\gamma,\eta}$ (essentially weighted versions of the spaces $\mathcal{D}^\gamma$ which allow coefficients to become singular near times 0) such that the following holds.

THEOREM 4.2. *For every $\alpha > -\frac{2}{3}$ there exist exponents γ and η such that, for every $\Phi_0 \in \mathcal{C}^\alpha$ and every admissible model, (4.2) admits a unique local solution in $\mathcal{D}^{\gamma,\eta}$. Furthermore, this solution depends continuously on both Φ_0 and the underlying model.*

REMARK 4.3. One hits some minor technical difficulties in the actual definition of local solutions to (4.2), but these are resolved in [**Hai14**]. One can also show that solutions can be continued either forever or up to a time where the $\mathcal{C}^\alpha$ solution of the solution explodes. In this particular instance, one can actually show that solutions do not blow up, as shown in [**MW15**].

Given a model of the type $\mathscr{L}(\xi)$ constructed as in the previous subsection for a continuous space-time function ξ, it then follows from (4.6b) and the admissibility of $\mathscr{L}(\xi)$ 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 of course as all the functions to which $\mathcal{R}$ is applied belong to $\mathcal{D}^\gamma$ for some $\gamma > 0$ so that $\mathcal{R}f$ is uniquely defined by (3.4). As a consequence, applying the reconstruction operator $\mathcal{R}$ to both sides of (4.2), we see that if Φ solves (4.2) then $\mathcal{R}\Phi$ solves the integral equation (4.1), which thus yields the unique classical solution to (Φ^4). Note however that this identification of solutions to (4.2) with solutions to (Φ^4) relies crucially on the identity $\mathcal{R}(\Phi^3) = (\mathcal{R}\Phi)^3$. This identity has no reason whatsoever to hold for a generic admissible model! In particular, while (4.2) makes sense for *any* admissible model, we do not know in general whether its solutions still solve a local PDE.

At this stage, the situation is as follows. For any *continuous* realisation ξ of the driving noise, we have factored the solution map $(\Phi_0, \xi) \to \Phi$ associated to (Φ^4) into maps

$$(\Phi_0, \xi) \to \big(\Phi_0, \mathscr{L}(\xi)\big) \to \Phi \to \mathcal{R}\Phi \,,$$

where the middle arrow corresponds to the solution to (4.2) built in Theorem 4.2. 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. We also hope that the reader is convinced that a similar construction can be performed for a very large class of semilinear stochastic PDEs. In particular, the KPZ equation can also be analysed in this framework.

Given this construction, one is lead naturally to the following question: given a sequence ξ_ε of "natural" regularisations of space-time white noise, do the lifts $\mathscr{L}(\xi_\varepsilon)$ converge in probably in a suitable space of admissible models? Unfortunately, unlike in the case of the theory of rough paths where this is very often the case (but see [**Hai13**] for an example where it fails there too), the answer to this question in the context of SPDEs is often an emphatic *no*. Indeed, if it were the case for the dynamical Φ_3^4 model, then one could have chosen the constant C_ε to be independent of ε in (1.2), which is certainly not the case.

The way in which we are able to circumvent 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 a model does not impose the identity (4.6b), 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 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 $\mathscr{L}(\xi_\varepsilon)$ 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 no *not* preserve the image of $\mathscr{L}$ in the space of admissible models. As a consequence, when solving the fixed point map (4.2) 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 however that this is again the case and the modified equation is precisely given by (1.2), where C_ε is some linear combination of the two 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?

4.4. The renormalisation group. In order to build the group $\mathfrak{R}$, it turns out to be appropriate to describe its action first at the level of $\mathbf{\Pi}$ rather than at the level of (Π, Γ). At this stage we note that if ξ happens to be a stationary stochastic process and $\mathbf{\Pi}$ is built from ξ by (4.12) and (4.13), then $\mathbf{\Pi}\tau$ is also a stationary stochastic process for every $\tau \in T$. In order to define $\mathfrak{R}$, it is 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 $\mathbf{\Pi}$, the only remaining operation is to form linear combinations between them. It is therefore natural to describe elements of $\mathfrak{R}$ by linear maps $M \colon T \to T$ and to postulate their action on admissible models by $\mathbf{\Pi} \mapsto \mathbf{\Pi}^M$ with

$$\mathbf{\Pi}^M \tau = \mathbf{\Pi} M \tau .$$

It is not clear *a priori* whether given such a map M and an admissible model (Π, Γ) there is a coherent way of building a new model (Π^M, Γ^M) such that $\mathbf{\Pi}^M$ is the map associated to (Π^M, Γ^M) as above. It turns out that one has the following statement:

PROPOSITION 4.4. *In the above context, for every linear map $M \colon T \to T$ commuting with $\mathcal{I}$ and multiplication by X^k, there exist unique linear maps $\Delta^M \colon T \to T \otimes T^+$ and $\hat{\Delta}^M \colon T^+ \to T^+ \otimes T^+$ such that if we set*

$$\Pi_x^M \tau = \left(\Pi_x \otimes f_x \right) \Delta^M \tau , \qquad \gamma_{xy}^M(\sigma) = (\gamma_{xy} \otimes f_x) \hat{\Delta}^M \sigma ,$$

then Π_x^M satisfies again (3.11) and the identity $\Pi_x^M \Gamma_{xy}^M = \Pi_y^M$.

At this stage it may look like *any* linear map $M \colon T \to T$ commuting with $\mathcal{I}$ and multiplication by X^k yields a transformation on the space of admissible models by Proposition 4.4. This however is not true since we have completely disregarded the *analytical* bounds that every model has to satisfy. It is clear from Definition 3.3 that these are satisfied in general if and only if, for every symbol τ, $\Pi_x^M \tau$ is a linear combination of the $\Pi_x \bar{\tau}$ only involving symbols $\bar{\tau}$ with $|\bar{\tau}| \geq |\tau|$. This suggests the following definition.

DEFINITION 4.5. The renormalisation group $\mathfrak{R}$ consists of the set of linear maps $M \colon T \to T$ commuting with $\mathcal{I}$ and with multiplication by X^k, such that for $\tau \in T_\alpha$ and $\sigma \in T_\alpha^+$, one has

(4.16) $\Delta^M \tau - \tau \otimes \mathbf{1} \in T_{>\alpha} \otimes T^+ , \qquad \hat{\Delta}^M \sigma - \sigma \otimes \mathbf{1} \in T_{>\alpha}^+ \otimes T^+ .$

Its action on the space of admissible models is given by Proposition 4.4.

REMARK 4.6. It turns out that the second condition of (4.16) is actually a consequence of the first one, see the appendix in [**HQ15**].

4.5. The renormalised equations. In the case of the dynamical Φ_3^4 model considered in this article, it turns out that in general (i.e. if we also want to cover the situation mentioned in the introduction where ξ is approximated by a rescaled non-Gaussian stationary process) we need a five-parameter subgroup of $\mathfrak{R}$ to renormalise the equations. More precisely, we

consider elements $M \in \mathfrak{R}$ of the form $M = \exp(-\sum_{i=1}^{5} C_i L_i)$, where the generators L_i are determined by the substitution rules

$$L_1 \colon \mathsf{V} \mapsto \mathbf{1} \,, \quad L_2 \colon \mathsf{V} \mapsto \mathbf{1} \,, \quad L_3 \colon \mathsf{V} \mapsto \mathbf{1} \,, \quad L_4 \colon \mathsf{V} \mapsto \mathbf{1} \,, \quad L_5 \colon \mathsf{V} \mapsto \mathbf{1} \,.$$

This should be understood in the sense that if τ is an arbitrary formal expression, then $L_1 \tau$ is the sum of all formal expressions obtained from τ by performing a substitution of the type $\mathsf{V} \mapsto 1$, and similarly for the other L_i. For example, one has

$$L_1 \mathsf{V} = 3{\uparrow} \,, \qquad L_1 \mathsf{V} = {\searrow} \,, \qquad L_3 \mathsf{V} = 3{\uparrow} \,, \qquad L_1 {\uparrow} = 0 \,.$$

The convention $\mathcal{I}(\mathbf{1}) = 0$ also suggests that one should set for example $L_2 \mathsf{V} = 0$, since the symbol $\searrow$, which is what the substitution $\mathsf{V} \mapsto \mathbf{1}$ creates, would contain a factor of $\mathcal{I}(\mathbf{1})$. It is easy to see that all of the L_i commute with each other so that, as an abstract group, the transformations we consider form a copy of $\mathbf{R}^5$ with addition as its group operation. This is not a general fact however. One then has the following result:

PROPOSITION 4.7. *The linear maps M of the type just described belong to $\mathfrak{R}$. Furthermore, if (Π, Γ) is an admissible model such that $\Pi_x \tau$ is a continuous function for every $\tau \in T$, then one has the identity*

$$(4.17) \qquad\qquad \left(\Pi_x^M \tau\right)(x) = \left(\Pi_x M \tau\right)(x) \,.$$

REMARK 4.8. Note that it is the same value x that appears twice on each side of (4.17). It is in fact *not* the case that one has $\Pi_x^M \tau = \Pi_x M \tau$! However, the identity (4.17) is all we need to derive the renormalised equations.

It is now rather straightforward to show the following:

PROPOSITION 4.9. *Let M be as above and let (Π^M, Γ^M) be the model obtained by acting with M on the canonical model $(\Pi, \Gamma) = \mathscr{L}(\xi)$ for some smooth function ξ. Let furthermore Φ be the solution to (4.2) with respect to the model (Π^M, Γ^M). Then, the function $u(t, x) = \left(\mathcal{R}^M \Phi\right)(t, x)$ solves the equation*

$$\partial_t u = \Delta u - u^3 + (3C_1 - 9C_3 - 6C_5)u - (C_2 + 3C_4) + \xi \,.$$

PROOF. By Theorem 3.17, it turns out that (4.2) can be solved locally in $\mathcal{D}^\gamma$ (or rather a weighted version of this space taking into account possible blowup near time 0) as soon as γ is a little bit greater than 1. Therefore, we only need to keep track of its solution Φ up to terms of homogeneity 1. By repeatedly applying the identity (4.3), we see that the solution Φ is necessarily of the form

$$(4.18) \qquad\qquad \Phi = {\uparrow} + \varphi\,\mathbf{1} - \mathsf{V} - 3\varphi\,\mathsf{Y} + \langle \nabla\varphi, X \rangle \,,$$

for some real-valued function φ and some $\mathbf{R}^3$-valued function $\nabla\varphi$. We emphasise that $\nabla\varphi$ is treated as an independent function here, we certainly do not suggest that the function φ is differentiable! Our notation is only by

analogy with the classical Taylor expansion. Similarly, the right hand side of the equation is given up to order 0 by

(4.19)
$$\Xi - \Phi^3 = \Xi - \mathcal{V} - 3\varphi \vee + 3\mathcal{V} - 3\varphi^2 \uparrow + 6\varphi \mathcal{V} + 9\varphi \mathcal{V} - 3\langle \nabla\varphi, \vee X\rangle - \varphi^3 \mathbf{1} \;.$$

Combining this with the definition of M, it is straightforward to see that, modulo terms of strictly positive homogeneity, one has

$$
\begin{aligned}
M(\Xi - \Phi^3) = {}& \Xi - (M\Phi)^3 + 3C_1\uparrow + C_2\mathbf{1} + 3C_1\varphi\mathbf{1} - 9C_3\uparrow - 3C_4\mathbf{1} \\
& - 6C_5\uparrow - 6C_5\varphi\mathbf{1} - 9C_3\varphi\mathbf{1} \\
= {}& \Xi - (M\Phi)^3 + (3C_1 - 9C_3 - 6C_5)M\Phi + (C_2 - 3C_4)\mathbf{1} \;.
\end{aligned}
$$

Combining this with (4.17) and applying as before the reconstruction operator $\mathcal{R}^M$ to both sides of (4.2), the claim now follows at once. $\qquad\square$

4.6. Convergence of the renormalised models for Gaussian approximations. We now argue that if $\xi_\varepsilon = \varrho_\varepsilon * \xi$ as in Theorem 1.1, then one expects to be able to find constants $C_1^{(\varepsilon)}$ and $C_3^{(\varepsilon)}$ (and set $C_i^{(\varepsilon)} = 0$ for $i \in \{2,4,5\}$) such that the sequence of renormalised models $M^\varepsilon \mathscr{L}(\xi_\varepsilon)$ defined as in the previous subsection 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 before. Since there are general arguments available to deal with all the expressions τ of positive homogeneity, we restrict ourselves to those of negative homogeneity which, leaving out Ξ which is easy to treat, are given by

$$\mathcal{V}, \; \vee, \; \mathcal{V}, \; \uparrow, \; \mathcal{V}, \; \mathcal{V}, \; X_i\vee \;.$$

REMARK 4.10. Even if we can show that $\hat{\boldsymbol{\Pi}}^\varepsilon \tau$ converges weakly to some limit as $\varepsilon \to 0$, this does not necessarily imply that the corresponding sequence of models converges in their natural topology. We will not try to address convergence in the correct topology here, although this is a nontrivial problem.

For this section, some elementary notions from the theory of Wiener chaos expansions are required, but we will try to hide this as much as possible. Recall that $\boldsymbol{\Pi}^\varepsilon\uparrow = K * \xi_\varepsilon = K_\varepsilon * \xi$, where the kernel K_ε is given by $K_\varepsilon = K * \varrho_\varepsilon$. This shows that, for $\varepsilon > 0$, one has

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

Here, we make use of the fact that, for any two distributions ζ and η, the product $\zeta \cdot \eta$ is well-defined as a distribution on the product space. It is only when one tries to define a pointwise product yielding again a distribution on the *same* space that one runs into trouble. 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

$$(4.20) \qquad I_k(f) = \int \cdots \int f(z_1, \ldots, z_k) \, \xi(z_1) \cdots \xi(z_k) \, dz_1 \cdots dz_k \; .$$

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 would then allow to extend (4.20) from smooth test functions to all functions $f \in L^2$. It is unfortunately *not* the case that one has such an isometry. Instead, one should first replace the products in (4.20) by *Wick products*, which are obtained by considering all possible *contractions* of the type

$$\xi(z_i)\xi(z_j) \mapsto \, :\!\xi(z_i)\,\xi(z_j)\!: + \, \delta(z_i - z_j) \; .$$

For example, the distribution $:\!\xi(z_i)\,\xi(z_j)\,\xi(z_k)\!:$ is defined by the identity

$$\xi(z_i)\xi(z_j)\xi(z_k) = \, :\!\xi(z_i)\xi(z_j)\xi(z_k)\!: + \, \xi(z_i)\delta(z_j - z_k)$$
$$+ \, \xi(z_j)\delta(z_k - z_i) + \xi(z_k)\delta(z_i - z_j) \; .$$

(See Section 5.1 below for a more general definition which covers Gaussian random variables as a special case.) If we then set

$$\hat{I}_k(f) = \int \cdots \int f(z_1, \ldots, z_k) \, :\!\xi(z_1) \cdots \xi(z_k)\!: \, dz_1 \cdots dz_k \; ,$$

which is well-defined for all smooth test functions f, one recovers indeed the Itô isometry

$$(4.21) \qquad \mathbf{E} \hat{I}_k(f) \hat{I}_k(g) = \langle f^{\mathrm{sym}}, g^{\mathrm{sym}} \rangle \; .$$

We refer to [**Nua95**] for a more thorough description of this construction, which also goes under the name of *Wiener chaos*, with random variables of the type $\hat{I}_k(f)$ said to belong to the chaos of kth order. While chaoses of different order are not independent, they are orthogonal in the sense that $\mathbf{E} \hat{I}_k(f) \hat{I}_\ell(g) = 0$ if $k \neq \ell$.

One very nice property is that 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

$$(4.22) \qquad \mathbf{E} |\hat{I}_k(f)|^p \leq C_{k,p} \|f^{\mathrm{sym}}\|_{L^2}^p \leq C_{k,p} \|f\|_{L^2}^p \; ,$$

where the second bound comes from the fact that symmetrisation is a contraction in L^2. In other words, the pth moment of a random variable belonging to a Wiener chaos of fixed order can be bounded by the corresponding power of its second moment. The reason why such a bound is very useful is the following Kolmogorov-type result:

THEOREM 4.11. *Let ξ be a distribution-valued random variable such that there exists $\alpha \leq 0$ such that, for every $p > 0$, the bound*

$$(4.23) \qquad \mathbf{E} |\xi(\varphi_x^\lambda)|^p \lesssim \lambda^{\alpha p} \; ,$$

holds uniformly over $\varphi \in \mathcal{B}_r$ for some $r > |\alpha|$, all $\lambda \in (0, 1]$, and locally uniformly in x. Then, for every $\beta < \alpha$, there exists a $\mathcal{C}^\beta$-valued version of ξ.

REMARK 4.12. Here, we say that $\tilde{\xi}$ is a version of ξ if, for every smooth test function φ, $\tilde{\xi}(\varphi) = \xi(\varphi)$ almost surely.

REMARK 4.13. There is an analogous result allowing to show the convergence of a sequence of models to a limiting model in the topology given by Definition 3.3, see [**Hai14**, Thm 10.7] for more details.

REMARK 4.14. As usual for Kolmogorov-type results, one can trade integrability for regularity in the sense that if (4.23) only holds for some fixed value $p > 2$, then we can only conclude that ξ has a version in some $\mathcal{C}^\beta$ for $\beta < \alpha - \frac{d}{p}$, where d denotes the scaling dimension of the underlying space. In our particular case one has $d = 5$ since we consider space-time distributions and "time counts double".

Returning to our problem, we first argue that it should be possible to choose M in such a way that $\hat{\mathbf{\Pi}}^\varepsilon \mathbb{V}$ converges to a limit as $\varepsilon \to 0$. The above considerations suggest that one should rewrite $\mathbf{\Pi}^\varepsilon \mathbb{V}$ as
(4.24)
$$\left(\mathbf{\Pi}^\varepsilon \mathbb{V}\right)(z) = \left(K * \xi_\varepsilon\right)(z)^2 = \int\!\!\int K_\varepsilon(z - z_1) K_\varepsilon(z - z_2) \,{:}\xi(z_1)\,\xi(z_2){:}\, dz_1\, dz_2 + C_\varepsilon\ ,$$

where the constant C_ε is given by

(4.25)
$$C_\varepsilon = \int K_\varepsilon^2(z_1)\, dz_1 = \int K_\varepsilon^2(z - z_1)\, dz_1\ .$$

At this stage, it is convenient to introduce a graphical notation "à la Feynman" for these multiple integrals since expressions otherwise rapidly became unwieldy. Similarly to [**Hai14, HP14, HS15**], we denote dummy integration variables by black dots, we write ● for the distinguished variable z, ○ for a Wick factor of the type $\xi(z_i)$, ----→ for the kernel K_ε evaluated at the difference between the variables representing its two endpoints, and we implicitly assume that all variables except for ● are integrated out. With this notation, (4.24) and (4.25) can be expressed in a much friendlier and more compact way as

(4.26)
$$\mathbf{\Pi}^\varepsilon \mathbb{V} = \text{} + \text{}\ , \qquad C_\varepsilon = \text{}\ .$$

This notation is also compatible with the graphical notation we are using for the various symbols: the graphical notation for $\mathbf{\Pi}^\varepsilon \tau$ is essentially the same as that for τ but, as a consequence of the definition of Wick products, one should sum over all possible graphs obtained by pairwise contractions of the noises ○.

Note now that K_ε is an ε-approximation of the kernel K which has the same singular behaviour as the heat kernel. In terms of the parabolic

distance, the singularity of the heat kernel scales like $K(z) \sim |z|^{-3}$ 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 heat kernel is bounded by $t^{-3/2}$.) This suggests that one has $K_\varepsilon^2(z) \sim |z|^{-6}$ for $|z| \gg \varepsilon$. Since parabolic space-time has scaling dimension 5 (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 correctly suggests that one has

$$C_\varepsilon \sim \frac{1}{\varepsilon} \; .$$

This already shows that one should not expect $\mathbf{\Pi}^\varepsilon \mathbb{V}$ to converge to a limit as $\varepsilon \to 0$. Indeed, as we will see presently, it turns out that the first term in (4.24) converges to a distribution-valued stationary space-time process, so that one would like to somehow get rid of this diverging constant C_ε. This is exactly where the renormalisation map M^ε (in particular the factor $\exp(-C_1^{(\varepsilon)} L_1)$) enters into play. Following the above definitions, we see that one has

$$\left(\hat{\mathbf{\Pi}}^\varepsilon \mathbb{V}\right)(z) = \left(\mathbf{\Pi}^\varepsilon M \mathbb{V}\right)(z) = \left(\mathbf{\Pi}^\varepsilon \mathbb{V}\right)(z) - C_1^{(\varepsilon)} \; .$$

This suggests that if we make the choice $C_1^{(\varepsilon)} = C_\varepsilon$, then $\hat{\mathbf{\Pi}}^\varepsilon \mathbb{V}$ does indeed converge to a non-trivial limit as $\varepsilon \to 0$. Writing $\longrightarrow$ for the kernel K, this limit is a distribution which can formally be described as

$$\mathbf{\Pi}^\varepsilon \mathbb{V} = \bigvee \; ,$$

with the implicit understanding that $\left(\mathbf{\Pi}^\varepsilon \mathbb{V}\right)(z)$ does not make sense as a random variable, but must first be integrated against a smooth test function in order to produce a random variable belonging to the second Wiener chaos. Using the scaling properties of the kernel K, it is not too difficult to show that this procedure works, and that the resulting random variables satisfy a bound of the type considered in Theorem 4.11 with $\alpha = -1$. Once we know that $\hat{\mathbf{\Pi}}^\varepsilon \mathbb{V}$ converges, it is immediate that $\hat{\mathbf{\Pi}}^\varepsilon X \mathbb{V}$ converges as well, since this amounts to just multiplying a distribution by a smooth function.

A similar argument to what we did for $\mathbb{V}$ allows to take care of $\tau = \mathbb{W}$ since one then has

$$\left(\mathbf{\Pi}^\varepsilon \mathbb{W}\right)(z) = \bigvee + 3 \; \bigcirc \; .$$

Noting that the second term in this expression is nothing but $3C_\varepsilon\left(\mathbf{\Pi}^\varepsilon \uparrow\right)(z)$, we see that in this case, provided again that $C_1^{(\varepsilon)} = C_\varepsilon$, $\hat{\mathbf{\Pi}}^\varepsilon \mathbb{W}$ is given by only the first term in the expression above, which turns out to converge to a non-degenerate limiting random distribution in a similar way to what happened for $\mathbb{V}$, but this time the Kolmogorov-type bound only holds for $\alpha = -\frac{3}{2}$.

Going down our list of terms of negative homogeneity, we see that it remains to consider Ψ, Ψ, and Ψ. It turns out that the last one is the most difficult, so we only discuss that one. The explicit expression for $\mathbf{\Pi}^\varepsilon \Psi$ is given in our graphical notation by

$$\mathbf{\Pi}^\varepsilon \Psi = \quad + 6 \quad + 3 \quad + $$

$$+ 3 \quad + 6 \quad + 6 \quad .$$

It turns out that *all* of these terms, except for the first two, are divergent as $\varepsilon \to 0$, so we have to hope that our definition of $\hat{\mathbf{\Pi}}^\varepsilon$ creates sufficiently many cancellations to take care of them! A simple calculation using our definition of M^ε shows that one has the identity

$$\hat{\mathbf{\Pi}}^\varepsilon \Psi = \mathbf{\Pi}^\varepsilon \Psi - 3C_1^{(\varepsilon)} \mathbf{\Pi}^\varepsilon \Psi - C_1^{(\varepsilon)} \mathbf{\Pi}^\varepsilon \Psi + 3(C_1^{(\varepsilon)})^2 \mathbf{\Pi}^\varepsilon \Psi - 3C_3^{(\varepsilon)} \mathbf{\Pi}^\varepsilon \Psi \, .$$

(Recall that we have set $C_i^{(\varepsilon)} = 0$ for $i \notin \{1,3\}$.) Inserting the definition of $\mathbf{\Pi}^\varepsilon$ into this expression and making use of the identity $C_1^{(\varepsilon)} = C_\varepsilon$ with C_ε given by (4.26), we see that many terms do indeed cancel out, finally yielding

$$(4.27) \qquad \hat{\mathbf{\Pi}}^\varepsilon \Psi = \quad + 6 \quad + 6 \quad - 3C_3^{(\varepsilon)} \quad .$$

This is still problematic: the penultimate term in this expression contains the kernel

$$(4.28) \qquad Q_\varepsilon = \quad .$$

(This is a slight abuse of our notation: the two variables $\bullet$ are fixed and Q_ε is evaluated at their difference.) As $\varepsilon \to 0$, Q_ε converges to a kernel Q, which has a non-integrable singularity at the origin, thus preventing the corresponding term to converge to a limit.

This is akin to the problem of making sense of integration against a one-dimensional kernel with a singularity of type $1/|x|$ at the origin. 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. In a way, the extra term can be interpreted as subtracting a Dirac distribution with an "infinite mass" located at the origin, thus cancelling out the divergence of the non-integrable singularity. 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 $W_\varepsilon \sim 1/\varepsilon$ otherwise), then $\mathscr{R}W^\varepsilon \to \mathscr{R}W$ in a distributional sense, and (using the usual correspondence between functions and distributions) one has

$$\mathscr{R}W^\varepsilon = W^\varepsilon - \hat{C}_\varepsilon \delta_0 \; , \qquad \hat{C}_\varepsilon = \int W^\varepsilon(x)\, dx \; .$$

The cure to the problem we are facing for showing the convergence of $\boldsymbol{\Pi}^\varepsilon \Psi$ is virtually identical. Indeed, by choosing

$$C_3^{(\varepsilon)} = 2 \;\; \raisebox{-0.5em}{} \;\; ,$$

(which is a constant as it does not depend on $\bullet$ by translation invariance) the term in the first homogeneous Wiener chaos for $\hat{\boldsymbol{\Pi}}^\varepsilon \Psi$ is precisely given by

$$6 \int \big(\mathscr{R}\hat{Q}_\varepsilon * K_\varepsilon\big)(z - z_2)\xi(z_2)\, dz_2 \; .$$

It turns out that the convergence of $\mathscr{R}\hat{Q}_\varepsilon$ to a limiting distribution $\mathscr{R}\hat{Q}$ takes place in a sufficiently strong topology to allow to conclude that $\hat{\boldsymbol{\Pi}}^\varepsilon \Psi$ does indeed converge to a non-trivial limiting random distribution.

It should be clear from this whole discussion that while the precise values of the constants $C_1^{(\varepsilon)}$ and $C_3^{(\varepsilon)}$ depend on the shape of the mollifier ϱ, 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 (4.2) and of the reconstruction operator $\mathcal{R}$ with respect to the underlying model, the conclusion of Theorem 1.1 follows.

5. Convergence of other smooth models to Φ_3^4

In this final section, we give a short overview of the ideas involved in showing that different kinds of smooth models (and not just the simplest one obtained by hitting the noise with a smooth mollifier) also converge to the same process.

5.1. Non-Gaussian approximations. We first discuss how the argument sketched in the previous section can be modified to deal with the case when ξ_ε does no longer denote some simple mollification of ξ, but instead denotes a suitable rescaling of an arbitrary stationary space-time stochastic process, i.e. one has

$$\xi_\varepsilon(t, x) = \varepsilon^{-\frac{5}{2}} \eta(t/\varepsilon^2, x/\varepsilon) \,,$$

for a smooth stationary process η on $\mathbf{R} \times \mathbf{R}^3$. As already mentioned in the introduction, it is then possible to find choices of constants $C_\varepsilon^{(i)}$ such that solutions to

$$\partial_t \Phi_\varepsilon = \Delta \Phi_\varepsilon + C_\varepsilon^{(1)} + C_\varepsilon^{(2)} \Phi_\varepsilon - \Phi_\varepsilon^3 + \xi_\varepsilon \,,$$

converge as $\varepsilon \to 0$ to the process Φ built in Theorem 1.1. In view of Proposition 4.9, this is again a consequence of the fact that the renormalised models $\hat{\mathbf{\Pi}}^\varepsilon$ built from ξ_ε as in Section 4.6 converge to the same limit as before, provided that the renormalisation constants $C_i^{(\varepsilon)}$ are suitably chosen. The purpose of this section is to show why all of the constants $C_i^{(\varepsilon)}$ with $i \in \{1, \ldots, 5\}$ should in general be set to non-zero values.

We essentially follow the arguments of [**HS15**] and we will completely ignore issues related to the fact that one wishes to consider the limiting equation on a torus rather than the whole space. We assume that the σ-fields generated by point evaluations of η in any two regions separated by a distance of at least 1 are independent. The kind of example for η one should have in mind is

$$\eta(z) = \int \varphi(z - \bar{z}) \, \mathbf{P}_\mu(d\bar{z}, d\varphi) \,,$$

where $\mathbf{P}_\mu$ denotes a compensated Poisson point process on $\mathbf{R}^4 \times \mathcal{C}_0^\infty$ with intensity measure $\mu(d\bar{z}, d\varphi) = d\bar{z} \, \nu(d\varphi)$, where ν is a finite measure charging only functions with support contained in the unit ball and satisfying suitable moment conditions to guarantee that η admits moments of all orders.

Since ξ_ε is no longer Gaussian, we no longer have Wiener chaos decomposition at our disposal. However, there exists an analogue to the Wick product for arbitrary collections of random variables, and this will be sufficient for our purpose. In order to introduce the Wick product, we first need to introduce joint cumulants. Given a collection of random variables $\mathcal{X} = \{X_\alpha\}_{\alpha \in A}$ for some index set A, and a subset $B \subset A$, we write $X_B \subset \mathcal{X}$ and X^B as shorthands for

$$X_B = \{X_\alpha \,:\, \alpha \in B\} \,, \qquad X^B = \prod_{\alpha \in B} X_\alpha \,.$$

Given a finite set B, we furthermore write $\mathcal{P}(B)$ for the collection of all partitions of B, i.e. if B is non-empty, then $\mathcal{P}(B)$ consists of all sets $\pi \subset \mathscr{P}(B)$ (the set of subsets of B) such that $\bigcup \pi = B$, $\emptyset \in \pi$, and such that any two distinct elements of π are disjoint. With this definition, we in particular have $\mathcal{P}(\emptyset) = \{\{\emptyset\}\}$. The joint cumulant $\mathbf{E}_c(X_B)$ for a collection X_B of random variables is then given by the following defining property.

DEFINITION 5.1. Given a collection $\mathcal{X}$ of random variables as above and any non-empty finite set $B \subset A$, one has

$$(5.1) \qquad \mathbf{E}(X^B) = \sum_{\pi \in \mathcal{P}(B)} \prod_{\bar{B} \in \pi} \mathbf{E}_c(X_{\bar{B}}) .$$

Note that this definition naturally enforces the convention that $\mathbf{E}_c(X_\emptyset) = 1$. We can now define the Wick product of an arbitrary finite collection of random variables:

DEFINITION 5.2. The Wick product $:X_B:$ for $B \subset A$ is defined recursively by

$$(5.2) \qquad X^A = \sum_{B \subset A} :X_B: \sum_{\pi \in \mathcal{P}(A \backslash B)} \prod_{\bar{B} \in \pi} \mathbf{E}_c(X_{\bar{B}}) .$$

Similarly to above, this enforces the natural convention that $\mathbf{E}:X_\emptyset: = 1$. In other words, to turn a product of random variables into a sum of Wick products, one now not only considers all possible ways of replacing pairs of random variables by their covariance, but all possible ways of replacing subsets of random variables by their joint cumulants.

Denote now by $\kappa_p^{(\varepsilon)}$ the pth joint cumulant of ξ_ε, namely
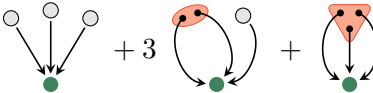
$$\kappa_p^{(\varepsilon)}(z_1, \ldots, z_p) = \mathbf{E}_c\big(\{\xi_\varepsilon(z_1), \ldots, \xi_\varepsilon(z_p)\}\big) .$$

By translation invariance, this function depends only on the differences between its p arguments. Furthermore, it satisfies the scaling relation
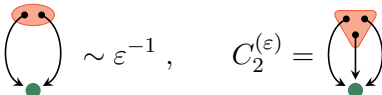
$$\kappa_p^{(\varepsilon)}(z_1, \ldots, z_p) = \varepsilon^{-\frac{5p}{2}} \kappa_p(S_\varepsilon z_1, \ldots, S_\varepsilon z_p) ,$$

where κ_p denotes the pth joint cumulant of η and the scaling operator S_ε acts on $\mathbf{R}^4$ as $S_\varepsilon(t, x) = (t/\varepsilon^2, x/\varepsilon)$. We then use the same graphical notation as in [**HS15**] for these cumulants: a red p-gon for instance for $p = 4$, represents the cumulant function $\kappa_p^{(\varepsilon)}(z_1, \ldots, z_p)$, with z_i given by the p integration variables represented as before by the p black dots. This time furthermore, $\circ$ denotes a Wick factor of ξ_ε instead of ξ.

As a consequence of Definitions 5.1 and 5.2, we then have for example the identity

$$(\mathbf{\Pi}^\varepsilon \mathbb{V})(z) = \quad \text{} \quad + 3 \quad \text{} \quad + \quad \text{} .$$

A powercounting argument shows that this additional term also diverges as $\varepsilon \to 0$, which suggests that in this case we should no longer set $C_2^{(\varepsilon)} = 0$ as in the Gaussian case but instead we should choose the relevant renormalisation constants as follows:

$$(5.3) \qquad C_1^{(\varepsilon)} = \quad \text{} \quad \sim \varepsilon^{-1} , \qquad C_2^{(\varepsilon)} = \quad \text{} \quad \sim \varepsilon^{-3/2} .$$

This choice of renormalisation constant then yields just as in the Gaussian case

$$\left(\hat{\mathbf{\Pi}}^{\varepsilon}\mathbb{V}\right)(z) = \quad\raisebox{-1em}{}\quad,$$

which can then be shown to converge to a finite limiting distribution as $\varepsilon \to 0$.

REMARK 5.3. The fact that $\hat{\mathbf{\Pi}}^{\varepsilon}\mathbb{V}$ is tight in some space of distributions as $\varepsilon \to 0$ can be shown quite easily by first testing against a test function and then bounding its second moment. Tightness in the correct topology and identification of the correct limit is more delicate and requires control on moments of all orders. Unfortunately, these can no longer be obtained as a corollary of a second moment bound since the analogue to (4.22) does not hold anymore in general.

We will henceforth ignore these difficulties and admit without proof the fact that if f is an L^2 kernel of k variables and f_ε is a sequence of "sufficiently nice" kernels converging to f, then the sequence of random variables

$$\int f_\varepsilon(z_1, \ldots, z_k) :\xi_\varepsilon(z_1)\cdots\xi_\varepsilon(z_k): dz_1\cdots dz_k ,$$

converges weakly to $I_k(f)$ as $\varepsilon \to 0$. This convergence furthermore holds jointly for any finite collection of such random variables. See [**HS15**] for more details.

The necessity of the remaining renormalisation constants can again be traced back to the analysis of $\hat{\mathbf{\Pi}}^{\varepsilon}\mathbb{V}$. This time, one has the much lengthier identity

At this stage, it starts to become clear that on should in principle start to use a more systematic approach, but for the sake of this article, and in order to give a feeling for the kind of bounds involved in the proofs, our more "brutal" approach is sufficient.

In order to bound $\mathbf{\Pi}^\varepsilon \Psi$ we first note that as a consequence of the fact that K integrates to 0, both the last term on the first line and the first term on the third line of the above expression vanish. This time, taking into account the fact that all of the $C_i^{(\varepsilon)}$ can be chosen to be non-zero, we obtain the identity

$$\hat{\mathbf{\Pi}}^\varepsilon \Psi = \mathbf{\Pi}^\varepsilon \Psi - 3C_1^{(\varepsilon)}\mathbf{\Pi}^\varepsilon \Psi - C_1^{(\varepsilon)}\mathbf{\Pi}^\varepsilon \Psi + 3(C_1^{(\varepsilon)})^2\mathbf{\Pi}^\varepsilon \Psi - 3C_3^{(\varepsilon)}\mathbf{\Pi}^\varepsilon \Psi$$
$$- C_4^{(\varepsilon)}\mathbf{\Pi}^\varepsilon \mathbf{1} - 2C_5^{(\varepsilon)}\mathbf{\Pi}^\varepsilon \Psi .$$

We claim that besides (5.3), the correct choice of renormalisation constants in order to obtain the same limiting model as in Section 4.6 is given by

With this choice, a tedious but straightforward graphical calculation yields the identity

(5.4)

The four terms appearing on the first line are direct analogues of the terms appearing in the Gaussian case (4.27). In view of Remark 5.3 it is therefore at least very plausible that they converge to the same limit. It remains to argue that the terms appearing on the last line vanish in the limit $\varepsilon \to 0$. For this, the following simple lemma (see [**HS15**, Lem. 4.7] for a slightly more general version) is useful:

LEMMA 5.4. *One has the bound*

$$\int \prod_{i=1}^{n} |K(z_i - \bar{z}_i)||\kappa_n^{(\varepsilon)}(\bar{z}_1, \ldots, \bar{z}_n)| \, d\bar{z}_1 \cdots d\bar{z}_n$$

$$\lesssim \varepsilon^{5(n/2-1)} \int_{\mathbf{R}^4} \prod_{i=1}^{n} \left(|z_i - \bar{z}| + \varepsilon\right)^{-3} d\bar{z} \,,$$

uniformly over $z_1, \ldots, z_n \in \mathbf{R}^4$ and $\varepsilon > 0$.

Recall that since we are in space dimension 3, both the heat kernel and its truncation K satisfy the bounds $|K(z)| \lesssim |z|^{-3}$, so that each factor appearing on the right hand side of this bound behaves essentially like a factor K_ε. In particular, writing

$$Q_\varepsilon^{(1)} = \quad, \qquad Q_\varepsilon^{(2)} = \quad, \qquad Q_\varepsilon^{(3)} = \quad,$$

with the same conventions as in (4.28), it follows that the kernels $Q_\varepsilon^{(i)}$ satisfy the bounds

$$|Q_\varepsilon^{(1)}(z)| \lesssim \varepsilon^\kappa |z|^{-\frac{9}{2}-\kappa}, \quad |Q_\varepsilon^{(2)}(z)| \lesssim \varepsilon^\kappa |z|^{-\frac{9}{2}-\kappa}, \quad |Q_\varepsilon^{(3)}(z)| \lesssim \varepsilon^\kappa(|z|+\varepsilon)^{-5-\kappa},$$

provided that one chooses $\kappa \geq 0$ sufficiently small. In particular, both $Q_\varepsilon^{(1)}$ and $Q_\varepsilon^{(2)}$ converge to 0 in L^p for p sufficiently small (but still greater than 1), which allows to show that the first two terms on the second line of (5.4) vanish in the limit. To show that the last two terms also vanish in the limit, one needs to exploit cancellations between them: taken separately they converge to finite non-vanishing limits. However, since $Q_\varepsilon^{(3)}$ is integrable with its integral remaining uniformly bounded as $\varepsilon \to 0$, it is easy to show that $Q_\varepsilon^{(3)}(\cdot) - C_{3,2}^{(\varepsilon)}\delta(\cdot)$ converges weakly to 0, where $C_{3,2}^{(\varepsilon)}$ denotes the second term appearing in the above definition of $C_3^{(\varepsilon)}$. While this in itself is not sufficient to guarantee that the sum of the last two terms in (5.4) indeed vanishes in the limit, it is a strong indication that it does, and is in fact not very difficult to show.

In a similar way, it is possible to show that $\hat{\mathbf{\Pi}}^\varepsilon \tau$ converges to the same limit as in Section 4.6 for every symbol $\tau \in \mathcal{W}$. With some additional effort, one can then show that this convergence actually takes place in the topology of the space of models for $(T, \mathcal{G})$ which, when combined with the continuity statement of Theorem 4.2, the identification of the renormalised solutions of Proposition 4.9, and the explicit expressions of the renormalisation constants, immediately implies that the solutions to (1.6) with constants as in (1.7) do indeed converge to the same limiting object as constructed in Theorem 1.1.

5.2. Non-cubic approximations. In this last section, we give a very short sketch of the main ideas appearing in the proof of Theorem 1.5. Recall that we are interested in the study of (1.3) as the effective potential $\langle V_\theta \rangle$ undergoes a pitchfork bifurcation. In order to concentrate on the main ideas, we restrict ourselves to the case where V is of degree 6. Then, normalising the equation such that $\partial_\theta \partial_x^2 \langle V_\theta \rangle(0) = -\frac{1}{2}$ and $\langle V_0 \rangle^{(4)}(0) = 6$, the derivative of the effective potential $\langle V_\theta \rangle$ is given to lowest order in θ by

$$(5.5) \qquad\qquad -\langle V_\theta \rangle'(\Phi) = \theta \Phi - \Phi^3 - a\Phi^5 \ ,$$

for some constant a. From now on we neglect terms of higher order in θ and we consider the case of a potential with derivative equal to (5.5).

Note that (5.5) is the derivative of the *effective* potential. After rescaling, the model (1.3) we consider is therefore given by

$$(5.6)$$
$$\partial_t \Phi_{\varepsilon,\theta} = \Delta \Phi_{\varepsilon,\theta} + \theta \varepsilon^{-1} \Phi - H_3(\Phi, C/\varepsilon) - a\varepsilon H_5(\Phi, C/\varepsilon) + \xi_\varepsilon$$
$$= \Delta \Phi_{\varepsilon,\theta} + \varepsilon^{-1}(\theta + 3C - 15aC^2)\Phi - (1 + 10aC)\Phi^3 - a\varepsilon\Phi^5 + \xi_\varepsilon \ ,$$

where C is the constant appearing in (1.4). In this equation, $H_n(x, c)$ denotes the nth Hermite polynomial with variance c, i.e. $H_3(x, c) = x^3 - 3cx$, etc. A special case of Theorem 1.5 is then given by, the following

THEOREM 5.5. *There exist values $a > 0$ and $b \in \mathbf{R}$, such that if one sets $\theta = a\varepsilon \log \varepsilon + b\varepsilon$, then the solution to (5.6) converges as $\varepsilon \to 0$ to the process Φ built in Theorem 1.1.*

The remainder of this section is devoted to a sketch of the proof of Theorem 5.5. It is not clear a priori how to fit (5.6) into the framework we developed in this article. One might think that one could try to change (4.2) into something like

$$(5.7) \qquad \Phi = \mathcal{P} \mathbf{1}_{t>0} \big(\Xi + c_1 \Phi - c_2 \Phi^3 - a\varepsilon \Phi^5 \big) + P\Phi_0 \ ,$$

for suitable choices of constants c_i, but one then immediately runs into two (related) problems:

- In order to interpret (5.7) as an equation in $\mathcal{D}^\gamma$ for some regularity structure, we would naturally have to replace the recursive step (4.4) by something like

$$\tau_i \in \mathcal{U} \quad \Rightarrow \quad \mathcal{I}(\tau_1 \tau_2 \tau_3 \tau_4 \tau_5) \in \mathcal{U} \ .$$

  The problem then is that the resulting collection of symbols no longer has the property that, for every $\gamma \in \mathbf{R}$, there are only finitely many symbols of homogeneity at most γ. This pretty much destroys the rest of the argument.
- The whole point of our construction was to build a solution map that no longer depends on ε and to hide all the singular ε-dependency of the solution in the convergence of a suitable sequence

of "models" to a limit. In the case of (5.7), we still have an ε-dependent fixed point problem. This is of course not a problem if its solution depends continuously on ε as $\varepsilon \to 0$, but this is precisely not expected here: Theorem 1.5 states that the effect of the quintic term is still felt in the limit since the constant appearing in front of the cubic term in the limit is given by the cubic term of the effective potential $\langle V_0 \rangle$ and not by that of V_0.

In order to circumvent these problems, the idea introduced in [**HQ15**] is to build a regularity structure which encodes the operation "multiplication by ε" as a non-trivial operation in $\mathcal{D}^\gamma$, thus allowing us again to set up an ε-independent fixed point problem and to shift all of the difficulties to that of proving that a suitable model converges to a limit.

For this, besides the symbols X, Ξ and the operator $\mathcal{I}$ used before, we introduce a new operator $\mathcal{E}$ and we consider instead of (5.7) a fixed point problem of the type

$$(5.8) \qquad \Phi = \mathcal{P}\mathbf{1}_{t>0}\big(\Xi - \Phi^3 - a\hat{\mathcal{E}}\Phi^5\big) + P\Phi_0 \;,$$

where $\hat{\mathcal{E}}$ will be an operator acting on $\mathcal{D}^\gamma$ built from $\mathcal{E}$ (in a way reminiscent of how $\mathcal{P}$ and $\mathcal{K}$ were built from $\mathcal{I}$). The recursive step (4.4) in the construction of the regularity structure is then replaced by

$$\tau_i \in \mathcal{U} \quad \Rightarrow \quad \mathcal{I}(\tau_1\tau_2\tau_3) \in \mathcal{U} \quad \& \quad \mathcal{I}(\mathcal{E}(\tau_1\tau_2\tau_3\tau_4\tau_5)) \in \mathcal{U} \;.$$

and the definition (4.5) of the collection of symbols $\mathcal{W}$ is replaced by

$$\mathcal{W} = \{\Xi\} \cup \{\tau_1\tau_2\tau_3 \,:\, \tau_i \in \mathcal{U}\} \cup \{\mathcal{E}(\tau_1\tau_2\tau_3\tau_4\tau_5) \,:\, \tau_i \in \mathcal{U}\} \;.$$

We also define a slightly larger collection of symbols

$$\mathcal{W}_{\mathrm{ex}} = \mathcal{W} \cup \{\tau_1\tau_2\tau_3\tau_4\tau_5 \,:\, \tau_i \in \mathcal{U}\} \;,$$

and we call similarly to before T and T_{ex} the linear spans of $\mathcal{W}$ and $\mathcal{E}_{\mathrm{ex}}$ respectively. In this construction, we consider the symbol $\mathcal{E}$ as an "integration operator of order 1" in the sense that we set $|\mathcal{E}(\tau)| = |\tau|+1$. This might seem strange at first sight: after all $\mathcal{E}$ is supposed to represent multiplication by the number ε, while the homogeneity is supposed to be associated to some kind of regularity, but we will see that this is actually quite natural.

The structure group $\mathcal{G}$ associated to the two graded spaces T and T_{ex} is then built in an analogous way to the construction of Section 4.2. The difference is that elements of T_+ are now not just polynomials in X and $\mathscr{I}_k(\tau)$ as in (4.7), but are also allowed to contain factors $\mathscr{E}_k(\tau)$, with the maps Δ and Δ^+ defined as above, but with the additional rule

$$\Delta\mathcal{E}(\tau) = (\mathcal{E} \otimes \mathrm{id})\Delta\tau + \sum_k \frac{X^k}{k!} \otimes \mathscr{E}_k(\tau) \;,$$

and similarly for Δ^+. Given a smooth space-time function ξ, we now have a *family* of canonical lifts $\mathscr{L}_\varepsilon(\xi) = (\Pi^{(\varepsilon)}, \Gamma^{(\varepsilon)})$ of ξ to models on $\mathscr{T}_{\mathrm{ex}}$ given as

before by (4.6) and (3.11), together with

$$(5.9) \qquad \big(\Pi_z^{(\varepsilon)}\mathcal{E}(\tau)\big)(\bar{z}) = \varepsilon\big(\Pi_z^{(\varepsilon)}\tau\big)(\bar{z}) - \varepsilon \sum_{|k|<|\tau|+1} \frac{(\bar{z}-z)^k}{k!}\big(D^k\Pi_z^{(\varepsilon)}\tau\big)(z) \ .$$

The operators $\Gamma_{z\bar{z}}^{(\varepsilon)}$ are also built in exactly the same way as in Section 4.2, with the additional definition

$$(5.10) \qquad f_z^{(\varepsilon)}(\mathcal{E}_\ell\tau) = -\varepsilon \sum_{|k+\ell|<|\tau|+1} \frac{(-z)^k}{k!}\big(D^{k+\ell}\Pi_z^{(\varepsilon)}\tau\big)(z) \ ,$$

which in particular allows us to rewrite (5.9) as

$$(5.11) \qquad \Pi_z^{(\varepsilon)}\mathcal{E}(\tau) = \varepsilon\Pi_z^{(\varepsilon)}\tau + \sum_{|k|<|\tau|+1} \Pi_z^{(\varepsilon)}\frac{(X+z)^k}{k!}f_z^{(\varepsilon)}(\mathcal{E}_k\tau) \ .$$

These definitions guarantee that if we define $\mathbf{\Pi}^{(\varepsilon)}$ as in Section 4.2, one has

$$(5.12) \qquad\qquad\qquad \mathbf{\Pi}^{(\varepsilon)}\mathcal{E}(\tau) = \varepsilon\mathbf{\Pi}^{(\varepsilon)}\tau \ ,$$

and it is in this sense that $\mathcal{E}$ represents multiplication by ε.

REMARK 5.6. These definitions deviate slightly from those given in [**HQ15**]. They are however equivalent and amount to a simple change of basis in T^+. The reason for our choice here is that it leads as in (4.14) to the nicer expression

$$\Delta^+\mathscr{E}_\ell(\tau) = (\mathscr{E}_\ell \otimes \mathrm{id})\Delta\tau + \sum_k \frac{X^k}{k!} \otimes \mathscr{E}_{\ell+k}(\tau) \ ,$$

for Δ^+, where we used similarly to before the convention that $\mathscr{E}_\ell(\tau) = 0$ unless $|\tau| + 1 > |\ell|$.

REMARK 5.7. At this stage, we would like to point out one fundamental difference between $\mathcal{E}$ and $\mathcal{I}$. In the case of $\mathcal{I}$, the fact that it represents K in the sense of (3.11) was made part of the definition of our notion of admissible model, and we only ever considered renormalisation procedures that preserve the set of admissible models. In the case of $\mathcal{E}$ however, we only impose (5.10) and (5.12) for the canonical lifts $\mathscr{L}_\varepsilon(\xi)$. In general, we do *not* impose that these identities are satisfied for an arbitrary admissible model. In particular, it is crucial to allow for renormalisation procedures that break them. This is reminiscent of the status of the product which is typically not preserved by renormalisation.

This construction also suggests a definition for the operator $\hat{\mathcal{E}}$: given an admissible model $(\mathbf{\Pi}, f)$, we set

$$\big(\hat{\mathcal{E}}U\big)(z) = \mathcal{E}U(z) - \sum_\ell \frac{(X+z)^\ell}{\ell!}f_z(\mathscr{E}_\ell(U(z))) \ .$$

Combining this with (5.11), it follows immediately that for a canonical model of the type $\mathscr{L}_\varepsilon(\xi)$, one has the identity

$$\mathcal{R}\hat{\mathcal{E}}U = \varepsilon\mathcal{R}U \ .$$

This however is a particular property of such canonical models. In general, there is no reason for $\mathcal{R}\hat{\mathcal{E}}U$ to be proportional to $\mathcal{R}U$! What is more, the operator $\hat{\mathcal{E}}$ has a regularising property in the $\mathcal{D}^\gamma$-scale of spaces. More precisely, we have the following result, the proof of which is given in [**HQ15**, Prop. 3.15].

PROPOSITION 5.8. *For the regularity structure $(T_{\mathrm{ex}}, \mathcal{G})$ just constructed, there exists $\delta > 0$ such that the operator $\hat{\mathcal{E}}$ is bounded from $\mathcal{D}^\gamma$ to $\mathcal{D}^{(\gamma+1)\wedge\delta}$.*

In particular, combining this with Theorem 3.17, it turns out that if we set $|\Xi| = -\frac{5}{2}-\kappa$ for $\kappa > 0$ small enough, then the map $\Phi \mapsto \mathcal{P}\hat{\mathcal{E}}\Phi^5$ is bounded from $\mathcal{D}^\gamma$ into $\mathcal{D}^\gamma$ as soon as $\gamma > 1 + 4\kappa$. Similarly to before one can show that, for any admissible model and for sufficiently regular initial conditions, the fixed point problem (5.8) admits short-time solutions, with the existence time of these solutions uniformly bounded over bounded balls in the space of admissible models, and solutions being continuous not just as a function of their initial condition but also as a function of the model. Furthermore, for models of the type $\mathscr{L}_\varepsilon(\xi)$ (with ξ continuous), these solutions are such that $\varphi = \mathcal{R}\Phi$ solves

$$\partial_t\varphi = \Delta\varphi - \varphi^3 - a\varepsilon\varphi^5 + \xi \ , \qquad \varphi(0,\cdot) = \Phi_0 \ .$$

REMARK 5.9. One crucial remark at this stage is that solutions to (5.8) do not depend on the full model on $(T_{\mathrm{ex}}, \mathcal{G})$, but only on its restriction to the smaller regularity structure $(T, \mathcal{G})$. In particular, these solutions make sense for *any* admissible model on $(T, \mathcal{G})$, even if such a model does not come from a model on $(T_{\mathrm{ex}}, \mathcal{G})$!

As before, one cannot expect that $\mathscr{L}_\varepsilon(\xi_\varepsilon)$ converges to a limit if ξ_ε is a mollification of white noise. However, it is possible to perform a renormalisation such that the models $M_\varepsilon\mathscr{L}_\varepsilon(\xi_\varepsilon)$ converge to a limit on the smaller structure $(T, \mathcal{G})$. Even after renormalisation, these models do *not* converge on $(T_{\mathrm{ex}}, \mathcal{G})$, but thanks to Remark 5.9 this does not matter for us. Let us now describe the renormalisation maps M suitable in this setting. We keep the same graphical notation for basis vectors of T/T_{ex} as previously, with $\mathcal{E}$ depicted by a circle. With this notation, one has for example

$$\mathcal{E}(\mathcal{I}(\Xi)^5) = \text{\vcenter{\hbox{⬥}}} \ , \qquad \mathcal{I}(\Xi)^2\mathcal{I}\big(\mathcal{E}(\mathcal{I}(\Xi)^5)\big) = \text{\vcenter{\hbox{⬥}}} \ .$$

Since we only work with Gaussian approximations, we only keep L_1 and L_3 from Section 4.5, which we rename $\tilde{L}_1$ and $\tilde{L}_2$ to avoid confusion. The operator $\tilde{L}_1$ is extended to all of T_{ex} in the same way as before by furthermore postulating that it commutes with $\mathcal{E}$, so that one has for example

$$\tilde{L}_1\text{\vcenter{\hbox{⬥}}} = \binom{5}{2}\text{\vcenter{\hbox{⬥}}} \ , \qquad \tilde{L}_1\text{\vcenter{\hbox{⬥}}} = \binom{5}{2}\text{\vcenter{\hbox{⬥}}} + \text{\vcenter{\hbox{⬥}}} \ .$$

In the case of $\tilde{L}_2$, we extend it to T_{ex} by simply postulating that it vanishes on all symbols except for ⚘ and ⚘, so we do for example set $\tilde{L}_2$⚘ $= 0$ rather than $\tilde{L}_2$⚘ $= \mathcal{E}(\mathbf{1})$. Finally, we add two more operators $\tilde{L}_3$ and $\tilde{L}_4$ given by

$$(5.13) \qquad \tilde{L}_3 \text{⚘} = \mathbf{1} , \qquad \tilde{L}_3 \text{⚘} = 4\!\uparrow , \qquad \tilde{L}_4 \text{⚘} = \mathbf{1} , \qquad \tilde{L}_4 \text{⚘} = 5\!\uparrow ,$$

as well as $\tilde{L}_i \tau = 0$ for all combinations of $i \in \{3, 4\}$ and $\tau \in \mathcal{W}$ not appearing in (5.13). The rationale here is that $\tilde{L}_3$ "contracts" all instances of ⚘, which appears 4 times in ⚘ and nowhere else, and similarly for $\tilde{L}_4$. Similarly to before, we then consider renormalisation maps of the form $\tilde{M} = \exp\!\big(-\sum_i C_i \tilde{L}_i\big)$. There is an analogue in this context to Proposition 4.7, which shows that these renormalisation maps are "legal" in the sense that they induce an action on admissible models with the property that a model represented by $\mathbf{\Pi}$ is mapped to one represented by $\mathbf{\Pi}\tilde{M}$.

By a calculation analogous to that of Proposition 4.9, one can then show the following.

PROPOSITION 5.10. *Let $\tilde{M}$ be as above and let $(\Pi^{\tilde{M}}, \Gamma^{\tilde{M}})$ be the model obtained by acting with $\tilde{M}$ on $(\Pi, \Gamma) = \mathscr{L}_\varepsilon(\xi)$ for some smooth function ξ and some $\varepsilon > 0$. Let furthermore Φ be the solution to (5.8) with respect to the model $(\Pi^{\tilde{M}}, \Gamma^{\tilde{M}})$. Then, the function $u(t, x) = \big(\mathcal{R}^{\tilde{M}}\Phi\big)(t, x)$ solves the equation*

$$\partial_t u = \Delta u - H_3(u, \tilde{C}_1) - a\varepsilon H_5(u, \tilde{C}_1) - (9\tilde{C}_2 + 20a\tilde{C}_3 + 25a^2\tilde{C}_4)u + \xi ,$$

where H_n denotes the nth Hermite polynomial as before.

Given $c \in \mathbf{R}$ and $\varepsilon > 0$, we now define $\tilde{M}_{c,\varepsilon} = \exp\!\big(-\sum_i C_i^{(c,\varepsilon)} \tilde{L}_i\big)$ by setting

$$\tilde{C}_1^{(c,\varepsilon)} = \tilde{C}_1^{(\varepsilon)} , \quad \tilde{C}_2^{(c,\varepsilon)} = \tilde{C}_2^{(\varepsilon)} , \quad \tilde{C}_3^{(c,\varepsilon)} = c\tilde{C}_3^{(\varepsilon)} , \quad \tilde{C}_4^{(c,\varepsilon)} = c^2\tilde{C}_4^{(\varepsilon)} ,$$

with

with the same graphical notations as in the previous section. Here, $\tilde{C}_1$ diverges like ε^{-1} and $\tilde{C}_2$ diverges logarithmically, but the remaining two constants actually converge to finite limits as $\varepsilon \to 0$. The reason for these definitions is the following result, which is the crucial point on which the proof of Theorem 5.5 hinges.

THEOREM 5.11. *Let $\xi_\varepsilon = \varrho_\varepsilon * \xi$ for some space-time white noise ξ and, for $c \in \mathbf{R}$ and $\varepsilon > 0$, let*

$$\mathfrak{M}_{c,\varepsilon} = \tilde{M}_{c,\varepsilon}\mathscr{L}_{c\varepsilon}(\xi_\varepsilon) \ ,$$

with $\tilde{M}_{c,\varepsilon}$ as above. Then, the limit $\lim_{\varepsilon \to 0} \mathfrak{M}_{c,\varepsilon}$ exists in the space of admissible models on $(T, \mathcal{G})$ and is independent of c.

REMARK 5.12. It is *not* true that the models $\mathfrak{M}_{c,\varepsilon}$ converge to a limit on the extended regularity structure $(T_{\mathrm{ex}}, \mathcal{G})$. Fortunately, this is not needed for our argument.

The proof of this result follows the same line of argument as in Section 4.6. The reason why the limit turns out to be independent of c is that the value of c only affects $\Pi_z \tau$ for symbols τ containing at least one instance of $\mathcal{E}$. Our choice of renormalisation constants is then precisely such that, in the limit $\varepsilon \to 0$, all of these terms vanish.

Combining Theorem 5.11 with Proposition 5.10 and the fact that solutions to (5.8) depend continuously on the underlying model, we conclude that the solutions to

$$\partial_t u = \Delta u - H_3(u, \tilde{C}_1^{(\varepsilon)}) - a\varepsilon H_5(u, \tilde{C}_1^{(\varepsilon)}) - (9\tilde{C}_2^{(\varepsilon)} + 20a\tilde{C}_3^{(\varepsilon)} + 25a^2\tilde{C}_4^{(\varepsilon)})u + \xi_\varepsilon \ ,$$

as well as those to

$$\partial_t u = \Delta u - u^3 + (3\tilde{C}_1^{(\varepsilon)} - 9\tilde{C}_2^{(\varepsilon)})u + \xi_\varepsilon \ ,$$

both converge to the same limit as $\varepsilon \to 0$. Since the limit of the latter is the Φ_3^4 model by definition, the claim of Theorem 5.5 follows at once.

References

[AHR01] S. ALBEVERIO, Z. HABA, and F. RUSSO. A two-space dimensional semilinear heat equation perturbed by (Gaussian) white noise. *Probab. Theory Relat. Fields* **121**, no. 3, (2001), 319–366. `doi:10.1007/s004400100153`

[AR91] S. ALBEVERIO and M. RÖCKNER. Stochastic differential equations in infinite dimensions: solutions via Dirichlet forms. *Probab. Theory Relat. Fields* **89**, no. 3, (1991), 347–386. `doi:10.1007/BF01198791`.

[BCD11] H. BAHOURI, J.-Y. CHEMIN, and R. DANCHIN. *Fourier Analysis and Nonlinear Partial Differential Equations*, vol. 343 of *Grundlehren Math. Wiss.*, Springer, Heidelberg, 2011.

[BG97] L. BERTINI and G. GIACOMIN. Stochastic Burgers and KPZ equations from particle systems. *Commun. Math. Phys.* **183**, no. 3, (1997), 571–607. `doi:10.1007/s002200050044`.

[BP08] J. BOURGAIN and N. PAVLOVIĆ. Ill-posedness of the Navier-Stokes equations in a critical space in 3D. *J. Funct. Anal.* **255**, no. 9, (2008), 2233–2247. `doi:10.1016/j.jfa.2008.07.008`.

[CC13] R. CATELLIER and K. CHOUK. Paracontrolled distributions and the 3-dimensional stochastic quantization equation. *ArXiv e-prints* (2013). `arXiv:1310.6869`.

[Col83] J.-F. COLOMBEAU. A multiplication of distributions. *J. Math. Anal. Appl.* **94**, no. 1, (1983), 96–115. `doi:10.1016/0022-247X(83)90007-0`.

[CW15] A. CHANDRA and H. WEBER. Stochastic PDEs, regularity structures, and interacting particle systems. *ArXiv e-prints* (2015). `arXiv:1508.03616`.

[DPD03] G. Da Prato and A. Debussche. Strong solutions to the stochastic quantization equations. *Ann. Probab.* **31**, no. 4, (2003), 1900–1916. `doi:10.1214/aop/1068646370`.

[Fal12] P. Falco. Kosterlitz-Thouless transition line for the two dimensional Coulomb gas. *Commun. Math. Phys.* **312**, no. 2, (2012), 559–609. `doi:10.1007/s00220-012-1454-7`.

[FH14] P. Friz and M. Hairer. *A Course on Rough Paths*. Universitext. Springer, 2014. With an introduction to Regularity Structures. `doi:10.1007/978-3-319-08332-2`.

[FS81] J. Fröhlich and T. Spencer. The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas. *Commun. Math. Phys.* **81**, no. 4, (1981), 527–602.

[GIP12] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *ArXiv e-prints* (2012). `arXiv:1210.2684v3`.

[GLP99] G. Giacomin, J. L. Lebowitz, and E. Presutti. Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems. In *Stochastic Partial Differential Equations: Six Perspectives*, vol. 64 of *Math. Surv. Monogr.*, 107–152. Amer. Math. Soc., Providence, RI, 1999.

[Gub04] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.* **216**, no. 1, (2004), 86–140. `doi:10.1016/j.jfa.2004.01.002`.

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

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

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

[Hai15] M. Hairer. Introduction to regularity structures. *Braz. J. Probab. Stat.* **29**, no. 2, (2015), 175–210. `arXiv:1401.3014`. `doi:10.1214/14-BJPS241`.

[HL15] M. Hairer and C. Labbé. Multiplicative stochastic heat equations on the whole space. *ArXiv e-prints* (2015). `arXiv:1504.07162`.

[HP13] M. Hairer and N. S. Pillai. Regularity of laws and ergodicity of hypoelliptic SDEs driven by rough paths. *Ann. Probab.* **41**, no. 4, (2013), 2544–2598. `arXiv:1104.5218`. `doi:10.1214/12-AOP777`.

[HP14] M. Hairer and É. Pardoux. A Wong-Zakai theorem for stochastic PDEs. *ArXiv e-prints* (2014). `arXiv:1409.3138`. *J. Math. Soc. Jpn.*, to appear.

[HQ15] M. Hairer and J. Quastel. A class of growth models rescaling to KPZ, 2015. Preprint.

[HS14] M. Hairer and H. Shen. The dynamical sine-Gordon model. *ArXiv e-prints* (2014). `arXiv:1409.5724`.

[HS15] M. Hairer and H. Shen. A central limit theorem for the KPZ equation. *ArXiv e-prints* (2015). `arXiv:1507.01237`.

[HX15] M. Hairer and W. Xu. Large scale behaviour of $3D$ continuous phase coexistence models, 2015. Preprint.

[JLM85] G. Jona-Lasinio and P. K. Mitter. On the stochastic quantization of field theory. *Commun. Math. Phys.* **101**, no. 3, (1985), 409–436. `doi:10.1007/BF01216097`.

[KPZ86] M. Kardar, G. Parisi, and Y.-C. Zhang. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56**, no. 9, (1986), 889–892. `doi:10.1103/PhysRevLett.56.889`.

[Kup15] A. Kupiainen. Renormalization group and stochastic PDEs. *Ann. Henri Poincaré* 1–39. `arXiv:1410.3094`. `doi:10.1007/s00023-015-0408-y`.

[LCL07] T. J. Lyons, M. Caruana, and T. Lévy. *Differential Equations Driven by Rough Paths*, vol. 1908 of *Lect. Notes Math.*, Springer, Berlin, 2007. Lectures

from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, with an introduction concerning the Summer School by Jean Picard.

[Lyo98]   T. J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14**, no. 2, (1998), 215–310.

[MW14]   J.-C. Mourrat and H. Weber. Convergence of the two-dimensional dynamic Ising-Kac model to Φ^4_2. *ArXiv e-prints* (2014). `arXiv:1410.1179`.

[MW15]   J.-C. Mourrat and H. Weber. Global well-posedness of the dynamic Φ^4 model in the plane. *ArXiv e-prints* (2015). `arXiv:1501.06191`.

[Nua95]   D. Nualart. *The Malliavin calculus and related topics. Probab. Appl. (N. Y.)*, Springer-Verlag, New York, 1995.

[PW81]   G. Parisi and Y. S. Wu. Perturbation theory without gauge fixing. *Sci. Sin.* **24**, no. 4, (1981), 483–496.

[Qua11]   J. Quastel. Introduction to KPZ. *Curr. Dev. Math.* **1**, (2011), 125–194. `doi:10.4310/CDM.2011.v2011.n1.a3`.

[Sch54]   L. Schwartz. Sur l'impossibilité de la multiplication des distributions. *C. R. Acad. Sci. Paris* **239**, (1954), 847–848.

[Sim97]   L. Simon. Schauder estimates by scaling. *Calc. Var. Partial Differ. Equ.* **5**, no. 5, (1997), 391–407. `doi:10.1007/s005260050072`.

[You36]   L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.* **67**, no. 1, (1936), 251–282. `doi:10.1007/BF02401743`.

THE UNIVERSITY OF WARWICK, UNITED KINGDOM
E-mail address: `m.hairer@warwick.ac.uk`