ANNALES
DE LA FACULTÉ
DES SCIENCES
DE TOULOUSE
Mathématiques

AJAY CHANDRA AND HENDRIK WEBER
Stochastic PDEs, Regularity structures, and interacting particle systems

Stochastic PDEs, Regularity structures, and interacting particle systems

Ajay Chandra [1] and Hendrik Weber [2]

Abstract. — These lecture notes grew out of a series of lectures given by the second named author in short courses in Toulouse, Matsumoto, and Darmstadt. The main aim is to explain some aspects of the theory of "Regularity structures" developed recently by Hairer in [27]. This theory gives a way to study well-posedness for a class of stochastic PDEs that could not be treated previously. Prominent examples include the KPZ equation as well as the dynamic Φ_3^4 model.

Such equations can be expanded into formal perturbative expansions. Roughly speaking the theory of regularity structures provides a way to truncate this expansion after finitely many terms and to solve a fixed point problem for the "remainder". The key ingredient is a new notion of "regularity" which is based on the terms of this expansion.

Résumé. — Ces notes sont basées sur trois cours que le deuxième auteur a donnés à Toulouse, Matsumoto et Darmstadt. L'objectif principal est d'expliquer certains aspects de la théorie des « structures de régularité » développée récemment par Hairer [27]. Cette théorie permet de montrer que certaines EDP stochastiques, qui ne pouvaient pas être traitées auparavant, sont bien posées. Parmi les exemples se trouvent l'équation KPZ et le modèle Φ_3^4 dynamique.

Telles équations peuvent être développées en séries perturbatives formelles. La théorie des structures de régularité permet de tronquer ce développement après un nombre fini de termes, et de résoudre un problème de point fixe pour le reste. L'idée principale est une nouvelle notion de régularité des distributions, qui dépend des termes de ce développement.

Ajay Chandra and Hendrik Weber

Lecture 1

In this introductory lecture we outline the scope of the theory of regularity structures. We start by discussing two important stochastic PDE (SPDE) coming from physics. The first is the Kardar–Parisi–Zhang (KPZ) equation which is *formally* given by

$$\partial_t h(t, x) = \partial_x^2 h(t, x) + \frac{1}{2} (\partial_x h(t, x))^2 + \xi(t, x) \ . \qquad \text{(KPZ)}$$

We will restrict ourselves to the case where the spatial variable x takes values in a one dimensional space. The term $\xi(t, x)$ denotes space-time white noise which is not an actual function but a quite irregular random (Schwartz) distribution. This equation was introduced in [32] in 1986 and is a model for the fluctuations of an evolving one dimensional interface which separates two competing phases of a physical system. An explanation for the presence of the individual terms on the right hand side of (KPZ) can be found in [43]. The KPZ equation has recieved a lot of attention from mathematicians in recent years: One major development was an exact formula for the one point distribution of solutions to (KPZ) which was found independently by [44] and [2]. This formula is susceptible to asymptotic analysis which reveals that the scaled one-point distributions converge to the Tracy–Widom distribution, a result that has been spectacularly confirmed by physical experiments [45].

Throughout these lectures, we will focus more on our second example, the dynamic Φ_d^4 model. Formally this model is given by

$$\partial_t \varphi(t, x) = \Delta \varphi(t, x) - \varphi^3(t, x) - m^2 \varphi(t, x) + \xi(t, x) \ . \qquad (\Phi_d^4)$$

Here the spatial variable x takes values in a d-dimensional space and ξ is again space-time white noise. The invariant measure of (Φ_d^4) was studied intensively in the seventies in the context of Constructive Quantum Field Theory (see e.g. [20, 13, 19, 7]). Formally this invariant measure is given by

$$\mu(\mathrm{d}\varphi) \propto \exp \left[-2 \int_{\mathbb{R}^d} \frac{1}{4} \varphi^4(x) + \frac{1}{2} m \varphi^2 \, \mathrm{d}x \right] \nu(\mathrm{d}\varphi) \qquad (1.1)$$

where ν is the law of Gaussian Free Field (GFF). The GFF can be thought of as a Gaussian random field on $\varphi : \mathbb{R}^d \to \mathbb{R}$ with covariance given by $\mathbf{E}_\nu[\varphi(x)\varphi(y)] = \frac{1}{2} G(x - y)$ where G is the Green's function of the d-dimensional Laplacian. However when $d > 1$ the measure ν is not supported on a space of functions so φ must actually be a distribution. A rigorous understanding of (1.1) then requires interpreting nonlinearities of distributions.

In addition to being a toy model for QFT the measure (1.1) can be seen as a continuum analog of the famous ferromagnetic Ising model. For example, in [21] the authors showed that the concrete measure corresponding to (1.1)

in $d = 2$ has a phase transition; their proof is a sophisticated version of the classical Peierls argument [41] for the Ising model. We will close the first lecture by describing how the equation (Φ_d^4) can be obtained as the scaling limit of a dynamical Ising model with long range interaction (at least for $d = 1, 2$).

An important remark is that the theory of regularity structures will be restricted to studying (Φ_d^4) in the regime $d < 4$ and (KPZ) for space dimension $d < 2$. These are both manifestations of a fundamental restriction of the theory which is the assumption of *subcriticality* which will be discussed later. Another important remark about the scope of the theory is that regularity structures deliver a robust mathematical theory for making sense of (KPZ) and (Φ_d^4) on compact space-time domains and describe their solutions on *very small scales*. The large scale behaviour of these solutions is mostly out of the current theory's scope (although some results have been obtained, see e.g. [29, 38]). This is quite important since it is primarily the large scale behaviour of solutions which makes the equations (KPZ) and (Φ_d^4) experimentally validated models of physical phenomena, in particular the macroscopic behaviour of *critical* systems. However, understanding small scale behaviour and proving well-posedness is a fundamental step towards having a complete theory for these SPDE[1]. As mentioned earlier, a large obstacle we must overcome is that the ∇h of (KPZ) and φ of (Φ_d^4) will in general be distributions. This makes interpreting the nonlinearities appearing in these equations highly non-trivial.

1.1. Random Distributions and Scaling Behaviour

1.1.1. Space-time white noise

We start by defining space-time white noise ξ which appeared in both (KPZ) and (Φ_d^4). Formally $\xi(t, x)$ is a random Gaussian function on $\mathbb{R} \times \mathbb{R}^d$, its covariance is given by

$$\mathbb{E}[\xi(t, x)\xi(t', x')] = \delta(t - t')\, \delta^d(x - x') \,, \tag{1.2}$$

where δ^d denotes the d-dimensional Dirac δ distribution. However for any fixed (t, x) one cannot rigorously interpret $\xi(t, x)$ as a random variable, there is no coordinate process. Instead ξ must be interpreted as a random distribution, a random element of $\mathcal{S}'(\mathbb{R} \times \mathbb{R}^d)$ whose law is centered Gaussian. For any $f \in \mathcal{S}'(\mathbb{R} \times \mathbb{R}^d)$ and smooth test function η on $\mathbb{R} \times \mathbb{R}^d$ we denote

---

[1] There are also some physical phenomena appearing in the scale regimes that regularity structures can access, such as *near*-critical systems at large volume.

by (f, η) the corresponding duality pairing. The quantity $(\xi, \bullet)$ is then the analog of the coordinate process for ξ and the rigorous version of (1.2) is given by

$$\mathbb{E}[(\xi, \eta_1)(\xi, \eta_2)] = \int_{\mathbb{R} \times \mathbb{R}^d} \eta_1(t, x)\eta_2(t, x) \, dt \, dx \quad \text{for any smooth } \eta_1, \eta_2. \quad (1.3)$$

Remark 1.1. — The formula (1.3) implies that $(\xi, \bullet)$ can be extended beyond smooth functions to an isometry from $L^2(\mathbb{R} \times \mathbb{R}^d)$ to $L^2(\Omega, \mathcal{F}, \mathbb{P})$ where $(\Omega, \mathcal{F}, \mathbb{P})$ is the underlying probability space. Adapting the definition to the case of $\mathbb{R}$ instead of $\mathbb{R} \times \mathbb{R}^d$ gives us the process called white noise, in this case one has

$$\mathbb{E}\big[(\xi, \mathbf{1}_{[0,s]})(\xi, \mathbf{1}_{[0,t]})\big] = \int_{\mathbb{R}} \mathbf{1}_{[0,s]}(r) \, \mathbf{1}_{[0,t]}(r) \, dr = s \wedge t \,,$$

so $(\xi, \mathbf{1}_{[0,t]})$ " $= \int_0^t \xi(r) \, dr$ " is a Brownian motion and we see that ξ can be thought of as the derivative of Brownian motion. In these lectures we will focus on equations driven by space-time noise processes so we will use the term white noise to refer to space-time white noise.

We will frequently be interested in the scaling behaviour of space-time distributions. Given a white noise ξ and positive parameters $\tau, \lambda > 0$ we can *define* a new random distribution $\xi_{\tau, \lambda}$ via

$$(\xi_{\tau, \lambda}, \eta) := (\xi, \mathcal{S}^{\tau, \lambda}\eta)$$

where for any smooth function η we have set

$$(\mathcal{S}^{\tau, \lambda}\eta)(t, x) := \tau^{-1}\lambda^{-d}\eta(\tau^{-1}t, \lambda^{-1}x) \,.$$

This is a simple rescaling operation, if ξ was an actual function then this would amount to setting $\xi_{\tau, \lambda}(t, x) = \xi(\tau t, \lambda x)$. By (1.3) one has

$$\mathbb{E}\big[(\xi_{\tau, \lambda}, \eta)^2\big] = \int_{\mathbb{R} \times \mathbb{R}^d} \tau^{-2}\lambda^{-2d}\eta(\tau^{-1}t, \lambda^{-1}x)^2 \, dt \, dx$$

$$= \tau^{-1}\lambda^{-d} \int_{\mathbb{R} \times \mathbb{R}^d} \eta(t, x)^2 \, dt \, dx \,. \quad (1.4)$$

Since ξ and $\xi_{\tau, \lambda}$ are centred Gaussian processes we can conclude that ξ is scale invariant in distribution, in particular $\xi_{\tau, \lambda} \overset{\text{law}}{=} \tau^{-\frac{1}{2}}\lambda^{-\frac{d}{2}}\xi$.

1.1.2. Scaling Behaviour for SPDEs and Subcriticality

Both (KPZ) and (Φ_d^4) are non-linear perturbations of a linear SPDE called the stochastic heat equation (SHE)

$$\partial_t Z(t, x) = \Delta Z(t, x) + \xi(t, x) \quad \text{(SHE)}$$

where as before $(t, x) \in \mathbb{R} \times \mathbb{R}^d$. As before, ξ cannot be evaluated pointwise and (SHE) has to be interpreted in the distributional sense. Since (SHE) is linear it follows that the solution Z will be Gaussian (for deterministic or Gaussian initial conditions).

Remark 1.2. — The equation (SHE) is sometimes called the *additive* stochastic heat equation in order to distinguish it from the *multiplicative* stochastic heat equation which is given by

$$\partial_t Z(t, x) = \Delta Z(t, x) + Z(t, x)\, \xi(t, x) \ .$$

The above equation has a close relationship to (KPZ) via a change of variables called the Cole–Hopf transformation. However we will not use this transformation nor investigate the multiplicative SHE in these notes. Whenever we refer to the stochastic heat equation we are always refering (SHE).

We now perform some formal computations to investigate the scaling behaviour of solutions (SHE). For $\lambda > 0$ and suitable scaling exponents $\alpha, \beta, \gamma \in \mathbb{R}$ we define $\hat{Z}(t, x) := \lambda^\alpha Z(\lambda^\beta t, \lambda^\gamma x)$ and $\hat{\xi} := \lambda^{\frac{\beta}{2}} \lambda^{\frac{d\gamma}{2}} \xi_{\lambda^\beta, \lambda^\gamma}$, it then follows that

$$\partial_t \hat{Z} = \lambda^{\beta - 2\gamma} \Delta \hat{Z} + \lambda^{\alpha + \frac{\beta}{2} - \frac{d\gamma}{2}} \hat{\xi} \ . \tag{1.5}$$

We have already shown that $\hat{\xi} \overset{\text{law}}{=} \xi$. Therefore, if we set

$$\alpha = \frac{d}{2} - 1 \ , \qquad \beta = 2 \ , \qquad \text{and} \qquad \gamma = 1 \tag{1.6}$$

then we see that $\hat{Z} \overset{\text{law}}{=} Z$ (ignoring boundary conditions) so the solution to (SHE) is also scale invariant.

In general non-linear equations like (KPZ) and (Φ_d^4) will not be scale invariant. If one rescales these equations according to the exponents given in (1.6) then the non-linearity will be multiplied by a prefactor which is some power of λ; the assumption of subcriticality then requires that this prefactor vanish as $\lambda \to 0$. Roughly speaking, this condition enforces that the solutions to (KPZ) and (Φ_d^4) both behave like the solution to the (SHE) *on small scales*. Let us illustrate this for (KPZ). We perform the same scaling as in (1.6) and set $\hat{h}(t, x) = \lambda^{-\frac{1}{2}} h(\lambda^2 t, \lambda x)$. This gives

$$\partial_t \hat{h} = \partial_x^2 \hat{h} + \frac{\lambda^{\frac{1}{2}}}{2} (\partial_x \hat{h})^2 + \hat{\xi} \ .$$

On small scales, i.e. for $\lambda \to 0$, the prefactor $\lambda^{\frac{1}{2}}$ of the non-linear term goes to zero. We perform the same calculation for (Φ_d^4), for this discussion the mass term $m^2 \varphi$ is irrelevant so we drop it. Applying the scaling (1.6), i.e. setting $\hat{\varphi}(t, x) = \lambda^{\frac{d}{2} - 1} \varphi(\lambda^2 t, \lambda x)$ we get

$$\partial_t \hat{\varphi}(t, x) = \Delta \hat{\varphi}(t, x) - \lambda^{4-d} \hat{\varphi}^3 + \hat{\xi} \ .$$

If the spatial dimension d is strictly less than 4 the prefactor λ^{4-d} vanishes in the limit $\lambda \to 0$. We call $d < 4$ the *subcritical* regime. If $d = 4$ the prefactor $\lambda^{4-d} = 1$; this is the *critical* regime. The regime $d \geqslant 5$ is called the *supercritical* regime.

We now state a crude "definition" of subcriticality which will be sufficient for these notes. The interested reader is referred to [27, Assumption 8.3] for a more precise definition of subcriticality which also extends to different types of noise ξ and a larger class of regularising linear operators.

DEFINITION 1.3. — *Consider the equation*

$$\partial_t u = \Delta u + F(u, \nabla u) + \xi \,, \tag{1.7}$$

in d spatial dimensions. Equation (1.7) *is called* subcritical *if under the scaling* (1.6) *the non-linear term $F(u, \nabla u)$ gets transformed into a term $F_\lambda(u, \nabla u)$ which formally goes to zero as λ goes to 0.*

The main result of [27] can roughly be paraphrased as follows.

METATHEOREM 1.4 ([27]). — *Assume that SPDE* (1.7) *is subcritical. We assume that x takes values in a compact subset of $\mathbb{R}^d$ with some boundary conditions. Furthermore, we prescribe an initial condition u_0 which has the same spatial regularity as we expect for the solution u.*

There is a natural notion of solution to (1.7) *and such solutions exist and are unique on a time interval $[0, T)$ for some random $T > 0$.*

Remark 1.5. — The assumption of subcriticality is not just a technical restriction. For example it has been proven that a non-trivial Φ_d^4 cannot exist for $d \geqslant 5$ (this result extends to $d \geqslant 4$ with some caveats) [1, 16].

Remark 1.6. — We will see below that the statement of Metatheorem 1.4 really consists of two independent statements: (i) for subcritical equations it is possible to build the algebraic and analytic structure that allows to formulate the equation and (ii) all the stochastic processes entering the expansion converge (after renormalisation). It is an astonishing fact that in the case of equations driven by white noise, the scaling conditions for these two statements to hold coincide. It is however possible to define a notion of subcriticality for more general equations driven by a more general noise term. This generalised notion still implies that it is possible to build the algebraic and analytic structure, but there are examples, where stochastic terms fail to be renormalisable.[(2)]

[(2)] This may be familiar to readers who know the theory of rough paths: In principle this theory allows to solve differential equations with a driving noise dW for W of arbitrary positive regularity by increasing the number of iterated integrals one considers. However, the stochastic calculations needed to actually construct the iterated integrals fail for fractional Brownian motion of Hurst index $H < \frac{1}{4}$ [10].

Remark 1.7. — For what follows we restrict ourselves to the subcritical regime. While the equations (KPZ) and (Φ_d^4) are not scale invariant themselves they do interpolate between two different scale invariant space-time processes, one governing small scales and another governing large scales. As mentioned before the small scale behaviour should be governed by the solution to (SHE). At large scales it is expected that (i) one must use different exponents then (1.6) to get a meaningful limit and (ii) the limiting object will be a non-Gaussian process scale invariant under these different exponents.

For (KPZ) one should apply the famous $1, 2, 3$-scaling

$$\hat{h}(t, x) = \lambda^{-\frac{1}{2}} h(\lambda^{\frac{3}{2}} t, \lambda x).$$

Then, setting $\hat{\xi} = \lambda \xi_{\lambda^{\frac{3}{2}}, \lambda}$ one has the equation

$$\partial_t \hat{h}(t, x) = \frac{1}{2}(\partial_x \hat{h})^2 + \lambda^{-\frac{1}{2}} \partial_x^2 \hat{h} + \lambda^{-\frac{1}{4}} \hat{\xi}(t, x). \tag{1.8}$$

Modulo the subtraction of a drift term, as $\lambda \to \infty$ the solution of (1.8) is conjectured to converge to an object called the *KPZ fixed point* (see [9]). This limiting object is not yet very well understood[3].

The behaviour of (Φ_d^4) at large scales is also of interest, but much less is known in this case.

Remark 1.8. — The main aim of these lectures is to show how the theory of regularity structures can be used to construct local-in-time solutions for Φ_3^4. Let us point out however, that after this result was first published by Hairer in [27], two alternative methods to obtain similar results have been put forward: In [23] Gubinelli, Imkeller and Perkowski developed the method of "paracontrolled distributions" to construct solutions to singular stochastic PDEs and this method was used in [8] to construct local in time solutions to Φ_3^4. Independently, in [35] Kupiainen proposed yet another method based on Wilsonian renormalization group analysis. The result for Φ_3^4 that can be obtained by the method of "paracontrolled distributions" is essentially equivalent to the result obtained in the framework of regularity structures and arguably this method is simpler because less machinery needs to be developed. However, the construction of a more comprehensive theory pays off when looking at more complicated models. For example, approximation results for the multiplicative stochastic heat equation such as obtained in [30] seem out of reach of the method of "paracontrolled distributions" for the moment.

---

[3] In [9] it was shown that this object does *not* coincide with the entropy solution of the Hamilton–Jacobi equation $\partial_t h = \frac{1}{2}(\partial_x h)^2$.

Remark 1.9. — At the time of writing these lectures there were at least three other works ([26], [28], and [15, Chap. 13–15]) that survey the theory of regularity structures. In particular, [28] gives a much more detailed exposition for many of topics we only briefly discuss in Lecture 4. The authors' goal for the present work was (i) to clarify certain central concepts via numerous concrete computations with simple examples and (ii) to give a panoramic view of how the various parts of the theory of regularity structure work together.

1.2. The need for renormalisation

We must clarify what is meant by *solution theory* and *uniqueness* in the *Metatheorem* 1.4. *Classical* solution theories for SPDEs (see e.g. [12, 24, 42]) do not apply here since the solutions are too irregular. For (KPZ) the solution $h(t, x)$ has the regularity of a Brownian motion in space (the mapping $x \mapsto h(t, x)$ for fixed t is almost surely α-Hölder continuous for every $\alpha < \frac{1}{2}$ but not for any $\alpha \geqslant \frac{1}{2}$). Remembering Remark 1.1 we expect that the distributional derivative $\partial_x h$ has the regularity of spatial white noise. For (Φ_d^4) the solution theory was already fairly understood only in $d = 1$ (there φ is α-Hölder for every $\alpha < \frac{1}{2}$ which is largely sufficient to define φ^3, see [17]). In the cases $d = 2, 3$ the subcriticality assumption stated in Definition 1.3 still applies but φ will not be regular enough to be a function.

A natural way to try to interpret nonlinear expressions involving highly irregular objects is regularization. In the context of our singular SPDE this means that if we show that solutions of *regularized* equations converge to some object as we remove the regularization then we can define this limiting object as the solution of the SPDE. Unfortunately this naive approach does not work, the solutions to the regularized equations will either fail to converge or converge to an uninteresting limit. We use the dynamic Φ_2^4 model as a concrete example of this. One natural regularization consists of replacing ξ by a smoothened noise process. Let ϱ be a smooth function on $\mathbb{R} \times \mathbb{R}^d$ which integrates to 1. For $\delta > 0$ we set

$$\varrho_\delta(t, x) := \delta^{-(2+d)} \varrho(\delta^{-2} t, \delta^{-1} x) . \tag{1.9}$$

We use the parabolic scaling $\delta^{-2} t$ and $\delta^{-1} x$ since it will be a convenient choice for later examples. For any $\delta > 0$ we define the regularized noise $\xi_\delta := \xi * \varrho_\delta$ where $*$ indicates space-time convolution. For any fixed positive δ proving (local) existence and uniqueness for the solution of

$$\partial_t \varphi_\delta = \Delta \varphi_\delta - \varphi_\delta^3 + \xi_\delta \tag{1.10}$$

poses no problem in any dimension since the driving noise ξ_δ is smooth. However in [31] this example was studied[4] on the two dimensional torus and it was shown that as $\delta \downarrow 0$ the solutions φ_δ converge to the trivial limit 0 for any initial condition! In order to obtain a non-trivial limit the equation (1.10) has to be modified in a δ dependent way. We will see that in dimensions $d = 2, 3$ if one considers

$$\partial_t \varphi_\delta = \Delta \varphi_\delta - (\varphi_\delta^3 - 3\mathfrak{c}_\delta \varphi_\delta) + \xi_\delta \,, \qquad (1.11)$$

for a suitable dimension dependent choice of *renormalisation* constants $\mathfrak{c}_\delta$, then the solutions φ_δ do indeed converge to a non-trivial limit φ. This constant $\mathfrak{c}_\delta$ will diverge as $\delta \downarrow 0$. For (1.11) in $d = 2$ one can take $C_1 \log(\delta^{-1})$ for a specific constant C_1, while for $d = 3$ one can take $\mathfrak{c}_\delta = C_1\delta^{-1} + C_2 \log(\delta^{-1})$ for specific constants C_1, C_2 where C_1 depends on the choice of ϱ. A similar *renormalisation* procedure is necessary for the KPZ equation. In [25] it was shown that solutions of

$$\partial_t h_\delta(t, x) = \partial_x^2 h_\delta(t, x) + \frac{1}{2}(\partial_x h_\delta(t, x))^2 - \mathfrak{c}_\delta + \xi_\delta(t, x) \qquad (1.12)$$

on the one-dimensional torus converge to a non-trivial limit h when one sets $\mathfrak{c}_\delta = C_1\delta^{-1}$ for a specific constant C_1. We call (1.11) and (1.12) *renormalized equations* and the limits of their corresponding solutions $\varphi := \lim_{\delta \downarrow 0} \varphi_\delta$ and $h := \lim_{\delta \downarrow 0} h_\delta$ are what we define to be solutions of (Φ_d^4) and (KPZ) in Metatheorem 1.4; such solutions are often called *renormalized solutions*.

We now turn to discussing uniqueness for these SPDE. For a fixed subcritical equation one can choose different renormalization schemes which yield different families of renormalized equations and different corresponding renormalized solutions. A simple example of this in the cases of (1.11) or (1.12) would be shifting $\mathfrak{c}_\delta$ by a finite constant independent of δ, this would change the final renormalized solution one arrives at. One could also change the renormalization scheme by using a different mollifier ϱ or use a non-parabolic scaling for a given mollifier. Even with all these degrees of freedom in choosing a renormalization scheme it turns out that for a given subcritical equation the corresponding family of possible renormalized solutions will be parameterized by a finite dimensional space. If a renormalization scheme yields a non-trivial limit renormalized solution then this solution will lie in this family. For (KPZ) and (Φ_d^4) the family of solutions is parameterized by a single real parameter.

---

[4] Actually in [31] a different regularisation of the noise is considered, but that does not change the result.

Remark 1.10. — The reader should compare the situation just described to the familiar problem one encounters when solving the stochastic differential equation

$$\dot{x}(t) = b(x(t)) + \sigma(x(t))\xi(t) \, ; \tag{1.13}$$

(which is more conventionally written as $dx_t = b(x_t)\,dt + \sigma(x_t)\,dW_t$). There it is well-known that different natural regularisations converge to different solutions. An explicit Euler scheme, for example, will lead to the solution in the *Itô* sense (see e.g. [33]) whereas smoothening the noise usually leads to the *Stratonovich* solution (see the e.g. the classical papers [46, 47]). There is a whole one-parameter family of meaningful solution-concepts to (1.13) and the question for uniqueness is only meaningful once it is specified which particular solution one is looking for.

Later in these lecture notes we will discuss how the theory of regularity structures gives a "recipe" for coming up with renormalization schemes which give non-trivial limits, we will also see that the limiting solution itself will be a fairly concrete object in the theory.

1.3. Approximation of renormalised SPDE by interacting particle systems

One might think that by introducing (1.11) and (1.12) we have turned our back on the original equations and physical phenomena they represent. This is not the case however. There is strong evidence, at least for KPZ and for Φ_d^4, that the *renormalised* solutions are the physical solutions. For the KPZ equation subtracting a diverging constant corresponds simply to a change of reference frame. Furthermore, it was shown in [4] that the solutions to KPZ arise as continuum limits for the weakly asymmetric simple exclusion process, a natural surface growth model.

We will now discuss how the dynamic Φ_d^4 model can be obtained as a continuum limit of an Ising model with long range interaction near criticality. In the one dimensional case (where no renormalisation is necessary) this is a well known result [5, 14] and the right scaling relations for spatial dimensions $d = 1, 2, 3$ were conjectured in [18]. One of the interesting features of these scaling relations is that the "infinite" renormalisation constant has a natural interpretation as shift of the critical temperature. The two dimensional convergence result was established only recently in [37]. We will now briefly discuss this result and show how the relevant scaling relations relate to the subcriticality assumption for (Φ_d^4).

For $N \geqslant 1$ let $\Lambda_N = \mathbf{Z}^d/(2N+1)\mathbf{Z}^d$ be the d-dimensional discrete torus. Denote by $\Sigma_N = \{-1, +1\}^{\Lambda_N}$ the set of spin configurations on Λ_N. For a spin configuration $\sigma = (\sigma(k), k \in \Lambda_N)$ we define the *Hamiltonian* as

$$\mathscr{H}_\gamma(\sigma) := -\frac{1}{2} \sum_{k,j \in \Lambda_N} \kappa_\gamma(k-j)\sigma(j)\sigma(k) .$$

$\gamma \in (0,1)$ is a model parameter which determines the interaction range between spins. It enters the model through the definition of the interaction kernel κ_γ which is given by

$$\kappa_\gamma(k) = c_\gamma \gamma^d \mathfrak{K}(\gamma k) ,$$

where $\mathfrak{K} \colon \mathbb{R}^d \to \mathbb{R}$ is a smooth, nonnegative function with compact support and c_γ is chosen to ensure that $\sum_{k \in \Lambda_N} \kappa_\gamma = 1$. One should think of this model as an interpolation between the *classical* Ising model where every spin interacts only with spins in a fixed neighbourhood (corresponding to the case $\gamma = 1$) and the *mean-field* model, where every spin interacts with every other spin and the geometry of the two-dimensional lattice is completely lost (corresponding to the case $\gamma = 0$).

Then for any *inverse temperature* $\beta > 0$ we can define the Gibbs measure λ_γ on Σ_N as

$$\lambda_\gamma(\sigma) := \frac{1}{\mathscr{Z}_\gamma} \exp\Big(- \beta \mathscr{H}_\gamma(\sigma) \Big) ,$$

where as usual

$$\mathscr{Z}_\gamma := \sum_{\sigma \in \Sigma_N} \exp\Big(- \beta \mathscr{H}_\gamma(\sigma) \Big) ,$$

denotes the normalisation constant that makes λ_γ a probability measure.

We want to obtain the SPDE (Φ_d^4) as a scaling limit for this model and therefore, we have to introduce dynamics. One natural choice is given by the *Glauber-dynamics* which are defined by the generator

$$\mathscr{L}_\gamma f(\sigma) = \sum_{j \in \Lambda_N} c_\gamma(\sigma, j)\big(f(\sigma^j) - f(\sigma)\big) ,$$

acting on functions $f \colon \Sigma_N \to \mathbb{R}$. Here $\sigma^j \in \Sigma_N$ is the spin configuration that coincides with σ except for a flipped spin at position j. The jump rates $c_\gamma(\sigma, j)$ are given by

$$c_\gamma(\sigma, j) := \frac{\lambda_\gamma(\sigma^j)}{\lambda_\gamma(\sigma) + \lambda_\gamma(\sigma^j)} .$$

It is easy to check that these jump rates are reversible with respect to the measure λ_γ.

In order to derive the right rescaling for the spin-field σ we write

$$X_\gamma(t,x) = \frac{1}{\delta} h_\gamma\left(\frac{t}{\alpha}, \frac{x}{\varepsilon}\right) \qquad x \in \Lambda_\varepsilon, \, t \geqslant 0 \, .$$

Here $h_\gamma(k,t) = \sum_{\ell \in \Lambda_N} \kappa_\gamma(k-\ell)\sigma(t,\ell)$ is a field of local spatial averages of the field σ and $\alpha, \delta, \varepsilon$ are scaling factors to be determined[5]. Let us sketch how to derive the right scaling relations for $\alpha, \delta, \varepsilon, \gamma$. We only sketch the calculation, the details can be found in [37]. If we apply the generator $\mathscr{L}_\gamma$ to the field X_γ an explicit calculation shows that

$$X_\gamma(t,x) = X_\gamma(0,x) + \int_0^t \left(\frac{\varepsilon^2}{\gamma^2} \frac{1}{\alpha} \Delta_\gamma X_\gamma(s,x) + \frac{(\beta-1)}{\alpha} K_\gamma * X_\gamma(s,x) \right.$$
$$\left. - \frac{\beta^3}{3} \frac{\delta^2}{\alpha} K_\gamma * X_\gamma^3(s,x) + K_\gamma * E_\gamma(s,x) \right) \mathrm{d}s + M_\gamma(t,x) \, , \quad (1.14)$$

for $x \in \Lambda_\varepsilon$. Here Δ_γ is a difference operator (based on the kernel κ_γ) which is scaled to approximate the Laplacian acting on the rescaled spatial variable x. K_γ is an approximation of a Dirac delta function, M_γ is a martingale and E_γ is a (small) error term. The second relevant relation concerns the noise intensity. This is determined by the quadratic variation of M_γ which is given by

$$\langle M_\gamma(\cdot, x), M_\gamma(\cdot, y)\rangle_t = 4 \frac{\varepsilon^d}{\delta^2 \alpha} \int_0^t \sum_{z \in \Lambda_\varepsilon} \varepsilon^2 K_\gamma(x-z) \, K_\gamma(y-z) \, C_\gamma(s,z) \, \mathrm{d}s \, ,$$

where $C_\gamma(s,z) := c_\gamma(\sigma(s/\alpha), z/\varepsilon)$.

In order to obtain (Φ_d^4) we need to choose scaling factors satisfying

$$1 \approx \frac{\varepsilon^2}{\gamma^2} \frac{1}{\alpha} \approx \frac{\delta^2}{\alpha} \approx \frac{\varepsilon^d}{\delta^2 \alpha} \, ,$$

which leads to

$$\varepsilon \approx \gamma^{\frac{4}{4-d}}, \qquad \alpha \approx \gamma^{\frac{2d}{4-d}}, \qquad \delta \approx \gamma^{\frac{d}{4-d}} \, .$$

It is striking to note, that these equations can be satisfied for spatial dimensions $d = 1, 2, 3$ but they cannot hold as soon as $d = 4$. This corresponds exactly to the criticality assumption for (Φ_d^4).

At first sight (1.14) suggests that β should be so close to one that $(\beta-1)/\alpha = O(1)$. Note that $\beta = 1$ is the critical temperature for the mean field model in our setup. But for $d \geqslant 2$ this naive guess is incorrect. As

---

[5] Working with a field of local averages rather than with the original field σ is more convenient technically, but a posteriori convergence for the original field σ in a weaker topology can be shown as well.

for the macroscopic equation the microscopic model has to be renormalised. Indeed, the main result of [37] states that for $d = 2$ if we set

$$(\beta - 1) = \alpha(\mathfrak{c}_\gamma - m^2) \,,$$

where the "mass" $m \in \mathbb{R}$ is fixed and the extra term $\mathfrak{c}_\gamma$ chosen in a suitable way (diverging logarithmically) as γ goes to 0, then (under suitable assumptions on the initial data) X_γ does indeed converge in law to the solution of (Φ_d^4). A similar result is expected to hold in three dimensions.

Lecture 2

We start this lecture by describing how we will keep track of the regularity of space-time functions *and* distributions. After that we give a review of classical solution techniques for semilinear (stochastic) PDEs. We will explain how a lack of regularity causes problems for these theories, using (Φ_d^4) in $d = 2, 3$ as our examples. We will then describe a *perturbative* approach to these equations. Divergences will be seen clearly in formal expansions of the solutions, this will motivate the choice of diverging renormalization constants appearing in the renormalized equations. We will also go through some calculations to make the objects at hand concrete; this will prepare us for Lecture 3 where we present more abstract parts of the theory.

2.1. Regularity

The functional spaces we use in these notes are a generalization of the usual family of Hölder spaces, these spaces will be denoted by $\mathcal{C}_\mathfrak{s}^\alpha$ where α is the analog of the Hölder exponent. We will measure space-time regularity in a parabolic sense which is why we write $\mathfrak{s}$ in the subscript of $\mathcal{C}_\mathfrak{s}^\alpha$ (the $\mathfrak{s}$ stands for "scaled"). For $z, z' \in \mathbb{R} \times \mathbb{R}^d$ we denote by $\|z' - z\|_\mathfrak{s}$ the parabolic distance between $\bar{z}$ and z. Writing $z' = (t', x')$ and $z = (t, x)$ we set

$$\|z' - z\|_\mathfrak{s} := |t' - t|^{\frac{1}{2}} + \sum_{j=1}^d |x'_j - x_j| \,.$$

Below it will also be useful to have the notion of *scaled dimension* $d_\mathfrak{s} = d + 2$ for space-time $\mathbb{R} \times \mathbb{R}^d$, i.e. the time variable counts for two dimensions.

DEFINITION 2.1. — *For $\alpha \in (0, 1)$ the space $\mathcal{C}_\mathfrak{s}^\alpha(\mathbb{R} \times \mathbb{R}^d)$ consists of all continuous functions $u \colon \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ such for every compact set $\mathfrak{K} \subseteq \mathbb{R} \times \mathbb{R}^d$ one has*

$$\sup_{\substack{z, z' \in \mathfrak{K} \\ z \neq z'}} \frac{|u(z) - u(z')|}{\|z - z'\|_\mathfrak{s}^\alpha} < \infty \,. \tag{2.1}$$

Remark 2.2. — In these notes the theory of regularity structures will be applied to problems in compact space-time domains. Accordingly we will be concerned with estimates that are uniform over compacts instead of trying to get estimates uniform over all of space-time.

In order to accomodate distributions we will want an analog of Hölder spaces where α is allowed to be negative. A natural choice are the (parabolically scaled) Besov spaces $\{\mathcal{B}^\alpha_{\infty,\infty}\}_{\alpha\in\mathbb{R}}$. In particular these spaces agree with our earlier definition for $\alpha \in (0,1)$. In analogy to the positive Hölder spaces we still denote these Besov spaces by $\mathcal{C}^\alpha_\mathfrak{s}$ when $\alpha < 0$.

There are several ways to characterise these Besov spaces (including Paley–Littlewood decomposition ([3]) or wavelet decompositions). For these notes we use a simple definition that is convenient for our purposes. First we need some more notation. For any positive integer r we define B_r to be the set of all smooth functions $\eta : \mathbb{R}^{d+1} \to \mathbb{R}$ with η supported on the unit ball of $\mathbb{R}^{d+1}$ (in the parabolic distance $\|\cdot\|_\mathfrak{s}$) and $\|\eta\|_{C^r} \leqslant 1$. Here $\|\cdot\|_{C^r}$ denotes the standard norm on C^r, that is

$$\|f\|_{C^r} := \sup_{\alpha,|\alpha|\leqslant r} \sup_{x\in\mathbb{R}^{d+1}} |D^\alpha f(x)|$$

where we used multi-index notation. We then have the following definition.

DEFINITION 2.3. — *Suppose that $\alpha < 0$. We define $\mathcal{C}^\alpha_\mathfrak{s}$ to be the set of all distributions $u \in \mathcal{S}'(\mathbb{R}^{d+1})$ such that for any compact set $\mathfrak{K} \subseteq \mathbb{R} \times \mathbb{R}^d$ one has*

$$\|u\|_{\mathcal{C}^\alpha_\mathfrak{s}(\mathfrak{K})} := \sup_{\substack{z\in\mathfrak{K} \\ \lambda\in(0,1]}} \sup_{\eta\in B_r} \left| \frac{\langle u, \mathcal{S}^\lambda_z \eta\rangle}{\lambda^\alpha} \right| < \infty$$

where have set $r = \lceil -\alpha \rceil$ and

$$\mathcal{S}^\lambda_z \eta(s,y) := \lambda^{-d-2} \, \eta\big(\lambda^{-2}(s-t), \lambda^{-1}(y-x)\big) . \tag{2.2}$$

One can adapt Definition 2.3 to the case $\alpha > 0$ (extending Definition 2.1). We first need to define the parabolic degree of a polynomial. Given a multi-index $k = (k_0, k_1, \ldots, k_d) \in \mathbf{N} \times \mathbf{N}^d$ we define the monomial z^k in the standard way, we also define the parabolic degree of this monomial to be $|k|_\mathfrak{s} := 2k_0 + \sum_{j=1}^d k_j$. We then define the parabolic degree of a polynomial $P(z)$ to be the maximum of the parabolic degree of all of its constituent monomials.

DEFINITION 2.4. — *Suppose that $\alpha \geqslant 0$. We define $\mathcal{C}^\alpha_\mathfrak{s}$ to be the set of all functions $u \in \mathcal{S}'(\mathbb{R}^{d+1})$ such that there exist polynomials $\{P_z\}_{z\in\mathbb{R}^{d+1}}$, each of parabolic degree less than α, such that for any compact set $\mathfrak{K} \subseteq \mathbb{R} \times \mathbb{R}^d$*

one has

$$\|u\|_{\mathcal{C}^\alpha_{\mathfrak{s}}(\mathfrak{K})} := \sup_{z \in \mathfrak{K}} \sup_{\substack{\eta \in B_0 \\ \lambda \in (0,1]}} \left| \frac{\langle u - P_z, \mathcal{S}^\lambda_z \eta \rangle}{\lambda^\alpha} \right| < \infty . \tag{2.3}$$

Remark 2.5. — It is easily checked that in the above definition P_z must just be the $\lfloor \alpha \rfloor$-th order Taylor expansion for the function u centered at z.

Remark 2.6. — Important theorems about $\mathcal{C}^\alpha_{\mathfrak{s}}$ spaces (like Theorem 2.8) fail to hold when α takes integer values[6]. We will implicitly assume that any $\mathcal{C}^\alpha_{\mathfrak{s}}$ space appearing in the assumption or conclusion of a theorem uses a non-integer value of α.

We now investigate the regularity of space-time white noise. A calculation similar to (1.4) shows that for $\lambda \in (0,1]$ one has

$$\mathbb{E}\langle \xi, \mathcal{S}^\lambda_z \eta \rangle^2 \lesssim \lambda^{-d-2} . \tag{2.4}$$

This suggests that ξ has regularity $\alpha = -\frac{d}{2} - 1$. The following "Kolmogorov like" theorem which is small variation of [38, Thm. 5.2] shows that this is almost true.

THEOREM 2.7. — *Suppose that we are given a $\mathcal{S}(\mathbb{R}^{d+1})$-indexed stochastic process $\xi(\bullet)$ which is linear (that is a linear map from $\mathcal{S}(\mathbb{R}^{d+1})$ to the space of random variables).*

Fix any $\alpha < 0$ and a $p \geqslant 1$. Suppose there exists a constant C such that for all $z \in \mathbb{R}^{d+1}$, and for all $\eta \in \mathcal{S}(\mathbb{R}^{d+1})$ which are supported on the unit ball of $\mathbb{R}^{d+1}$ and satisfy $\sup_{z' \in \mathbb{R}^{d+1}} |\eta(z')| \leqslant 1$ one has

$$\mathbb{E}\big|\xi(\mathcal{S}^\lambda_z \eta)\big|^p \leqslant C\lambda^{\alpha p} \text{ for any } \lambda \in (0,1] \tag{2.5}$$

then there exists a random distribution $\tilde{\xi}$ in $\mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$ such that for all η we have $\xi(\eta) = \tilde{\xi}(\eta)$ almost surely. Furthermore, for any $\alpha' < \alpha - \frac{d+2}{p}$ and any compact $\mathfrak{K} \subseteq \mathbb{R} \times \mathbb{R}^d$ we have

$$\mathbb{E}\|\tilde{\xi}\|^p_{\mathcal{C}^{\alpha'}_{\mathfrak{s}}(\mathfrak{K})} < \infty .$$

Sketch of proof. — We start by recalling the argument for the classical Kolmogorov criterion for a stochastic process $X(t)$ indexed by $t \in \mathbb{R}$ (ignoring all questions of suitable modifications). The first step is the purely deterministic observation that for any continuous function X we have

$$\sup_{s \neq t \in [-1,1]} \frac{|X(s) - X(t)|}{|s - t|^{\alpha'}} \lesssim \sup_{k \geqslant 0} \sup_{s \in 2^{-k}\mathbf{Z} \cap [-1,1)} 2^{k\alpha'}|X(s + 2^{-k}) - X(s)| .$$

---

[6] For example, when α is a positive integer the Besov space $\mathcal{B}^\alpha_{\infty,\infty}$ differs from the classical Hölder space with parameter α.

This implies (still in a purely deterministic fashion) that

$$
\left(\sup_{s \neq t \in [-1,1]} \frac{|X(s) - X(t)|}{|s-t|^{\alpha'}} \right)^p
$$

$$
\lesssim \sup_{k \geqslant 0} \sup_{s \in 2^{-k}\mathbf{Z} \cap [-1,1)} 2^{k\alpha' p} |X(s + 2^{-k}) - X(s)|^p
$$

$$
\lesssim \sum_{k \geqslant 0} \sum_{s \in 2^{-k}\mathbf{Z} \cap [-1,1)} 2^{k\alpha' p} |X(s + 2^{-k}) - X(s)|^p .
$$

The only stochastic ingredient consists of taking the expectation of this expression which yields

$$
\mathbb{E} \left(\sup_{s \neq t \in [-1,1]} \frac{|X(s) - X(t)|}{|s-t|^{\alpha'}} \right)^p
$$

$$
\lesssim \sup_{s \neq t} \left(\frac{1}{|s-t|^{\alpha p}} \mathbb{E}|X(s) - X(t)|^p \right) \sum_{k \geqslant 0} 2^k 2^{k\alpha' p} 2^{-k\alpha p} ,
$$

and summing the geometric series.

The argument for Theorem 2.7 follows a very similar idea. The crucial deterministic observation is that the Besov norm $\|\xi\|_{\mathcal{C}_{\mathfrak{s}}^{\alpha'}}$ can be controlled by

$$
\|\xi\|_{\mathcal{C}_{\mathfrak{s}}^{\alpha'}(\mathfrak{K})} \lesssim \sup_{k \geqslant 0} \sup_{x \in 2^{-2k}\mathbf{Z} \times 2^{-k}\mathbf{Z}^d \cap \bar{\mathfrak{K}}} 2^{-k\alpha'} \left(\xi, \mathcal{S}_x^{2^{-k}} \eta \right) ,
$$

where $\bar{\mathfrak{K}}$ is another compact set that is slightly larger than $\mathfrak{K}$ and η is a *single*, well chosen test function. There are different ways to construct such a function η e.g. using wavelets (as in [27]) or using the building blocks of the Paley–Littlewood decomposition (as in [38]). The argument then follows the same strategy replacing the sup by a sum and taking the expectation in the last step only. $\square$

Going back to the discussion of white noise, we recall the basic fact that for a Gaussian random variable X control of the second moment gives control on all moments. For all positive integers p there exists a constant C_p such that

$$
\mathbb{E}[|X|^p] \leqslant C_p \big(\mathbb{E}[X^2] \big)^{p/2}.
$$

It follows that for Gaussian processes once one has (2.5) for $p = 2$ then a similar bound holds for all p. Thus we can conclude that ξ has regularity $\mathcal{C}_{\mathfrak{s}}^{-\frac{d}{2}-1-\kappa}$ for every $\kappa > 0$.

2.2. Linear theory

From now on we will assume periodic boundary conditions in space: the spatial variable x will take values in the d-dimensional torus $\mathbf{T}^d$ (which is identified with $[-\pi, \pi]^d$). When more convenient we will sometimes view a function or distribution defined on $\mathbb{R} \times \mathbf{T}^d$ as defined on $\mathbb{R} \times \mathbb{R}^d$ and periodic in space. We first recall Duhamel's principle or the *variation of constants formula*. Consider the inhomogeneous heat equation given by

$$\partial_t u = \Delta u + f$$
$$u(0, \cdot) = u_0$$

where f is a space-time function and u_0 is a spatial initial condition. Under very general regularity assumptions on f and u_0 the solution is given by the formula

$$u(t, x) = \int_0^t \int_{\mathbf{T}^d} K(t - s, x - y)\, f(s, y)\, \mathrm{d}y\, \mathrm{d}s + \int_{\mathbf{T}^d} K(t, x - y) u_0(y)\,. \quad (2.6)$$

Here K is the heat kernel on the torus, which for $t > 0$ and $x \in \mathbf{T}^d$ is given by

$$K(t, x) = \sum_{k \in 2\pi \mathbf{Z}^d} \frac{1}{(4\pi t)^{\frac{d}{2}}} \exp\left(-\frac{(x - k)^2}{t} \right).$$

We extend K to negative times t by defining it to be 0 on $((-\infty, 0] \times \mathbf{T}^d) \setminus \{(0,0)\}$. We will then view K as smooth function on $\mathbb{R} \times \mathbf{T}^d \setminus \{0,0\}$. When f is a space-time distribution and/or u_0 is a space distribution the right hand side of (2.6) is a formal expression but in many cases it can be made rigorous via the duality pairing between distributions and test functions. More precisely, we say that ξ is a distribution on $(0, \infty) \times \mathbf{T}^d$ if it is a distribution on $\mathbb{R} \times \mathbf{T}^d$ which vanishes when tested against test-functions that are supported in $(-\infty, 0] \times \mathbf{T}^d$. Note that in general it is not possible, to multiply a distribution with the indicator function $\mathbf{1}_{(0,\infty)}(t)$, so that even in a distributional sense the integral over $(0, t)$ cannot always be defined (think e.g. of the distribution $\varphi \mapsto \mathrm{P.V.} \int \frac{1}{t} \varphi \, \mathrm{d}t$ on $\mathbb{R}$). However, for white noise ξ it is easy to define $\xi \mathbf{1}_{(0,\infty)}(t)$ as an element of $\mathcal{C}^{-\frac{d+2}{2} - \kappa}$.

To keep our exposition simple we will always assume that the initial condition u_0 is zero. We now give an important classical result (essentially a version of the parabolic Schauder estimates, see e.g. [34], for a statement which implies our version see [27, §5]). In what follows Λ_t denotes the domain of integration in (2.6), that is $\Lambda_t := (0, t) \times \mathbf{T}^d$ and we use $\Lambda = (0, \infty) \times \mathbf{T}^d$.

Ajay Chandra and Hendrik Weber

THEOREM 2.8 (Schauder Estimate). — *For f in $\mathcal{C}_{\mathfrak{s}}^{\alpha}(\Lambda)$ define*

$$u(s,x) := \int_{\Lambda_s} K(s-r, x-y)\, f(r,y)\, \mathrm{d}y\, \mathrm{d}r \qquad (2.7)$$

interpreted in a distributional sense if $\alpha < 0$. Then if $\alpha \notin \mathbf{Z}$, we have

$$\|u\|_{\mathcal{C}_{\mathfrak{s}}^{\alpha+2}(\Lambda_t)} \lesssim \|f\|_{\mathcal{C}_{\mathfrak{s}}^{\alpha}(\Lambda_t)} \,.$$

The Schauder estimate shows that the use of parabolic scaling of space-time is natural when measuring regularity. We do not give a proof of this result here; compare however to the discussion of the integration map for regularity structures in Section 4.4 below.

We now apply Duhamel's principle to the stochastic heat equation (SHE) (again with vanishing initial condition). Formally the solution is given by

$$Z(t,x) = \int_{\Lambda_t} K(t-s, x-y)\, \xi(s,y)\, \mathrm{d}y\, \mathrm{d}s \,. \qquad (2.8)$$

The standard approach is to view Z as a stochastic integral (the resulting object is called a stochastic convolution). However we can also define Z deterministically for each fixed realization of white noise. Each such realization of white-noise will be an element of $\mathcal{C}_{\mathfrak{s}}^{-\frac{d}{2}-1-\kappa}(\mathbb{R} \times \mathbf{T}^d)$, the Schauder estimate then implies that $Z \in \mathcal{C}_{\mathfrak{s}}^{1-\frac{d}{2}-\kappa}(\mathbb{R} \times \mathbf{T}^d)$ for every $\kappa > 0$. It follows that Z is a continuous function in $d = 1$ while for $d \geqslant 2$ we expect Z to be a distribution.

Instead of using the Schauder estimate we can also get a handle on the regularity of Z by establishing the estimate (2.5) for $p = 2$ (since Z is Gaussian). This is an instructive calculation and it gives us a good occasion to introduce *graphical notation* in familiar terrain. From now on we denote the process Z introduced in (2.8) by $\uparrow$. This is to be read as a simple graph, where the circle at the top represents an occurrence of the white noise and the line below represents an integration against the heat kernel. As above we will use the convention to combine the space and time variable into a single variable $z = (t,x)$. With these conventions testing $\uparrow$ against the rescaled test-function $\mathcal{S}_z^{\lambda}\eta$, defined as above in (2.2), yields

$$(\uparrow, \mathcal{S}_z^{\lambda}\eta) = \int_{\Lambda}\int_{\Lambda} \mathcal{S}_z^{\lambda}\eta(z_1)\, K(z_1 - z_2)\, \mathrm{d}z_1\, \xi(\mathrm{d}z_2) \,. \qquad (2.9)$$

Then, using the characterising property (1.3) of white noise we get

$$\mathbb{E}\big[(\uparrow, \mathcal{S}_z^{\lambda}\eta)^2\big]$$
$$= \int_{\Lambda}\int_{\Lambda}\int_{\Lambda} \mathcal{S}_z^{\lambda}\eta(z_1)\mathcal{S}_z^{\lambda}\eta(\bar{z}_1) K(z_1 - z_2) K(\bar{z}_1 - z_2)\, \mathrm{d}z_2\, \mathrm{d}z_1 \mathrm{d}\bar{z}_1 \,. \quad (2.10)$$

The only property of the kernel K that enters our calculations is how quickly it grows near the origin. This motivates the following definition.

DEFINITION 2.9. — *Given a real number ζ, we say a function $G :$ $\mathbb{R}^{d+1} \setminus \{0\} \mapsto \mathbb{R}$ is a kernel of order ζ if*

$$\|G\|_\zeta := \sup_{z \in \mathbb{R}^{d+1} \setminus \{0\}} |G(z)| \times \|z\|_{\mathfrak{s}}^\zeta < \infty .$$

Then one has the following easy lemma.

LEMMA 2.10. — *The kernel $K(z)$ of (2.8) and (2.9) is of order $-(d_{\mathfrak{s}}-2)$ where $d_{\mathfrak{s}} = d + 2$ is the scaled dimension introduced above.*

We now introduce a graphical representation of the *integral* (2.9). Of course at this stage one could still evaluate the integral by hand easily, but this formalism becomes very convenient for more complicated integrals. In this graphical formalism we represent (2.10) by

$$\mathbb{E}\big[(\uparrow, \mathcal{S}_z^\lambda \eta)^2\big] \;=\; \underset{-d_{\mathfrak{s}}+2}{\bigcirc}\!\!\rule[0.3ex]{2em}{0.4pt}\!\blacksquare\!\rule[0.3ex]{2em}{0.4pt}\!\!\underset{-d_{\mathfrak{s}}+2}{\bigcirc} . \tag{2.11}$$

Again, each line represents an occurrence of the kernel K and the order is denoted below. The black square in the middle represents an integration over the space-time Λ and the grey vertices at the sides represent an integration against the scaled test-function $\mathcal{S}_z^\lambda \eta$. Note that there is a simple "graphical derivation" of (2.11) which consists of "gluing" the dots in two copies of $\uparrow$ together.

The following lemma (essentially [27, Lem. 10.14]) is simple but extremely helpful, because it permits to perform the analysis of integrals, which are potentially much more complicated than (2.10) on the level of graphs, by recursively reducing the complexity of the graph, keeping track only of the relevant information.

LEMMA 2.11. — *Let K_1, K_2 be kernels of order $\zeta_1, \zeta_2 \in \mathbb{R}$ with compact support. Then their product $K_1 K_2$ is a kernel of order $\zeta_1 + \zeta_2$ and we have $\|K_1 K_2\|_{\zeta_1 + \zeta_2} \leqslant \|K_1\|_{\zeta_1} \|K_2\|_{\zeta_2}$. If furthermore,*

$$\zeta_1 \wedge \zeta_2 > -d_{\mathfrak{s}} \qquad and \qquad \zeta = \zeta_1 + \zeta_2 + d_{\mathfrak{s}} < 0 , \tag{2.12}$$

then $K_1 * K_2$ (where $*$ denotes the convolution on $\mathbb{R} \times \mathbf{T}^d$) is a kernel of order ζ and we have $\|K_1 * K_2\|_\zeta \lesssim \|K_1\|_{\zeta_1} \|K_2\|_{\zeta_2}$.

Remark 2.12. — The first condition in (2.12) is necessary in order to ensure that the convolution $K_1 * K_2$ is well defined. The integration is restricted to a *compact* space-time domain, so that we only have to deal with convergence at the singularities, but of course the constant depends on the choice of domain. The second condition ensures that the resulting kernel

actually does have a blowup at the origin. In the case $\zeta_1 + \zeta_2 + d_\mathfrak{s} = 0$ it is in general *not* true that $K_1 * K_2$ is bounded. However, one can obtain a bound with only a logarithmic divergence at the origin (we will see that in our discussion of the two dimensional stochastic heat equation below). There is in general no reason to expect that if $\zeta_1 + \zeta_2 + d_\mathfrak{s} > 0$ we will have $K_1 * K_2(0) = 0$. In this case $K_1 * K_2$ is not the correct object to work with, one must also subtract a partial Taylor expansion.

We now apply this result to the integral over z_2 appearing in (2.10). Note that if η has compact support this integration is over a compact space-time domain (depending on the point z). For $d \geqslant 3$, (i.e. $d_\mathfrak{s} \geqslant 5$) condition (2.12) is satisfied and we can replace the convolution of the kernels by a single kernel of order $-d_\mathfrak{s} + 4$. In our convenient graphical notation this can be written as

$$\circ\!\!\underset{-d_\mathfrak{s}+2}{\rule{2cm}{0.4pt}}\!\!\blacksquare\!\!\underset{-d_\mathfrak{s}+2}{\rule{2cm}{0.4pt}}\!\!\circ \quad \lesssim \quad \circ\!\!\underset{-d_\mathfrak{s}+4}{\rule{2cm}{0.4pt}}\!\!\circ \; . \tag{2.13}$$

At this stage it only remains to observe that for η with compact support and for $\zeta > -d_\mathfrak{s}$

$$\int_\Lambda \int_\Lambda \mathcal{S}_z^\lambda \eta(z_1) \, \mathcal{S}_z^\lambda \eta(\bar{z}_1) \|z_1 - \bar{z}_1\|_\mathfrak{s}^\zeta \mathrm{d}z_1 \, \mathrm{d}\bar{z}_1 \lesssim \lambda^\zeta \; ,$$

and we have derived (2.5) and therefore the right regularity of $\uparrow$ at least for $d \geqslant 3$. For $d = 2$ we are in the critical case $-d_\mathfrak{s} + 4 = 0$. According to Remark 2.12 the inequality (2.13) remains valid if we interpret a kernel of order 0 as a logarithmically diverging kernel.

For $d = 1$ condition (2.12) fails and we cannot use the same argument to derive the desired $\frac{1}{2}-$ regularity. This is due to the fact that in order to obtain *positive* regularity, $(\uparrow, \mathcal{S}_z^\lambda \eta)$ is plainly the wrong quantity to consider. As observed in Definition 2.4 rather than bounding the blowup of local averages of $\uparrow$ near z we need to control how fast these local averages go to zero if a suitable *polynomial approximation* (the Taylor polynomial) is subtracted. In the case of $\uparrow$ we aim to show $\frac{1}{2}-$ regularity, so we need to control how quickly $(\uparrow - \uparrow(z), \mathcal{S}_z^\lambda \eta)$ goes to zero for small λ. This observation may seem harmless, but we will encounter it again and again below. Arguably much of the complexity of the theory of regularity structures is due to the extra terms we encounter when we want to obtain bounds on a quantity of *positive* regularity (or *order*). In this particular case it is not too difficult to modify the graphical argument to get a bound on $(\uparrow - \uparrow(z), \mathcal{S}_z^\lambda \eta)$. The integral (2.10)

turns into

$$\mathbb{E}\big[(\uparrow - \uparrow(z), \mathcal{S}_z^\lambda \eta)^2\big]$$
$$= \int_\Lambda \int_\Lambda \int_\Lambda \mathcal{S}_z^\lambda \eta(z_1) \mathcal{S}_z^\lambda \eta(\bar{z}_1)$$
$$\times \big(K(z-z_2)-K(z_1-z_2)\big)\big(K(z-z_2)-K(\bar{z}_1-z_2)\big)\, \mathrm{d}z_2\, \mathrm{d}z_1 \mathrm{d}\bar{z}_1\ .$$

Now we need to use the fact that not only K has the right blowup near the diagonal but also its derivatives. More precisely, for every multi-index k we have that

$$|D^k K(z)| \lesssim \|z\|_{\mathfrak{s}}^{-d_{\mathfrak{s}}+2-|k|_{\mathfrak{s}}}\ .$$

In fact, these additional bounds are imposed in the version of Definition 2.9 found in [27] and also appear in some statements of harmonic analysis relating to singular kernels. In [27, Lem. 10.18] it is shown how the kernel $K(z - z_2) - K(z_1 - z_2)$ can be replaced by a "Taylor approximation" $DK(z_1 - z_2)(z - z_1)$. The factor $(z - z_1)$ can then be pulled out of the convolution integral over z_2 and the "graphical algorithm" can be applied to the convolution of two copies of DK which *do* satisfy (2.12).

2.3. Nonlinear equations

For non-linear equations Duhamel's principle turns into a fixed point problem. We illustrate this for equation (Φ_d^4) in one spatial dimension where one gets

$$\varphi(t,x) = \int_0^t \int_{\mathbf{T}^1} K(t-s, x-y)\, \xi(s,y)\, \mathrm{d}y\, \mathrm{d}s$$
$$- \int_0^t \int_{\mathbf{T}^1} K(t-s, x-y)\, \varphi^3(s,y)\, \mathrm{d}y\, \mathrm{d}s\ . \quad (2.14)$$

For simplicity we have dropped the mass term $m^2 \varphi$ and set the initial condition to be 0. The Schauder estimate tells us that the first term on the right hand side of (2.14) is in $\mathcal{C}_{\mathfrak{s}}^{\frac{1}{2}-\kappa}(\mathbb{R}_+ \times \mathbf{T}^1)$ for any κ. The following theorem characterizes when we can understand products like φ^3 classically.

THEOREM 2.13 ([27, Prop. 4.11], see also [3, §2.6]). — *Suppose that $\alpha + \beta > 0$, then there exists a bilinear form $B(\,\cdot\,, \cdot\,) : \mathcal{C}_{\mathfrak{s}}^\alpha \times \mathcal{C}_{\mathfrak{s}}^\beta \to \mathcal{C}_{\mathfrak{s}}^{\alpha \wedge \beta}$ such that*

- *For smooth functions f, g one has that $B(f,g)$ coincides with the point-wise product of f and g.*
- *For arbitrary $f \in \mathcal{C}_{\mathfrak{s}}^\alpha$, $g \in \mathcal{C}_{\mathfrak{s}}^\beta$ one has*

$$\|B(f,g)\|_{\mathcal{C}_{\mathfrak{s}}^{\alpha \wedge \beta}} \lesssim \|f\|_{\mathcal{C}_{\mathfrak{s}}^\alpha} \times \|g\|_{\mathcal{C}_{\mathfrak{s}}^\beta}\ .$$

Additionally, if $\alpha + \beta \leqslant 0$ then no bilinear form $B(\,\cdot\,,\,\cdot\,) : \mathcal{C}_{\mathfrak{s}}^{\alpha} \times \mathcal{C}_{\mathfrak{s}}^{\beta} \to \mathcal{C}_{\mathfrak{s}}^{\alpha \wedge \beta}$ satisfying both of the above statements exists.

It is then natural to treat (2.14) as a fixed point problem in $\mathcal{C}_{\mathfrak{s}}^{\frac{1}{2}-\kappa}(\mathbb{R}_+ \times \mathbf{T}^1)$ (by Theorem 2.13 the definition of φ^3 poses no difficulty in this space). For any fixed realisation of ξ there exists $T(\xi) > 0$ such that the operator

$$\Psi \; : \; \varphi \mapsto \int_0^{\cdot} \int_{\mathbf{T}^1} K(\,\cdot\, - s, \cdot - y)\, \xi(s,y)\, \mathrm{d}y\, \mathrm{d}s$$

$$- \int_0^{\cdot} \int_{\mathbf{T}^1} K(\,\cdot\, - s, \cdot - y)\, \varphi^3(s,y)\, \mathrm{d}y\, \mathrm{d}s \quad (2.15)$$

is a contraction on bounded balls in $\mathcal{C}_{\mathfrak{s}}^{\frac{1}{2}-\kappa}([0,T] \times \mathbf{T}^1)$. An important observation is that $v = \varphi - \uparrow = -\int_0^t \int_{\mathbf{T}^1} K(t - s, x - y)\, \varphi^3(s,y)\, \mathrm{d}y\, \mathrm{d}s$ is much more regular than φ itself, in fact the Schauder estimate implies that it is a $\mathcal{C}_{\mathfrak{s}}^{\frac{5}{2}-\kappa}$ function. It is important to note that this argument does *not* make use of the sign of the nonlinear term $-\varphi^3$. Of course, this sign is essential when deriving bounds that imply non-explosion, the existence of invariant measures for solutions, or even getting existence and uniqueness when $\mathbf{T}^1$ is replaced by $\mathbb{R}$.

For $d \geqslant 2$ it is not so easy to solve the fixed point problem (2.14) (with the one-dimensional torus $\mathbf{T}^1$ replaced by $\mathbf{T}^d$). As we have seen above the stochastic convolution $\uparrow$ only takes values in the distributional spaces $\mathcal{C}_{\mathfrak{s}}^{-\frac{2-d}{2}-\kappa}$ but there is no canonical way to define the mapping $\varphi \mapsto \varphi^3$ for $\varphi \in \mathcal{C}_{\mathfrak{s}}^{\alpha}$ with $\alpha < 0$. We will now try to find a way around this issue in the case of $d \geqslant 2$, we start by running a Picard iteration step by step. More precisely we set $\varphi_0 = 0$ and aim to study the behaviour of the sequence $\{\varphi_n\}_{n=0}^{\infty}$ defined recursively as

$$\varphi_{n+1} = \Psi(\varphi_n)\,,$$

where Ψ is defined in (2.15) (with $\mathbf{T}^1$ replaced by $\mathbf{T}^d$).

With our choice of $\varphi_0 = 0$ the first step in the Picard iteration yields $\varphi_1 = \uparrow$ which is of regularity $\mathcal{C}_{\mathfrak{s}}^{\frac{2-d}{2}-\kappa}$. When going to φ_2 we immediately run into trouble when we try to apply Ψ to $\uparrow$ since this requires us to define $\uparrow^3$ for which Theorem 2.13 is of no use.

So far our analysis of (2.15) could be performed entirely deterministically (occuring for a fixed realization of ξ) but at this point it becomes essential to use a probablistic approach. While there is no canonical way of cubing an arbitrary distribution of negative regularity, we will now see that there are ways to define polynomials in $\uparrow$ by exploiting its Gaussian structure.

2.3.1. Construction of Wick powers

We will define $\uparrow^3$ by approximation. The calculations will be performed in the framework of *iterated stochastic integrals*. Definition and elementary properties of these are recalled in Appendix A. Let ϱ_δ be a smoothing kernel on scale δ (as was used in (1.9)) and set

$$\uparrow_\delta(z) := \uparrow * \varrho_\delta(z) = \int_\Lambda K * \varrho_\delta(z - \bar{z})\, \xi(\mathrm{d}\bar{z})\,. \tag{2.16}$$

For every $\delta > 0$ the random function $\uparrow_\delta(z)$ is smooth and we can define $\uparrow_\delta(z)^3$ without ambiguity. To analyse the behaviour of $\uparrow_\delta^3$ as $\delta \to 0$ we interpret $\uparrow_\delta$ as a stochastic integral against ξ and apply (A.5) which gives

$$\uparrow_\delta(y)^3 = \int \mathcal{W}_\delta^{(3)}(y; z_1, z_2, z_3)\xi(\mathrm{d}z_1)\,\xi(\mathrm{d}z_2)\,\xi(\mathrm{d}z_3)$$
$$+ 3 \int \mathcal{W}_\delta^{(1)}(y; z_1)\,\xi(\mathrm{d}z_1)\,, \tag{2.17}$$

where

$$\mathcal{W}_\delta^{(3)}(y; z_1, z_2, z_3) = \prod_{j=1}^{3} \left[K * \varrho_\delta(y - z_j) \right],$$

and

$$\mathcal{W}_\delta^{(1)}(y; z_1) = K * \varrho_\delta(y - z_1) \int \left(K * \varrho_\delta(y - z) \right)^2 \mathrm{d}z\,. \tag{2.18}$$

As before, we will introduce a graphical notation to analyse these expressions. In this notation (2.17) becomes

$$\uparrow_\delta(y)^3 = \mathbf{\Psi}_\delta + 3 \ \ \mathbf{O}_\delta \ \uparrow_\delta\,.$$

As before, each black dot represents an occurrence of the space-time white noise, and each line represents an integration against a singular kernel. The black square appearing in the second term is a dummy variable which is integrated out. The subscript δ appearing in all the graphs represents the fact that all singular kernels are regularised at scale δ, i.e. K is replaced by $K * \varrho_\delta$ which satisfies

$$|K * \varrho_\delta(z)| \lesssim \frac{1}{(\|z\|_{\mathfrak{s}} + \delta)^{d_{\mathfrak{s}} - 2}}\,.$$

Applying the same graphical rule as above, we get

$$\mathbb{E}\big[(\Psi_\delta, \mathcal{S}_z^\lambda \eta)^2\big] \quad = \quad \text{} \quad \lesssim \quad \text{}$$

$$\lesssim \quad \text{}, \qquad (2.19)$$

uniformly in δ. For $d = 3$, i.e. for $d_{\mathfrak{s}} = 5$, we have $-3d_{\mathfrak{s}} + 12 = 3 > -d_{\mathfrak{s}}$, which yields the uniform-in-delta bound

$$\mathbb{E}\big[(\Psi_\delta, \mathcal{S}_z^\lambda \eta)^2\big] \lesssim \lambda^{-3}\ ,$$

while in the case $d = 2$ we get as above

$$\mathbb{E}\big[(\Psi_\delta, \mathcal{S}_z^\lambda \eta)^2\big] \lesssim |\log(\lambda)|^3.$$

However the lower order *Itô correction* $3 \int \mathcal{W}_\delta^{(1)}(y; z_1)\, \xi(\mathrm{d}z_1)$ will be a problem in the $\delta \downarrow 0$ limit. The explicit form (2.18) of the kernel $\mathcal{W}_\delta^{(1)}$ shows that it can be rewritten as $3C_\delta Z$ where C_δ is a constant given by

$$C_\delta := \int \big(K * \varrho_\delta(z)\big)^2 \mathrm{d}z\ . \qquad (2.20)$$

For $\delta \downarrow 0$ these C_δ diverge logarithmically for $d = 2$ and like $\frac{1}{\delta}$ for $d = 3^{(7)}$. To overcome this problem we simply remove the diverging term $3C_\delta Z$. From our second moment bound, the Nelson estimate (see (A.6) in Appendix A), and Theorem 2.7 one can then show that the limit

$$\Psi \ := \lim_{\delta \downarrow 0} \big(\uparrow_\delta^3 - 3C_\delta \uparrow_\delta\big)$$

exists as random elements of $\mathcal{C}_{\mathfrak{s}}^{-\frac{3}{2}-\kappa}$ for $d = 3$ and as random elements of $\mathcal{C}_{\mathfrak{s}}^{-\kappa}$ for $d = 2$, where the convergence holds for every stochastic L^p space. The subtraction implemented above is called *Wick renormalization*, and the object Ψ is called the third Wick power of $\uparrow$ and is sometimes denoted by $: Z^3 :$ (we could write $: \uparrow^3 :$ to be more consistent with our graphical notation).

---

[7] Actually, there is a slight cheat in (2.20) because we do not specify the domain of integration. In each case C_δ does not depend on the spatial variable y, but if we define C_δ as an integral over Λ_t then it actually depends on t which one may consider ugly. But the integral over Λ_t can be decomposed into a part which does not depend on t and which diverges as $\delta \to 0$ (e.g. the integral over $[0,1] \times \mathbf{T}^d$) and a part which depends on t but remains bounded in δ and which can be ignored in the renormalization procedure. There are many ways to choose C_δ in a t-independent way. None of these choices is canonical but all only differ by a quantity that remains bounded as $\delta \to 0$.

The general recipe for defining Wick powers is as follows: to define $: Z^n :$ one applies the n-th order analog of the identity (A.5) to Z_δ^n, drops *all* lower-order Itô corrections[8] to get an object we denote $: Z_\delta^n :$, and then takes the $\delta \downarrow 0$ limit. The graphical analysis of n-th Wick powers is very similar to (2.19), the only difference being that there are n edges connecting the left and right vertices. In this way one obtains a singular kernel of order $-nd_{\mathfrak{s}} + 4n$ as a final result. Hence for $d = 2$ the blowup of this kernel on the diagonal can be bounded by $|\log(z_1 - \bar{z}_1)|^n$ which is integrable for every n. For $d = 3$ however we get a polynomial blowup $|z_1 - \bar{z}_1|^{-n}$ which fails to be integrable for $n \geqslant 5$. For $d = 2$ we can define arbitrary Wick powers of $Z = \uparrow$ while for $d = 3$ we can only define Wick powers up to $n = 4$.

Remark 2.14. — Our reasoning shows that in the three dimensional case we can define Wick powers up to order $n = 4$ as space-time distributions. It is however *not* possible to evaluate these distribution for fixed t in the cases $n \geqslant 3$. Only space time averages are well defined.

2.3.2. Back to the Picard iteration

We now return to our Picard iteration, still working formally. The process $: \uparrow^3 :$ is denoted by $\mathbf{\Psi}$, where again each dot represents an occurrence of white noise and each line represents one integration against a kernel. The fact that they are merged at the bottom corresponds to multiplication. For now we will just replace the Z^3 that would have appeared in φ_2 with $\mathbf{\Psi}$ so that we have

$$\varphi_2 = \uparrow - \mathbf{\Psi}$$

where $\mathbf{\Psi} = K * \mathbf{\Psi}$. In the next step of the Picard iteration we would get

$$\varphi_3 = \uparrow - K * \left(\mathbf{\Psi} - 3\uparrow^2 K * (\mathbf{\Psi}) + 3\uparrow \left(K * \mathbf{\Psi} \right)^2 - \left(K * \mathbf{\Psi} \right)^3 \right).$$

If we restrict to the case $d = 2$ then almost all of these terms are well defined. Indeed, according to the Schauder estimates $K * \mathbf{\Psi}$ is a function of class $\mathcal{C}_{\mathfrak{s}}^{2-\kappa}$ for any $\kappa > 0$. And this is enough to define most of the products. The only term that causes a problem is the term $\uparrow^2 = Z^2$, however the corresponding Wick power $\mathbf{V} := \lim_{\delta \downarrow 0}(\uparrow_\delta^2 - C_\delta)$ is well defined.

It turns out that these are all the terms that need to be renormalised when $d = 2$, and that after modifying these first few steps, the Picard iteration can actually be closed. Of course we have been working somewhat formally here, instead of replacing certain powers Z^n with Wick powers $: Z^n :$ one

---

[8] By lower order we mean all the terms involving strictly less than n factors of ξ.

should instead modify the equation so it automatically generates the needed Wick renormalizations. In the next section we will explain in more detail, how the above method of treating Φ_2^4 can be implemented and we explain why this approach fails for Φ_3^4.

Lecture 3

The renormalization and the Picard iteration for Φ_2^4 were performed in a very elegant way in [11] by a method we call the Da Prato–Debussche argument. We start this lecture by discussing this argument and then sketch why it fails for Φ_3^4. This motivates us to turn to a more robust approach, the theory of regularity structures [27]. In particular we will introduce some of basic objects of the theory: *regularity structures*, *models*, and *modelled distributions*.

3.1. The Da Prato–Debussche Argument

At the end of Lecture 2 we expanded the solution φ of the Φ_2^4 equation in terms of objects built out of the linear solution $\uparrow$. A key observation is that the most singular term in our partial expansion of φ was $\uparrow$, if we write $\varphi = \uparrow + v$ then we expect the remainder v to be of better regularity. While we are unable to directly treat Φ_2^4 equation as fixed point problem in a $\mathcal{C}_{\mathfrak{s}}^\alpha$ space, it turns out that one can renormalize the original equation so that it generates Wick powers of $\uparrow$ and then solve a fixed point equation in a nicer space for the remainder v. As we already announced in (1.11), the renormalized equation is

$$\partial_t \varphi_\delta = \Delta \varphi_\delta - (\varphi_\delta^3 - 3C_\delta \varphi_\delta) + \xi_\delta , \qquad (3.1)$$

where C_δ is given by (2.20). Now we write $\varphi_\delta = \uparrow_\delta + v_\delta$ where $\uparrow_\delta$ is given by (2.16) so that it solves $\partial_t \uparrow_\delta = \Delta \uparrow_\delta + \xi_\delta$. Subtracting this linear equation from (3.1) gives us

$$\begin{aligned}
\partial_t v_\delta &= \Delta v_\delta - \big((v_\delta + \uparrow_\delta)^3 - 3C_\delta(v_\delta + \uparrow_\delta)\big) \\
&= \Delta v_\delta - v_\delta^3 - 3\uparrow_\delta v_\delta^2 - 3(\uparrow_\delta^2 - C_\delta)v_\delta - (\uparrow_\delta^3 - 3C_\delta \uparrow_\delta) .
\end{aligned} \qquad (3.2)$$

This equation looks more promising since the rough driving noise ξ_δ has dropped out and from the previous lecture we know that the polynomials in $\uparrow_\delta$ appearing above converge in probability to the corresponding Wick powers as $\delta \to 0$. We pass to the limit and try to solve the fixed point equation

$$v = K * \left[-v^3 - 3\uparrow v^2 - 3\mathcal{V}v - \mathcal{V} \right] . \qquad (3.3)$$

Recall that $\uparrow$, V, and $\mathsf{\Psi}$ are in $\mathcal{C}_{\mathfrak{s}}^{-\kappa}$ for any (small) $\kappa > 0$, when $d = 2$. Using Theorem 2.8 and Theorem 2.13 we can formulate (3.3) in $\mathcal{C}_{\mathfrak{s}}^{2-\kappa}$, the key point being that all products on the right hand side of (3.3) make sense in $\mathcal{C}_{\mathfrak{s}}^{2-\kappa}$. By exploiting the sign of v^3 in (3.2) one can also show global in time existence for v (and therefore for φ as well), see [38].

Remark 3.1. — Remarkably a similar argument was originally discovered by Bourgain in the context of the two dimensional non-linear Schrödinger equation with defocussing cubic non-linearity. More precisely, in [6] Bourgain studied the deterministic PDE

$$i\partial_t \varphi = \Delta\varphi - \varphi^3 \ .$$

When written in the mild form

$$\varphi(t) = e^{-i\Delta t}\varphi(0) - \int_0^t e^{-(t-s)i\Delta}\varphi^3(s)\,\mathrm{d}s \tag{3.4}$$

it resembles (2.14) with the important difference that unlike the heat semigroup the Schrödinger semigroup $e^{-it\Delta}$ does not improve differentiability. Bourgain studied (3.4) when the initial datum $\varphi(0)$ is a complex Gaussian free field on the torus in which case $z(t) = e^{-i\Delta t}\varphi(0)$ is a Gaussian evolution with regularity properties identical to those of the process $\uparrow$. He then performed the same Wick renormalisation for the square and the cube of $z(t)$ and showed that the equation for the remainder $v = \varphi - z$ can be solved as a fixed point problem in a space of function of positive differentiability. This is a remarkable result because, as said above, the Schrödinger semigroup does not usually improve regularity. See e.g. [39] for recent work in this direction.

The above argument does not apply for Φ_3^4. In this case $\mathsf{\Psi} \in \mathcal{C}_{\mathfrak{s}}^{-3/2-\kappa}$ so we expect the remainder v to be no better than $\mathcal{C}_{\mathfrak{s}}^{1/2-\kappa}$. Since $\mathsf{V} \in \mathcal{C}_{\mathfrak{s}}^{-1-\kappa}$ we fall far short of the regularity needed to define the product $\mathsf{V}v$. One might try to defeat this obstacle by pushing the expansion further, writing $\varphi = \uparrow - \mathsf{Y} + v$ and solving for v. The new fixed point equation is

$$v = K * \left[-v^3 - 3\uparrow v^2 - 3\mathsf{V}v - 3(\mathsf{Y})^2v - 3(\mathsf{Y})v^2 - 6\uparrow(\mathsf{Y})v - 3\mathsf{V}(\mathsf{Y}) - 3\uparrow(\mathsf{Y})^2\right].$$

Since we have pushed the expansion further we do not see the term $\mathsf{\Psi}$ anymore. However we are now confronted with the product $\mathsf{V}\mathsf{Y}$ which cannot be defined using Theorem 2.13 since $\mathsf{Y} \in \mathcal{C}_{\mathfrak{s}}^{1/2-\kappa}$ and $\mathsf{V} \in \mathcal{C}_{\mathfrak{s}}^{-1-\kappa}$. In fact, this ill-defined product is the reason for the second logarithimically diverging renormalization constant for Φ_3^4. But after defining this product by inserting another renormalization constant by hand, we are still unable to close the Picard iteration. The real problematic term is the product $\mathsf{V}v$ which creates a vicious circle of difficulty. If we could define the product $\mathsf{V}v$ then it would have regularity $\mathcal{C}_{\mathfrak{s}}^{-1-\kappa}$, this means at best one could have $v \in \mathcal{C}_{\mathfrak{s}}^{1-\kappa}$. However,

this is not enough regularity to define the product $\mathsf{V}v$ and so we are unable to close the fixed point argument.

3.2. Regularity Structures

The Da Prato–Debussche argument for Φ_2^4 consisted of using stochastic analysis to control a finite number of explicit objects built out of the linear solution followed by the application of a completely deterministic fixed point argument in order to solve for a relatively smooth remainder term. For Φ_3^4 we saw that regardless of how far one expands φ, writing

$$\varphi = \mathsf{1} - \mathsf{Y} + \cdots + v \,, \tag{3.5}$$

the product $\mathsf{V}v$ always prevents us from formulating a fixed point argument for v. We cannot make the remainder v arbitrarily smooth just by pushing the expansion further.

In the theory of regularity structure we will again postulate an expansion for φ which looks more like

$$\Phi(z) = \Phi_{\mathsf{1}}(z)\mathsf{1} + \Phi_{\mathsf{Y}}(z)\mathsf{Y} + \cdots + \Phi_{\mathbf{1}}(z)\mathbf{1} \,. \tag{3.6}$$

One immediately visible difference is that the expansion (3.6) allows varying coefficients in front of various stochastic objects. Instead of solving a fixed point equation for a single function v, we will instead solve a fixed point equation for a family of functions $(\Phi_{\mathsf{1}}, \Phi_{\mathsf{Y}}, \ldots, \Phi_{\mathbf{1}})$. We will also be interested in something called the "order"[9] of objects $\mathsf{1}, \ldots, \mathsf{Y}, \ldots$ in (3.6) instead of just their regularity. The order of an object describes how much we expect it to vanish or blow up when evaluated at small scales, one of the main goals of this section is to clarify the concept of "order". Finally, while the objects $\mathsf{1}, \ldots, \mathsf{Y}, \ldots$ appearing in (3.6) are related to the corresponding stochastic objects in (3.5), they will turn out to be a totally different sort of object so we have distinguished them by coloring them blue.

In [27] the fixed point problem associated with SPDE is solved in an abstract setting. The rest of this lecture will be devoted to introducing the hierarchy of objects that populate this abstract setting and we begin by defining the most basic object.

DEFINITION 3.2 ([27, Def. 2.1]). — *A regularity structure $\mathcal{T}$ consists of a triple (A, T, G).*

- *$A \subseteq \mathbb{R}$ is an indexing set which is bounded from below and has no accumulation points.*

---

[9] What we call order is referred to as homogeneity by Hairer in [28] and [27].

- $T = \bigoplus_{\alpha \in A} T_\alpha$ *is a graded vector space where each T_α is a finite dimensional*[10] *real vector space which comes with a distinguished basis and norm.*
- *G is a family of linear transformations on T with the property that for every $\Gamma \in G$, every $\alpha \in A$, and every $\tau \in T_\alpha$ one has*

$$\Gamma \tau - \tau \in T_\alpha^- \ , \qquad (3.7)$$

where we have set

$$T_\alpha^- := \bigoplus_{\beta < \alpha} T_\beta \ .$$

Additionally we require that G form a group under composition.

In the triple (A, T, G) the set A is an indexing set that lists the orders of the objects that we allow to appear in our expansions. We will always assume $0 \in A$. For any $\alpha \in A$ an element $\tau \in T_\alpha$ should be thought of as an abstract symbol that represents an object of order α. For such a "homogenous" element τ we write $|\tau| = \alpha$. We denote by $\|\cdot\|_\alpha$ the norm on T_α[11]. For general $\tau \in T$ we set $\|\tau\|_\alpha := \|\mathcal{Q}_\alpha \tau\|_\alpha$ where $\mathcal{Q}_\alpha : T \to T_\alpha$ is just projection onto the α-component.

Returning to (3.6), the objects $\uparrow$ and $\mathbb{Y}$ no longer represent fixed space-time distributions but instead are abstract symbols which are homogenous elements of T. The object Φ in (3.6) is actually a map $\Phi : \mathbb{R}^{d+1} \to T$. The family of linear transformations G, called the *structure group*, will play an important role in the theory but we will introduce it slowly as we introduce examples of increasing complexity[12].

3.3. An abstract generalization of Taylor expansions

While (3.5) is a perturbative expansion generated by Picard iteration, one should think of (3.6) as a jet[13], at each space-time point this expansion represents the solution as a linear combination of objects that vanish (or blow up) at controlled rates when evaluated near that space-time point. We will

---

[10] Actually, in [27] these spaces are note required to be finite-dimensional, but in most examples we are aware of they are even of very low dimension.

[11] Since all norms on such T_α are equivalent we may not fix a specific one when defining a regularity structure.

[12] In practice we will not explicitly define the *entire* structure group G when we encounter more complex regularity structures $\mathcal{T}$, only a small subgroup germane to our discussion.

[13] More specifically, a collection of Taylor expansions indexed by space-time "base-points".

now show how the actual Taylor expansions familiar to a calculus student can be formulated in the theory of regularity structures.

We claim that the statement a function $f : \mathbb{R}^{d+1} \to \mathbb{R}$ belongs to $\mathcal{C}_{\mathfrak{s}}^{\alpha}(\mathbb{R}^{d+1})$ for some $\alpha > 0$ is equivalent to requiring that (i) for any multi-index j with $|j|_{\mathfrak{s}} \leqslant \alpha$, $D^j f$ exists and is continuous, and (ii) for every $z \in \mathbb{R}^{d+1}$ one has the bound

$$\sup_{\substack{\bar{z} \\ \|\bar{z}-z\|_{\mathfrak{s}} \leqslant 1}} \left| f(\bar{z}) - \sum_{|k|_{\mathfrak{s}} \leqslant \alpha} \frac{1}{k!} D^k f(z)(\bar{z}-z)^k \right| \leqslant C\|\bar{z}-z\|_{\mathfrak{s}}^{\alpha} . \qquad (3.8)$$

It is not hard to check that together the conditions (i) and (ii) are equivalent to the requirements of Definition 2.4. Moreover, estimate (3.8) implies that for any multi-index j with $|j|_{\mathfrak{s}} \leqslant \alpha$ one has the bound

$$\sup_{\substack{\bar{z} \\ 0<\|\bar{z}-z\|_{\mathfrak{s}} \leqslant 1}} \left| D^j f(\bar{z}) - \sum_{|k|_{\mathfrak{s}} \leqslant \alpha-|j|_{\mathfrak{s}}} \frac{1}{k!} D^{j+k} f(z)(\bar{z}-z)^k \right| \leqslant C\|\bar{z}-z\|_{\mathfrak{s}}^{\alpha-|j|_{\mathfrak{s}}} .$$

$$(3.9)$$

Our point is that the statement $f \in \mathcal{C}_{\mathfrak{s}}^{\alpha}(\mathbb{R}^{d+1})$ corresponds to the existence of a family of polynomials indexed by $\mathbb{R}^{d+1}$ which do a sufficiently good job of describing f locally. To implement this in our setting we will define a regularity structure, denoted $\bar{\mathcal{T}}$, which we call the regularity structure of abstract polynomials.

More precisely, $\bar{\mathcal{T}} = (A, T, G)$ where $A = \mathbf{N}$ and T is the algebra of polynomials in the commuting indeterminates $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_d$. We write $\mathbf{X}$ for the associated $(d+1)$-dimensional vector indeterminant. For any $l \in \mathbf{N}$ we set T_l to be the Banach space whose basis is the set of monomials $\mathbf{X}^k$ of parabolic degree l (i.e. $|k|_{\mathfrak{s}} = l$). For a general $\tau \in T$ and monomial $\mathbf{X}^k$ we denote by $\langle \tau, \mathbf{X}^k \rangle$ the coefficient of $\mathbf{X}^k$ in the expansion of τ. We will explicitly describe the structure group G for $\bar{\mathcal{T}}$ a little later.

Given any $f \in \mathcal{C}_{\mathfrak{s}}^{\alpha}$ we can associate with it the function $F : \mathbb{R}^{d+1} \to T$ given by

$$F(z) = \sum_{|k|_{\mathfrak{s}} \leqslant \alpha} \frac{1}{k!} D^k f(z)\mathbf{X}^k . \qquad (3.10)$$

The object F should be thought of as a lift, or enhancement, of f. The original function is easily recovered since $f(z) = \langle F(z), \mathbf{1} \rangle$, where we have used the notation $\mathbf{1} := \mathbf{X}^0$. However, at each space time point F also provides additional local information about f.

Next, we make a connection between the abstract polynomials of $\bar{\mathcal{T}}$ and concrete polynomials on $\mathbb{R}^{d+1}$. We define a family of linear maps $\{\Pi_z\}_{z \in \mathbb{R}^{d+1}}$

where for any $z \in \mathbb{R}^{d+1}$ one has $\Pi_z : T \to \mathcal{S}'(\mathbb{R}^{d+1})$. The map Π_z takes an element $\tau \in T$ and returns a concrete space-time distribution which is "based at z". In this section these space-time distributions will just be polynomials so we can specify them pointwise, for any $z \in \mathbb{R}^{d+1}$ and multi-index k we set

$$\left(\Pi_z \mathbf{X}^k\right)(\bar{z}) := (\bar{z} - z)^k ,$$

where $\bar{z}$ is just a dummy variable. We then extend Π_z to all of T by linearity. The concrete Taylor polynomial for f with base point z is then given by $(\Pi_z F(z))(\,\cdot\,)$.

A key ingredient of the theory of regularity structures is a notion of smoothness for space-time distributions that are classically thought of as very singular. This requires lifting a space-time distribution to a family of local expansions at each space-time point, the notion of smoothness will then be enforced by comparing these local expansions at nearby space-time points. We make this analogy more concrete by showing how conditions (3.8) and (3.9) on f can be elegantly encoded in terms of more abstract conditions on F.

Directly comparing $F(z)$ and $F(z')$ for two close space-time points z and z' is quite unnatural since each of these local expansions are based at different space-time points. What we need is an analog of the parallel transport operation of differential geometry, we must transport a local description at one space-time point to another space-time point. For every pair $x, y \in \mathbb{R}^{d+1}$ we will define a linear map $\Gamma_{xy} : T \mapsto T$ which plays the role of parallel transport. Γ_{xy} takes something which is written as a local object at the space-time point y and "transports" it to x. This property is encoded in the algebraic relation

$$\Pi_y \tau = \Pi_x \Gamma_{xy} \tau \text{ for all } \tau \in T, \ x, y \in \mathbb{R}^{d+1} . \tag{3.11}$$

The structure group G will provide all the operators Γ_{xy}. For any $h \in \mathbb{R}^{d+1}$ we set

$$\Gamma_h \mathbf{X}^k := (\mathbf{X} - h)^k ,$$

and we extend this definition to all of T by linearity. G is defined to be the collection of operators $\{\Gamma_h\}_{h \in \mathbb{R}^{d+1}}$, one can easily check this satisfies the necessary conditions (and that G is isomorphic to $\mathbb{R}^{d+1}$ as a group). If $\Gamma_{xy} := \Gamma_{y-x}$ then (3.11) is satisfied.

With all this in place we can give the following characterization of $\mathcal{C}_\mathfrak{s}^\alpha$ spaces for $\alpha \geqslant 0$.

THEOREM 3.3. — *Let $\bar{\mathcal{T}} = (A, T, G)$ be the regularity structure of abstract polynomials (in $d+1$ components). Suppose that $\alpha > 0$. Then a function $f : \mathbb{R}^{d+1} \to \mathbb{R}$ is a member of $\mathcal{C}_\mathfrak{s}^\alpha$ (as in Definition 2.4) if and only if*

there exists a function $F : \mathbb{R}^{d+1} \to T_\alpha^-$ *such that* $\langle F(z), \mathbf{1} \rangle = f(z)$ *and for every compact set* $\mathfrak{K} \subseteq \mathbb{R}^{d+1}$ *and every* $\beta \in A$ *with* $\beta < \alpha$ *one has*

$$\sup_{x \in \mathfrak{K}} \|F(x)\|_\beta + \sup_{\substack{x,y \in \mathfrak{K} \\ x \neq y}} \frac{\|F(x) - \Gamma_{xy}F(y)\|_\beta}{\|x - y\|_{\mathfrak{s}}^{\alpha - \beta}} < \infty \,, \qquad (3.12)$$

For checking that $f \in \mathcal{C}_{\mathfrak{s}}^\alpha$ implies (3.12) one defines F as in (3.10) and check that the case of $\beta = 0$ encodes (3.8) and more generally the case of $\beta = l$ encodes (3.9) where $|j|_{\mathfrak{s}} = l$. For example, if $\beta = 1$ (and $\alpha > 1$) then one can check that

$$\mathcal{Q}_1 \Gamma_{xy} F(y) = \mathcal{Q}_1 \left(\sum_{|k|_{\mathfrak{s}} \leqslant \alpha} \frac{1}{k!} D^k f(y) (\mathbf{X} - (y - x))^k \right)$$

$$= \mathcal{Q}_1 \left(\sum_{|k|_{\mathfrak{s}} \leqslant \alpha} \sum_{j \leqslant k} \frac{1}{k!} D^k f(y) \frac{k!}{j!(k-j)!} \mathbf{X}^j (x - y)^{k-j} \right)$$

$$= \sum_{|j|_{\mathfrak{s}} = 1} \mathbf{X}^j \sum_{|k|_{\mathfrak{s}} \leqslant \alpha - 1} \frac{1}{k!} D^{j+k} f(y) (x - y)^k \,.$$

We can assume that the $\|\cdot\|_1$ norm on T_1 is an ℓ_1 type norm (with respect to the basis of monomials $\mathbf{X}^j$ with $|j|_{\mathfrak{s}} = 1$) and so we have

$$\|F(x) - \Gamma_{xy}F(y)\|_1 = \sum_{|j|_{\mathfrak{s}} = 1} \left| D^j f(x) - \sum_{|k|_{\mathfrak{s}} \leqslant \alpha - 1} \frac{1}{k!} D^{j+k} f(y) (x - y)^k \right|$$

which combined with (3.9) gives us (3.12) for all multi-indices j with $|j|_{\mathfrak{s}} = 1$. Showing that the existence of such an F implies $f \in \mathcal{C}_{\mathfrak{s}}^\alpha$ is quite similar, (3.12) implies that $\{\Pi_z F(z)\}_{z \in \mathbb{R}^{d+1}}$ is a family of sufficiently good polynomial approximations for f.

3.4. Models and Modelled Distributions

We now give a more general and axiomatic description for some of the new objects we encountered in the last section. The first concept is that of a *model* which is what allowed us to go from abstract symbols in a regularity structure to concrete space-time distributions.

DEFINITION 3.4. — *Let* $\mathcal{T} = (A, T, G)$ *be a regularity structure. A* model *for* $\mathcal{T}$ *on* $\mathbb{R}^{d+1}$ *consists of a pair* (Π, Γ) *where*

- Γ *is a map* $\Gamma \colon \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \to G$ *which we write* $(x, y) \mapsto \Gamma_{xy}$*. We require that* $\Gamma_{xx} = I$ *and* $\Gamma_{xy} \Gamma_{yz} = \Gamma_{xz}$ *for all* $x, y, z \in \mathbb{R}^{d+1}$*.*

- $\Pi = \{\Pi_x\}_{x \in \mathbb{R}^{d+1}}$ *is a family of linear maps* $\Pi_x \colon T \to \mathcal{S}'(\mathbb{R}^{d+1})$.
- *One has the algebraic relation*

$$\Pi_y = \Pi_x \Gamma_{xy} \text{ for all } x, y \in \mathbb{R}^{d+1} . \tag{3.13}$$

Finally, for any $\alpha \in A$ *and compact set* $\mathfrak{K} \subseteq \mathbb{R}^{d+1}$ *we also require that the bounds*

$$\left|(\Pi_x \tau)(\mathcal{S}_x^\lambda \eta)\right| \lesssim \|\tau\|_\alpha \lambda^\alpha \text{ and } \sup_{\beta < \alpha} \frac{\|\Gamma_{xy}\tau\|_\beta}{\|x - y\|_{\mathfrak{s}}^{\alpha - \beta}} \lesssim \|\tau\|_\alpha \tag{3.14}$$

hold uniformly over all $\tau \in T_\alpha$, $\lambda \in (0, 1]$, *space-time points* $x, y \in \mathfrak{K}$, *and test functions* $\eta \in B_r$ *for* $r := \lceil -\min A \rceil$.

Given a fixed regularity structure $\mathcal{T}$, let $\mathcal{M}$ be the set of all models on $\mathcal{T}$. For any compact set $\mathfrak{K} \subseteq \mathbb{R}^{d+1}$ one can define a "seminorm" $\|\cdot\|_{\mathfrak{K}}$ on $\mathcal{M}$ by defining $\|(\Pi, \Gamma)\|_{\mathfrak{K}}$ to be the smallest real number K such that the inequalities of (3.14) hold over $x, y \in \mathfrak{K}$ with proportionality constant K[14]. One can then define a corresponding metric on $\mathcal{M}$. While we do not explicitly give the metric here, the corresponding notion of convergence on $\mathcal{M}$ is very important and will be referenced when we introduce more of the machinery of regularity structures.

Remark 3.5. — It is straightforward to check that the (Π, Γ) introduced last section satisfies the conditions to be a model for the regularity structure of abstract polynomials $\bar{\mathcal{T}}$.

Remark 3.6. — Given $\tau \in T_\alpha$ and a model (Π, Γ) it is not necessarily the case that $\Pi_z \tau \in \mathcal{C}_{\mathfrak{s}}^\alpha$. The key point here is that the first bound of (3.14) is only enforced for test functions centered at z.

Another thing we did in the previous section was develop a notion of regularity for families of local expansions $F : \mathbb{R}^{d+1} \to T$. More generally, such families F with good regularity properties will be called *modelled distributions*.

DEFINITION 3.7. — *Fix a regularity structure* $\mathcal{T}$ *and a model* (Π, Γ) *for* $\mathcal{T}$. *Then for any* $\gamma \in \mathbb{R}$ *the space of* modelled distributions $\mathcal{D}^\gamma$ *consists of all functions* $F : \mathbb{R}^{d+1} \to T_\gamma^-$ *such that for any compact set* $\mathfrak{K} \subseteq \mathbb{R}^{d+1}$

$$\|F\|_{\gamma;\mathfrak{K}} := \sup_{x \in \mathfrak{K}} \sup_{\beta < \gamma} \|F(x)\|_\beta + \sup_{\substack{x, y \in \mathfrak{K} \\ 0 < \|x-y\|_{\mathfrak{s}} \leqslant 1}} \sup_{\beta < \gamma} \frac{\|F(x) - \Gamma_{xy}F(y)\|_\beta}{\|x - y\|_{\mathfrak{s}}^{\gamma - \beta}} < \infty . \tag{3.15}$$

---

[14] We used the word seminorm in quotation marks since $\mathcal{M}$ is *not* a linear space due to the algebraic constraint (3.13).

The definition above generalizes the idea behind Theorem 3.3. In the next section we will see a scenario where a certain class of functions with classical regularity $\mathcal{C}_{\mathfrak{s}}^{\alpha}$ for $\alpha \in (\frac{1}{3}, \frac{1}{2})$ can be thought of as more regular via the construction of lifts to modelled distributions in a $\mathcal{D}^{\gamma}$ space with $\gamma = 2\alpha$. This corroborates our earlier remark that objects with bad classical regularity can be thought of as more regular via a lift to a well behaved family of local expansions. In the next lecture we will see how this point of view actually pays off. Even if two space-time distributions f, g are too irregular to define their product fg via Theorem 2.13, we will in fact be able to make sense of their product if we can lift them to a appropriate $\mathcal{D}^{\gamma}$ spaces.

Remark 3.8. — As was the case with the $\mathcal{C}_{\mathfrak{s}}^{\alpha}$ spaces, certain theorems for $\mathcal{D}^{\gamma}$ spaces fail when $\gamma \in \mathbf{Z}$ or more general, when $\gamma \in A$ (in particular, the abstract Schauder estimate in second part of Theorem 4.20 fails). Therefore we implicitly assume that any $\mathcal{D}^{\gamma}$ space entering the assumptions or conclusion of theorem involve a value $\gamma \notin A$.

The machinery of regularity structures operates with a *fixed* regularity structure $\mathcal{T}$ and *varying* models (Π, Γ). Therefore it is very important to remember that the definition of a $\mathcal{D}^{\gamma}$ space strongly depends on the choice of model (even though their constituent objects $F : \mathbb{R}^{d+1} \to T$ don't make reference to any model). We will sometimes use the notation $\mathcal{D}^{\gamma}[(\Pi, \Gamma)]$ to make the dependence of this space on the choice of model explicit. Furthermore, we will sometimes be interested in comparing modelled distributions that live in different $\mathcal{D}^{\gamma}$ spaces coming from different models. We define $\mathcal{M} \ltimes \mathcal{D}^{\gamma}$ to be the set of triples (Π, Γ, F) such that $(\Pi, \Gamma) \in \mathcal{M}$ and $F \in \mathcal{D}^{\gamma}[(\Pi, \Gamma)]$. There is a natural way to turn $\mathcal{M} \ltimes \mathcal{D}^{\gamma}$ into a metric space, we do not describe this here but refer the reader to [28, Rem. 2.11].

3.5. Controlled rough paths

The theory of rough paths was originally developed by Lyons in [36], in this section we will see how the theory of regularity structures is related to a variant of Lyons' rough paths due to Gubinelli [22] called *controlled rough paths*. For the purposes of this section we will work with $\mathbb{R}^d$ valued functions defined on $[0, 1]$ instead of the real-valued functions defined on space-time we looked at earlier. Modifying Definitions 3.4 and 3.7 to this setting is straightforward.

Gubinelli was interested in defining the Riemann–Stieltjes type integral

$$\int_0^1 f \cdot \mathrm{d}g \,, \tag{3.16}$$

for functions f, g in the Hölder space $\mathcal{C}^\gamma([0,1], \mathbb{R}^d)$ for some $\gamma \in (\frac{1}{3}, \frac{1}{2})$. The classical theory breaks down in this regime for familiar reasons, morally dg is in $\mathcal{C}^{\gamma-1}$ and the product $f \, dg$ is not canonically defined since $2\gamma - 1 < 0$.

The strategy of controlled rough paths can be paraphrased as follows. If g is a well understood stochastic process one might be able to define the objects

$$\int_0^\bullet g_i \, \mathrm{d}g_j \tag{3.17}$$

for $1 \leqslant i, j \leqslant d$ via some probabilistic procedure (this is analogous to our construction of Wick powers of Z earlier). Then based on a completely deterministic analysis, the integral (3.16) can be constructed for a whole class of functions f which admit a type of local expansion in terms of g.

DEFINITION 3.9 (Gubinelli [22]). — *A function $f : [0,1] \to \mathbb{R}^d$ is controlled by a function $g \in \mathcal{C}^\gamma([0,1], \mathbb{R}^d)$ if there exists a function $D_g f : [0,1] \to \mathbb{R}^{d \times d}$ such that one has the bounds*

$$\left| f(t) - [f(s) + D_g f(s)(g(t) - g(s))] \right| \lesssim |t - s|^{2\gamma} \tag{3.18}$$

and

$$|D_g f(t) - D_g f(s)| \lesssim |t - s|^\gamma \tag{3.19}$$

uniformly over $s, t \in [0,1]$. Above $D_g f(s)$ is being thought as a $d \times d$ matrix acting on $\mathbb{R}^d$.

The requirements (3.18) and (3.19) should be seen as analogs of (3.8) and (3.9) above and the object $D_g f$ is analogous to a derivative. Gubinelli's observation was that although f will only be a $\mathcal{C}^\gamma$ function, the fact that f is controlled by g allows one to treat f as if it had $\mathcal{C}^{2\gamma}$ regularity.

Fix a choice of $\gamma \in (\frac{1}{3}, \frac{1}{2})$. We will now define a regularity structure $\mathcal{T}$ and an associated model (Π, Γ) built using a function $g \in \mathcal{C}^\gamma([0,1], \mathbb{R}^d)$. In this setting the statement that $f : [0,1] \to \mathbb{R}^d$ is controlled by g will be equivalent to the existence of a lift of f to modelled distribution in $\mathcal{D}^{2\gamma}$. One difference we will see here versus Section 3.3 will be in the action of the structure group G and the Γ_{xy} of the model. The interested reader can also look at [28, §3.2] where it is shown how enlarging the regularity structure given here and doing the same for the model (which is where one needs a definition for (3.17)) allows one to define the integral (3.16).

The regularity structure $\mathcal{T} = (A, T, G)$ we use has indexing set $A = \{0, \gamma\}$ where $\gamma \in (\frac{1}{3}, \frac{1}{2})$. We set $T_0 = \mathbb{R}^d$ with distinguished basis $\{\mathbf{E_i}\}_{i=1}^d$ and $T_\gamma = \mathbb{R}^{d \times d}$ with distinguished based $\{\mathbf{M_{i,j}}\}_{i,j=1}^d$. We now turn to defining the structure group G, for any $h \in \mathbb{R}^{d+1}$ we define $\Gamma_h : T \to T$ by setting

$$\Gamma_h \mathbf{E_j} = \mathbf{E_j}, \quad \Gamma_h \mathbf{M_{i,j}} = \mathbf{M_{i,j}} + h_j \mathbf{E_i} \, ,$$

and extending by linearity. We then set $G := \{\Gamma_h\}_{h \in \mathbb{R}^d}$. It is an easy exercise to check that G satisfies the necessary properties to be a structure group and is in fact ismorphic to $\mathbb{R}^d$.

If a function f is controlled by g we can lift f to vector $F: [0,1] \to T^d$ of modelled distributions by setting

$$F_i(s) = f_i(s)\mathbf{E_i} + \sum_j D_g f_{i,j}(s)\mathbf{M_{i,j}} \ . \tag{3.20}$$

We now describe a way to build a (vector-valued) model (Π, Γ) for this regularity structure for any fixed $g \in \mathcal{C}^\gamma$. For $t \in [0,1]$ we set

$$(\Pi_t \mathbf{E_i})(r) = e_i \ ,$$
$$(\Pi_t \mathbf{M_{i,j}})(r) = (g_j(r) - g_j(t))e_i$$

where $r \in [0,1]$ is a dummy variable and $\{e_i\}_{i=1}^d$ are the standard basis vectors for $\mathbb{R}^d$ (these are concrete vectors, as opposed to the abstract symbols $\{\mathbf{E_i}\}_{i=1}^d$). Finally we define the second part of the model as follows, for $s, t \in [0,1]$ we set $\Gamma_{st} = \Gamma_{g(t)-g(s)} \in G$. One can then check that (Π, Γ) satisfy the algebraic and analytic conditions to be a model. Finally one has the following theorem.

THEOREM 3.10. — *Let $\mathcal{T}$ be the regularity structure defined above and let (Π, Γ) be a model built out a fixed $g \in \mathcal{C}^\gamma([0,1], \mathbb{R}^d)$. Then a function $f : [0,1] \to \mathbb{R}^d$ is controlled by g if and only if there exists a modelled distribution $F \in \mathcal{D}^{2\gamma}$ with $\mathcal{Q}_0 F(t) = \sum_{i=1}^d f_i(t)\mathbf{E_i}$.*

3.6. Regularity Structures for SPDEs

We take a moment to discuss the vector space T that appears in regularity structures $\mathcal{T}$ used for solving equations like (1.7). The space T will be formed by the linear span of abstract symbols. We denote by $\mathfrak{T}$ the set of all abstract symbols appearing in T. $\mathfrak{T}$ contains the symbol Ξ which represents the driving noise ξ, since ξ is taken to be space-time white noise we set $|\Xi| = -d/2 - 1 - \kappa$ where $\kappa > 0$ can be taken arbitrary small. $\mathfrak{T}$ will also have the symbol $\mathcal{I}[\Xi]$ which represents the solution to the underlying linear equation. More generally, given a symbol $\tau \in \mathfrak{T}$ it will *sometimes* be the case that $\mathfrak{T}$ also contains the abstract symbol $\mathcal{I}[\tau]$ which represents the "integral" of τ, that is "$K * \tau$". Inspired by the Schauder estimate (Theorem 2.8) we would then require $|\mathcal{I}[\tau]| = |\tau| + 2$. However, we do not allow any symbol of the form $\mathcal{I}[\mathbf{X^k}]$[15].

---

[15] Said differently, we assume that $\mathcal{I}[\,\cdot\,]$ applied to any abstract polynomial vanishes, Section 4.4 will clarify this.

Given symbols $\tau_1, \tau_2 \in \mathfrak{T}$ it will *sometimes* be the case that $\mathfrak{T}$ will contain the abstract symbol $\tau_1 \tau_2$, which represents a commutative product of τ_1 and τ_2. In this case we will require $|\tau_1 \tau_2| = |\tau_1| + |\tau_2|$. This condition on products is an important way that the concept of order differs from that of regularity[16].

The symbols mentioned above are generated recursively: one starts with a set of primitive symbols which consists of Ξ and various powers of $\mathbf{X}$ and then builds more complicated symbols by using $\mathcal{I}[\,\cdot\,]$ and our abstract product. The graphical notation we used for Φ^4 should be seen as a shorthand for the symbols we have described. For example, we have

$$\Psi = \mathcal{I}\big[\mathcal{I}[\Xi]^3\big]\mathcal{I}[\Xi] \,.$$

However, the regularity structures one encounters in practice do not contain *all* the symbols generated by the recursive procedure sketched above, doing so would usually create problems for the first and second conditions of Definition 3.2.

To construct the right list of symbols $\mathfrak{T}$ we start by iteratively applying a particular set of rules $\mathfrak{R}_F$ determined by the structure of the non-linearity F appearing in (1.7). The list of rules for Φ^4 equations is given by

$$\mathfrak{R}_{\Phi^4} := \{\mathbf{X}^k \mathcal{I}[\,\cdot\,], \ \mathbf{X}^k \mathcal{I}[\,\cdot\,]\mathcal{I}[\,\cdot\,], \ \mathcal{I}[\,\cdot\,]\mathcal{I}[\,\cdot\,]\mathcal{I}[\,\cdot\,]\} \,. \tag{3.21}$$

Above and in what follows k represents an arbitrary multi-index, sometimes subject to a stated constraint. We set $\mathfrak{T}_0 := \{\Xi, \ \mathbf{X}^k\}$ to be the set of primitive symbols. Then for $j \geqslant 1$, the set $\mathfrak{T}_j$ is formed by taking the union of $\mathfrak{T}_0$ with the set of all the symbols that one gets by applying any of the operations listed in the given rule $\mathfrak{R}_F$ to any of the elements of $\mathfrak{T}_{j-1}$. For example, in the case of Φ^4 it is the case that

$$\tau_1, \tau_2, \tau_3 \in \mathfrak{T}_j \Rightarrow \mathcal{I}[\tau_1]\mathcal{I}[\tau_2]\mathcal{I}[\tau_3] \in \mathfrak{T}_{j+1} \,.$$

An important consequence of subcriticality of the equation (1.7) is the following: if one defines the sets of symbols $\mathfrak{T}_j$ using $\mathfrak{R}_F$ then there will exist some $\beta > 0$ such that for all $j \geqslant 1$ one has

$$\min_{\tau \in \mathfrak{T}_j \setminus \mathfrak{T}_{j-1}} |\tau| > \min_{\tau \in \mathfrak{T}_{j-1}} |\tau| + \beta \,.$$

This means that as we iterate the application of the rule the new symbols we generate are increasing in order. This guarantees that if we set $\mathfrak{T}$ equal to $\cup_{j=0}^{\infty} \mathfrak{T}_j$ then the corresponding list of orders A will be bounded below and will not contain any accumulation points.

---

[16] In particular, this will allow us to bypass circular issues like the product $v\Psi$ in Section 3.1.

However, $\mathfrak{T}$ would still include an infinite list of symbols. In practice one wants $\mathfrak{T}$ to be a finite set. To do this we fix a constant γ which is the upper limit on what order symbols we include in our regularity structure[17]. We can then modify our previous construction. For $j \geqslant 1$ we define the sets $\bar{\mathfrak{T}}_j$ by taking the union of $\mathfrak{T}_0$ with the set of all the symbols that one gets by applying any of the operations listed in the given rule $\mathfrak{R}_F$ to any of the elements of $\bar{\mathfrak{T}}_{j-1}$, but now with the convention that $\mathcal{I}[\tau] > \gamma$ then $\mathcal{I}[\tau]$ is considered to vanish[18]. We then set

$$\mathfrak{T} := \left\{ \tau \in \mathfrak{T}_0 \cup \left(\bigcup_{j=1}^{\infty} \bar{\mathfrak{T}}_j \right) \; : \; |\tau| \leqslant \gamma \right\}.$$

3.7. The regularity structure and model for Φ_2^4

When defining the regularity structure $\mathcal{T}$ for the Φ_2^4 equation the list of symbols is given by $\mathfrak{T}$ defined as above with the rule $\mathfrak{R}_{\Phi^4}$ and γ taken positive but sufficiently small ($\gamma > 2\kappa$ suffices).

Table 3.1. Symbols for Φ_2^4

Symbol	Order
Ξ	$-2 - \kappa$
$\mathord{\uparrow}$	$-\kappa$
V	-2κ
$\mathsf{\Psi}$	-3κ
$\mathbf{1}$	0

Any realization of the driving noise ξ can then be lifted to a model $(\hat{\Pi}^\delta, \hat{\Gamma}^\delta)$ as follows. For any $z \in \mathbb{R}^{2+1}$ we set:

$$\left(\hat{\Pi}_z^\delta \Xi \right)(\bar{z}) = \xi_\delta(\bar{z})$$
$$\left(\hat{\Pi}_z^\delta \mathord{\uparrow} \right)(\bar{z}) = \mathord{\uparrow}_\delta(\bar{z})$$

$$\left(\hat{\Pi}_z^\delta \mathsf{V} \right)(\bar{z}) = \mathsf{V}_\delta(\bar{z})$$
$$\left(\hat{\Pi}_z^\delta \mathsf{\Psi} \right)(\bar{z}) = \mathsf{\Psi}_\delta(\bar{z}) \tag{3.22}$$
$$\left(\hat{\Pi}_z^\delta \mathbf{1} \right)(\bar{z}) = 1 \, .$$

Here $\mathord{\uparrow}_\delta, \mathsf{V}_\delta, \mathsf{\Psi}_\delta$ are the approximate Wick powers introduced in Lecture 2. A key simplification with Φ_2^4 is that the maps Π_z do not depend on z, this means

---

[17] γ will need to be sufficiently large to allow one to pose the abstract fixed point problem, see Section 4.3.2.

[18] In particular, no symbol can contain $\mathcal{I}[\tau]$ as a subsymbol.

we can set $\Gamma_{xy} = \text{Id}$ where Id is the identity on T. It can be checked that the models $(\Pi^\delta, \Gamma^\delta)$ satisfy the conditions of Definition 3.4. Additionally, one can remove the regularization and show that the models $(\Pi^\delta, \Gamma^\delta)$, viewed as random elements of $\mathcal{M}$, converge in probability as $\delta \downarrow 0$ to a limiting random model (Π, Γ) given by[19]

$$\left(\hat{\Pi}_z \Xi\right)(\bar{z}) = \xi(\bar{z})$$

$$\left(\hat{\Pi}_z \uparrow\right)(\bar{z}) = \uparrow(\bar{z})$$

$$\left(\hat{\Pi}_z \mathsf{V}\right)(\bar{z}) = \mathsf{V}(\bar{z})$$

$$\left(\hat{\Pi}_z \mathsf{\Psi}\right)(\bar{z}) = \mathsf{\Psi}(\bar{z})$$

$$\left(\hat{\Pi}_z \mathbb{1}\right)(\bar{z}) = 1$$

and with $\Gamma_{xy} = \text{Id}$ for all $x, y \in \mathbb{R}^{d+1}$.

Although the full regularity structure $\mathcal{T}$ is required to formulate the fixed point problem, the solution Φ will be of the form $\Phi(z) = \Phi_{\mathbf{1}}\mathbf{1} + \uparrow$ which is similar to the decomposition seen in the Da Prato–Debussche argument. The fact that the structure group can be chosen to be trivial is why the Da Prato–Debussche argument works for Φ_2^4.

3.8. The regularity structure and model for Φ_3^4

For Φ_3^4, it suffices to define $\mathfrak{T}$ by taking γ slightly greater than 1. We include a table of these symbols below[20]. Whenever a factor $\mathbf{X}^k$ appears in a symbol the multi-index k can vary but is constrained by the condition that the symbol's order be less than γ.

Again, the approach is to define a family of random models $(\hat{\Pi}^\delta, \hat{\Gamma}^\delta)$, defined via lifting ξ, which converge in probability to a limiting random model $(\hat{\Pi}, \hat{\Gamma})$ as $\delta \downarrow 0$. For $\tau = \Xi, \uparrow, \mathsf{V},$ and $\mathsf{\Psi}$, we define $\hat{\Pi}_z^\delta \tau$ in the same way as we did for (3.22) (where the objects are replaced by their $d = 3$ counterparts). We will not explicitly describe all of the model $(\hat{\Pi}, \hat{\Gamma})$, the goal for our present discussion is to show how the Γ_{xy}^δ's and structure group G are forced to be non-trivial.

This is easily seen with the symbol Y. A naive definition one might make is

$$\left(\hat{\Pi}_z^{\delta}\mathsf{Y}\right)(\bar{z}) = \mathsf{Y}_\delta(\bar{z}) = \int_{\mathbb{R}^{d+1}} \mathrm{d}u \, K(\bar{z} - u)\mathsf{\Psi}_\delta(u) \tag{3.23}$$

---

[19] We continue to abuse notation here, using point-wise equalities for singular distributions.

[20] The algorithm for construction $\mathfrak{T}$ that we have given may produce extraneous symbols and indeed the last three symbols given in Table 3.2 are unnecessary to set up a fixed point problem for Φ_3^4. See [27, §8.1] for an algorithm that will give a minimal list of symbols. Also see 4.5 of these lecture notes for a discussion of relevant terms.

Table 3.2. Symbols for Φ^4_3

Symbol	Order
Ξ	$-5/2 - \kappa$
$\mathbf{X}^k$	$\|k\|_{\mathfrak{s}}$
$\mathord{\uparrow}\mathbf{X}^k$	$\|k\|_{\mathfrak{s}} - 1/2 - \kappa$
$\vee\mathbf{X}^k$	$\|k\|_{\mathfrak{s}} - 1 - 2\kappa$
ψ	$-3/2 - 3\kappa$
$\curlyvee$	$1/2 - 3\kappa$
$\searrow$	$1/2 - 3\kappa$
$\curlyvee$	$1 - 2\kappa$
ψ	-4κ
ψ	$-1/2 - 5\kappa$
ψ	-4κ
$\curlyvee\ \curlyvee$	$1 - 6\kappa$
$\mathbf{X}^k\psi$	$-4\kappa + \|k\|_{\mathfrak{s}}$
ψ	$1/2 - 3\kappa$

for any $z \in \mathbb{R}^{3+1}$. However this definition will *not* satisfy the first bound of (3.14). While the objects on the right hand side of (3.23) is of regularity $\frac{1}{2} - 3\kappa$, it does not satisfy the bound

$$\left|\left(\hat{\Pi}_z \curlyvee\right)(\mathcal{S}_z^\lambda \eta)\right| \lesssim \lambda^{\frac{1}{2} - 3\kappa} \tag{3.24}$$

uniformly in $\lambda \in (0, 1]$ for an arbitrary test function η. A way to reconcile this difference was already seen in Definition 2.4. If we want to see a space-time function of regularity $\gamma > 0$ vanish at order γ then we should subtract a suitable Taylor polynomial. We will get the bound (3.24) if (3.23) is changed to

$$\left(\hat{\Pi}_z^\delta \curlyvee\right)(\bar{z}) = \int \mathrm{d}u\ K(\bar{z} - u)\vee_\delta(u)\ -\ \int \mathrm{d}u\ K(z - u)\vee_\delta(u)\,.$$

However $\left(\Pi_z^\delta \curlyvee\right)(\bar{z})$ now has a dependence on z which means that the structure group G cannot be chosen to act trivially on $\curlyvee$. The compatibility condition (3.13) determines completely how Γ acts on $\curlyvee$. Indeed, $\hat{\Pi}_x^\delta \hat{\Gamma}_{xy}^\delta = \hat{\Pi}_y^\delta$ gives us that

$$\left[\hat{\Pi}_x^\delta\left(\mathrm{Id} - \hat{\Gamma}_{xy}^\delta\right)\curlyvee\right](\bar{z}) = \left(\hat{\Pi}_x^\delta \curlyvee\right)(\bar{z}) - \left(\hat{\Pi}_y^\delta \curlyvee\right)(\bar{z})$$

$$= \int \mathrm{d}u\ K(y - u)\vee_\delta(u) - \int \mathrm{d}u\ K(x - u)\vee_\delta(u)\,.$$

Therefore, we set

$$\hat{\Gamma}^{\delta}_{xy}\mathbf{\Psi} \;=\; \mathbf{\Psi} + \left(\int \mathrm{d}u \; K(y-u)\mathbf{V}_{\delta}(u) - \int \mathrm{d}u \; K(x-u)\mathbf{V}_{\delta}(u) \right) \mathbf{1} \,.$$

The group action on all the other symbols is determined by similar considerations for integration and the compatibility condition for products, given in (4.11) below.

Remark 3.11. — In general, terms involving $\mathbf{X}$ appear in a model when an abstract integration leads to a symbol of positive order. It is worth mentioning that these extra terms do not occur in Gubinelli's approach [23] to singular SPDEs using "paracontrolled distributions".

Lecture 4

4.1. Construction of Canonical Models

In the last lecture we discussed regularity structures and models associated with controlled rough paths, Φ^4_2, and Φ^4_3. In this section we will show that for any fixed regularity structure $\mathcal{T}$ which is created by a set of formal rules like (3.21), there exists a canonical way to map each fixed realization of a smoothed noise ξ_δ to a corresponding model $(\Pi^\delta, \Gamma^\delta)$. This model is called a *canonical model* and it will be defined recursively with respect to $\mathcal{I}[\,\cdot\,]$ and the abstract product on T. After that we will discuss more systematically how to perform the renormalisation procedure which leads to the *renormalized* models, examples of which we have already encountered in the previous lecture.

In order to motivate concepts we ignore Definition 3.4 for a moment. There is a naive approach to assigning a concrete space-time function (built out of ξ_δ) to each of the abstract symbols appearing in our regularity structure, one can recursively define a *single* linear map $\mathbf{\Pi}^\delta : T \to \mathcal{S}'(\mathbb{R}^{d+1})$ by setting:

$$\big(\mathbf{\Pi}^\delta \Xi\big)(\bar{z}) := \xi_\delta(\bar{z})$$

$$\big(\mathbf{\Pi}^\delta \mathcal{I}[\tau]\big)(\bar{z}) := \int_{\mathbb{R}^{d+1}} \mathrm{d}y \; K(\bar{z} - y)\big(\mathbf{\Pi}^\delta \tau\big)(y) \qquad (4.1)$$

$$\big(\mathbf{\Pi}^\delta \tau_1 \tau_2\big)(\bar{z}) := \big(\mathbf{\Pi}^\delta \tau_1\big)(\bar{z}) \times \big(\mathbf{\Pi}^\delta \tau_1\big)(\bar{z}) \,.$$

The map $\mathbf{\Pi}^\delta$ is a much simpler object than a model but it encodes less structure. In particular, it does not directly encode anything about the order of objects. The additional structure that models encode is what makes the machinery of Sections 4.3.1 and 4.4 continuous with respect to models.

Ajay Chandra and Hendrik Weber

We have already seen above that when a regularity structure includes symbols of positive order, the maps $\{\Pi_z\}$ in a model must be allowed to be z-dependent, if we want the first bound of (3.14) to hold. Keeping this in mind, we now describe how the maps Π_z^δ of the canonical model $(\Pi^\delta, \Gamma^\delta)$ are defined. For any $z \in \mathbb{R}^{d+1}$ one sets:

$$
\begin{aligned}
\left(\Pi_z^\delta \mathbf{X}^k \tau\right)(\bar{z}) &:= (\bar{z} - z)^k \times \left(\Pi_z^\delta \tau\right)(\bar{z}) \\
\left(\Pi_z^\delta \Xi\right)(\bar{z}) &:= \xi_\delta(\bar{z}) \\
\left(\Pi_z^\delta \mathcal{I}[\tau]\right)(\bar{z}) &:= \int_{\mathbb{R}^{d+1}} \mathrm{d}y \ K(\bar{z} - y)\left(\Pi_z^\delta \tau\right)(y) \\
&\qquad - \sum_{\substack{k \\ |k|_\mathfrak{s} < |\tau|+2}} \frac{(\bar{z} - z)^k}{k!} \int_{\mathbb{R}^{d+1}} \mathrm{d}y \ (D^k K)(z - y)\left(\Pi_z^\delta \tau\right)(y) \\
\left(\Pi_z^\delta \tau_1 \tau_2\right)(\bar{z}) &:= \left(\Pi_z^\delta \tau_1\right)(\bar{z}) \times \left(\Pi_z^\delta \tau_1\right)(\bar{z}) \,.
\end{aligned}
\tag{4.2}
$$

The key point here is that the application of $\mathcal{I}[\,\cdot\,]$ can produce a new object of positive order from an old one of negative order. This is why in the third line of (4.2) the subtraction we have implemented is just the subtraction of a partial Taylor expansion when $|\mathcal{I}[\tau]| \geqslant 0$. Of course, multiplication can also produce new objects of positive order but this is dealt with automatically when we enforce the product property, see Section 4.3.2 below, in particular (4.11).

In Section 3.8 we described models[21] $(\hat{\Pi}^\delta, \hat{\Gamma}^\delta)$ where for $\tau \in \{\Xi, \uparrow, \vee, \curlyvee\}$ the function $\hat{\Pi}_z^\delta \tau$ had no z-dependence. This was possible because these abstract symbols are of negative order *and* the latter three objects are built (using $\mathcal{I}$ and the abstract product) out of objects which are all of negative order. However the symbol $\curlyvee = \mathcal{I}[\mathcal{I}[\Xi]^3]$ is of positive order so $\Pi_z^\delta \curlyvee$ had to be z-dependent in order for the first bound of (3.14) to hold. We also remark that $\hat{\Pi}_z^\delta \curlyvee$ will also be z-dependent even though $\curlyvee$ is of negative order. This is because $\curlyvee$ is a product of $\vee$ and $\curlyvee$ and the latter is of positive order.

The recursive definition (4.2) is convenient to state and useful in many proofs. One can also recursively define the corresponding operators Γ_{xy}^δ as we have already sketched above. For models like the canonical model for Φ_3^4 one could in principle check the algebraic properties by hand, but this can easily become unwieldy with more complicated examples. In [27] the connection between a recursive definition of the Γ_{xy}^δ's and their algebraic properties is made clear in an elegant way by using the language of Hopf algebras (see [28, §5.3]). This formulation also becomes very useful when

---

[21] We remark again that these models were *not* canonical models, see the next section.

one wants to go beyond the canonical model and construct the renormalized models of the next section.

4.2. Convergence of Models and Renormalization

Most of the machinery we discuss in this lecture is completely deterministic, applied separately for each fixed realization of the noise ξ. However there is one major obstacle that this deterministic analysis cannot overcome: in general the canonical models $(\Pi^\delta, \Gamma^\delta)$ built from a δ-smoothing of a fixed realization of ξ will *not* converge in the limit $\delta \downarrow 0$. The canonical models associated to Φ_2^4 and Φ_3^4 are examples of this: we have already seen in Lecture 2 that the random space-time distributions $\uparrow_\delta^3$ (which play the role of $\Pi_z^\delta \mathcal{I}[\Xi]^3$ in the canonical model based on ξ_δ) do not converge as $\delta \to 0$. The Da Prato–Debussche argument for Φ_2^4 overcame this by implementing renormalization subtractions and the approach in regularity structures is similar.

We first discuss a criterion for the stochastic convergence of models. Let $\mathcal{T}$ be a regularity structure defined as in Section 3.6 and let $\mathfrak{T}_-$ be the set of abstract symbols *of negative order* appearing in $\mathcal{T}$. We seek conditions that ensure that a sequence $(\Pi^\delta, \Gamma^\delta)$ of random models converge in probability to a random limiting model (Π, Γ). The key stochastic estimates to show this are the following: for every $\tau \in \mathfrak{T}_-$, test function η, there should exist $\nu, \nu' > 0$ such that the bounds

$$\mathbb{E}\Big[\big|(\Pi_z^\delta \tau)(\mathcal{S}_z^\lambda \eta)\big|^p\Big] \lesssim \lambda^{p(|\tau|+\nu)} \tag{4.3}$$

and

$$\mathbb{E}\Big[\big|(\Pi_z^\delta \tau - \Pi_z \tau)(\mathcal{S}_z^\lambda \eta)\big|^p\Big] \lesssim \delta^{\nu'} \lambda^{p(|\tau|+\nu)} \tag{4.4}$$

hold for every $p \in \mathbf{N}$, uniformly over $\delta, \lambda \in (0, 1]$, uniformly over a suitable class of test-functions η and locally uniformly in $z^{(22)}$.

Note that the conditions above only involve a finite number of symbols τ. This is similar to what we saw in the Da Prato–Debussche argument and as before is a consequence of subcriticality. We also remark that under some natural assumptions on the sequence of models $(\Pi^\delta, \Gamma^\delta)$ one does not have to perform any stochastic estimates on the $\hat\Gamma^\delta$'s. Finally, as we have already seen in Lecture 2, one can win a major simplification when the driving noise ξ is Gaussian. In the Gaussian case one can apply Nelson's Estimate (A.6) and it suffices to establish the stochastic estimates just for $p = 2$. The reader

---

[22] For a precise statement of all required conditions in the Gaussian case see [27, Thm. 10.7].

has already seen the derivation of the bound (4.3) for $\tau = \Psi$ in Φ_2^4 and Φ_3^4. The simple graphical approach based on convolutions presented there is sufficient for Φ_3^4.

It is important to observe that the bound (4.3) consists of *two statements*: One is showing that the given quantity satisfies the right type of upper bound in λ which determines the *order* of the limiting object. The other statement is that the quantity remains finite as $\delta \downarrow 0$ and relates to the so called ultraviolet divergences and renormalization. These two issues are essentially orthogonal to each other. In particular, if the canonical models have divergent behaviour as $\delta \downarrow 0$ then one will already see this at the level of the simpler map $\mathbf{\Pi}^\delta$ of (4.1); it is conceptually simpler to first try to renormalize this map. One does this by defining a new map $\hat{\mathbf{\Pi}}^\delta := \mathbf{\Pi}^\delta M^\delta$ where $M^\delta : T \mapsto T$ is a linear map which performs renormalization subtractions at the level of the regularity structure $\mathcal{T}$. For example, for Φ_3^4 one would have

$$M^\delta \Psi = \Psi - C_\delta$$
$$M^\delta \mathbf{X_j}\Psi = \Psi\mathbf{X_j} - C_\delta \mathbf{X_j} \ \text{ for } j = 1,2,3$$
$$M^\delta \Psihook = \Psihook - C_\delta Y - \tilde{C}_\delta \mathbf{1}$$
$$M^\delta \Psibig = \Psibig - 3C_\delta \Psismall - C_\delta Y + 3C_\delta^2 \setminus - 3\tilde{C}_\delta \uparrow$$

$$M^\delta \Psiv = \Psiv - 3C_\delta \uparrow$$
$$M^\delta \Psidown = \Psidown - C_\delta \searrow$$
$$M^\delta \Psiw = \Psiw - 3C_\delta \searrow$$

where $C_\delta \sim \frac{1}{\delta}$ and $\tilde{C}_\delta \sim -\log(\delta)$[23]. M^δ is given by the identity on all remaining abstract symbols for this regularity structure[24].

The map M^δ has been defined so that if one views the objects $\hat{\mathbf{\Pi}}^\delta \tau$ as *random* space-time distributions then they converge in probability in the $\delta \downarrow 0$ limit. The canonical model $(\Pi^\delta, \Gamma^\delta)$ is an enhancement of the simpler map $\mathbf{\Pi}^\delta$. It is possible to postulate suitably flexible conditions on M^δ (for details see [27, §8.3]) which would guarantee that one can build another model $(\hat{\Pi}^\delta, \hat{\Gamma}^\delta)$, called a renormalized model, which is an analogous enhancement of the map $\hat{\mathbf{\Pi}}^\delta$. In the case of the Φ_3^4 and many other examples the canonical and renormalized models will satisfy the "diagonal identity"

$$(\hat{\Pi}_z^\delta \tau)(z) = (\Pi_z^\delta M^\delta \tau)(z). \tag{4.5}$$

---

[23] The renormalization procedure given for $\Psibig$ is a bit inconsistent, one should also include terms $-C_\delta c_K \Psi + C_\delta^2 c_K \mathbf{1}$ on the RHS, here c_K is a δ-independent finite constant which is formally given by $c_K = \int dz\ K(z)$. We will later make the convention that $\mathcal{I}$ should encode integration with just a piece of the heat kernel, defined so that it annihilates constants which means c_K will vanish. See Section 4.4.

[24] Note that in order to allow for renormalization of the canonical model this regularity structure has more symbols than those listed in Table 3.2, in particular it suffices to take $\gamma = 3/2$ when defining $\mathfrak{T}$. Taking γ larger does not create any new technical difficulties since $\mathfrak{T}_-$ remains the same.

Defining this process for general regularity structures is non-trivial (see [28, §7.1] for a summary). We remark that the relationship between the renormalized model and canonical model is fairly complex, in general one *does not* have the equality $\hat{\Pi}_z^\delta \tau = \Pi_z^\delta M^\delta \tau$.

The renormalization procedure can be seen as a deformation of the multiplicative structure of the canonical model. The first three relations of (4.2) are essential properties that in practice we always require from models. The maps $\hat{\Pi}_z^\delta$ will satisfy the first three relations of (4.2) but they will not satisfy the last relation. In the limit $\delta = 0$ the $\hat{\Pi}_z \tau$ will be distributions and it is a priori not even clear how this condition could be interpreted. One should then really view the left hand side of this expression as a *definition* of the right hand side.

4.3. The deterministic analysis

It is an amazing fact about the theory of regularity structures, that once the stochastic calculations that lead to the renormalised model have been performed, the deterministic argument that shows the short time existence of solutions, follows "automatically". We will now proceed to show how.

Instead of solving the desired equation in a space of $\mathbb{R}$ valued function, we will now solve the equation in a space of *modelled distributions*, i.e. functions taking values in T. We stress again that this space depends on the specific realisation of the model (which is in turn constructed as random variables on a suitable probability space). In the case of Φ_2^4 this "lifted" solution will take values in the linear space spanned by $\mathbf{1}$ and $\uparrow$. Furthermore, the coefficient for $\uparrow$ will be one, so that we will have

$$\Phi(z) = \uparrow + \Phi_{\mathbf{1}}(z)\mathbf{1} .$$

Here, the function $\Phi_{\mathbf{1}}$ which describes the solution Φ at order 0 corresponds exactly to the remainder v we already saw in the Da Prato–Debussche argument in Section 3.1. In the case of Φ_3^4 we need more terms that describe the solution Φ and again several of these will be constant. In the end, we have

$$\Phi(z) = \uparrow + \Phi_{\mathbf{1}}(z)\mathbf{1} - \mathsf{Y} + \Phi_{\mathsf{Y}}(z)\mathsf{Y} + \langle \Phi_{\mathbf{X}}(z), \mathbf{X} \rangle , \qquad (4.6)$$

where we have used the notation $\langle \Phi_{\mathbf{X}}, \mathbf{X} \rangle = \sum_{j=1}^{3} \Phi_{\mathbf{X}_j} \mathbf{X}_j$. Furthermore, we will see that the structure of the equation dictates that $\Phi_{\mathsf{Y}} = -3\Phi_{\mathbf{1}}$, so that solving for Φ really involves solving for a system of two functions (one real valued and one vector-valued) $\Phi_{\mathbf{1}}$ and $\Phi_{\mathbf{X}}$. We will justify this particular form of the expansion in Section 4.5 below. It is important to note that in both cases the description of the solution Φ requires much fewer symbols

than contained in the regularity structure. The remaining symbols will be used to define the non-linear operations in the fixed point map. The three operations we need are reconstruction, multiplication and integration. We will now proceed to explain each of these operations.

4.3.1. The Reconstruction Theorem

The fundamental link between the abstract definitions/machinery we have introduced and the concrete results described by Metatheorem 1.4 is the Reconstruction Theorem which states that there is a reconstruction operator $\mathcal{R}$ which establishes a correspondence between modelled distributions of strictly positive regularity and actual space-time distributions.

THEOREM 4.1 ([27, Thm. 3.10]). — *Let $\mathcal{T} = (A, T, G)$ be a regularity structure, let $\alpha := \min A$, and $r := \lceil -\alpha \rceil$. Let (Π, Γ) be a model for $\mathcal{T}$. For any $\gamma > 0$ we set $\mathcal{D}^\gamma$ to be the corresponding space of modelled distributions based on the model (Π, Γ).*

Then there exists a continuous linear map $\mathcal{R} : \mathcal{D}^\gamma \to \mathcal{C}^\alpha_{\mathfrak{s}}$ with the property that for any compact set $\mathfrak{K} \subseteq \mathbb{R}^{d+1}$ and any $F \in \mathcal{D}^\gamma$ one has that $\mathcal{R}F$ is the unique space-time distribution satisfying

$$\left| (\mathcal{R}F)(\mathcal{S}^\lambda_z \eta) - (\Pi_z F(z))(\mathcal{S}^\lambda_z \eta) \right| \lesssim \lambda^\gamma \tag{4.7}$$

uniformly over test functions $\eta \in B_r$, $\lambda \in (0, 1]$ and uniformly over compacts in z.

Furthermore, if the given model (Π, Γ) takes values in continuous functions (this means that $(\Pi_z \tau)$ is always a continuous function) then $\mathcal{R}F$ will also be a continuous function and one has the identity

$$(\mathcal{R}F)(z) = (\Pi_z F(z))(z) . \tag{4.8}$$

For a given modelled distribution F the space-time distribution $\mathcal{R}F$ is constructed as the limit $\mathcal{R}F := \lim_{n\to\infty} \mathcal{R}^n F$ where $\mathcal{R}^n F$ is built by "stitching together" the distributions $\{(\Pi_z F(z))\}_{z \in \Lambda_n}$ where $\Lambda_n \subseteq \mathbb{R}^{d+1}$ is a discrete set of grid points with resolution 2^{-n}. More precisely, $\mathcal{R}^n F$ is defined as

$$(\mathcal{R}^n F)(z) = \sum_{x \in \Lambda_n} (\Pi_x F(x))(\psi^n_x) \psi^n_x(z) \tag{4.9}$$

where the functions $\{\psi^n_x\}_{x \in \Lambda_n}$ are appropriately scaled and translated copies of a fairly regular function[25] [26] ψ. Establishing the convergence

---

[25] In particular, the proof of the Reconstruction Theorem in [27, 28] uses wavelet analysis.

[26] The reader is encouraged to compare (4.9) with the proof of Theorem 2.7.

$\lim_{n\to\infty} \mathcal{R}^n F$ uses the algebraic and analytic conditions imposed by Definitions 3.4 and 3.7 in a very direct manner. In fact, the Reconstruction Theorem served as the initial motivation for the abstract setting of regularity structures. A nice exposition of the proof can be found in [28].

One can also let the models in Theorem 4.1 vary, then the theorem gives the existence of a reconstruction map $\tilde{\mathcal{R}}$ acting on triples $(\Pi, \Gamma, F) \in \mathcal{M} \ltimes \mathcal{D}^\gamma$. The following theorem gives another essential property of the reconstruction operation.

THEOREM 4.2. — *Let $\mathcal{T}$ and α be as in Theorem 4.1. Then for any $\gamma > 0$ the reconstruction map $\tilde{\mathcal{R}} : \mathcal{M} \ltimes \mathcal{D}^\gamma \to \mathcal{C}_{\mathfrak{s}}^\alpha$ is continuous.*

4.3.2. Multiplication

We aim to *lift* the non-linear fixed point problem (2.14) to the level of modelled distributions. This will take the form

$$\Phi = -\mathcal{K}_\gamma \Phi^3 + \mathcal{K}_\gamma \Xi \,, \tag{4.10}$$

where $\mathcal{K}_\gamma$ is a linear operator acting on a space of modelled distribution corresponding to the convolution with the heat kernel K. We will discuss the definition of $\mathcal{K}_\gamma$ in the next section. For now we start with the definition of the operation $U \mapsto U^3$, i.e. we have to define multiplication of certain modelled distributions.

The product of modelled distributions is defined pointwise on the level of the regularity structure.

DEFINITION 4.3 ([27, Def. 4.1]). — *A mapping $T \times T \ni (a, b) \mapsto ab \in T$ is called a* product *if it is bilinear and*

- *It respects the orders in the sense that for $a \in T_\alpha$ and $b \in T_\beta$ we have $ab \in T_{\alpha+\beta}$,*
- *we have $\mathbf{1}\, a = a\, \mathbf{1} = a$ for all $a \in T$.*

We have already seen the actual construction of this product as a part of the construction of the regularity structure. In the case of the regularity structure for Φ_3^4 we have, for example

$$\mathbf{X}^k\, \mathbf{X}^\ell = \mathbf{X}^{k+\ell}, \quad \mathord{\uparrow}\,\mathord{\uparrow} = \mathrm{V}, \quad \mathrm{Y}\,\mathord{\uparrow} = \mathrm{V}, \dots$$

and the product is extended in a bilinear way. It is important to observe, that many products that could be built from the entries in Table 3.2 do not have a natural definition. For example, we have not included a symbol for $\mathrm{V}\,\mathrm{V}$ or for any product involving Ξ in the regularity structure. This is

because the regularity structure is tailor-built to include only those symbols that we actually need in the construction of the fixed point map. In the same way, we will set $\mathbf{X}^k \mathbf{X}^\ell = 0$ as soon as the order $|k + \ell|_{\mathfrak{s}} \geqslant 2$. In order to satisfy the assumption of Definition 4.3 we can always define such products to be zero.

Our aim for this section is to prove a "multiplicative" inequality in the spirit of Theorem 2.13 for modelled distributions. To this end we need to make sure that the product is compatible with spatial translations, represented by the group G. More precisely, we need to assume that the product is γ-*regular* which means that the identity

$$\Gamma(a\,b) = \Gamma a\,\Gamma b\,. \tag{4.11}$$

holds for all $\Gamma \in G$ and all "relevant" $a, b \in T$ of order $\leqslant \gamma$, where γ is the order up to which we aim to describe the product[27]. In the case of $\bar{\mathcal{T}}$, the regularity structure of polynomials, this condition reduces to the trivial identity $(x - h)^k\,(x - h)^\ell = (x - h)^{k+\ell}$. In the recursive definition of the canonical model (4.11) in conjunction with the action of the integration map $\mathcal{I}$ completely determines the action of the Γ_{xy}. However, it is non-trivial to construct the renormalised models $(\hat{\Pi}^\delta, \hat{\Gamma}^\delta)$ in such a way that (4.11) remains true.

In order to state the main result of this section, we need to introduce one more notion. Let $F \in \mathcal{D}^\gamma$ be a modelled distribution as defined in (3.7). We will say that $F \in \mathcal{D}^\gamma_\alpha$ if F takes values in a subspace of T which is spanned by symbols of order $\geqslant \alpha$. Note that a non-trivial modelled distribution must have a component of order $\leqslant 0$ so that necessarily $\alpha \leqslant 0$.

THEOREM 4.4 ([27, Thm. 4.7]). — *Let F, G be modelled distributions over a regularity structure endowed with a product as explained above. If $F \in \mathcal{D}^{\gamma_1}_{\alpha_1}$ and $G \in \mathcal{D}^{\gamma_2}_{\alpha_2}$ we have $FG \in \mathcal{D}^\gamma_{\alpha_1+\alpha_2}$ for $\gamma = (\gamma_1 + \alpha_2) \wedge (\gamma_2 + \alpha_1)$. Furthermore, we have for every compact set $\mathfrak{K}$*

$$\|FG\|_{\gamma,\mathfrak{K}} \lesssim \|F\|_{\gamma_1,\mathfrak{K}}\,\|G\|_{\gamma_1,\mathfrak{K}}\,. \tag{4.12}$$

Remark 4.5. — Unlike Theorem 2.13, this theorem does not require any condition on the exponents. Indeed, the product is always well-defined pointwise, independently of the choice of γ_1 and γ_2. However, we encounter a condition on γ when applying the *reconstruction operator*. Although the product is always defined as an abstract expansion, it is only for $\gamma > 0$ that this expansion uniquely describes a *real* distribution.

Example 4.6. — We have seen above in (3.10) that any $\mathcal{C}^\gamma_{\mathfrak{s}}$ function f can be lifted naturally to a modelled distribution in $\mathcal{D}^\gamma$ (which should be

---

[27] See [27, Def. 4.6] for precise definition.

denoted by $\mathcal{D}_0^\gamma$ here) by setting

$$F(z) = \sum_{|k|_{\mathfrak{s}} \leqslant \gamma} \frac{1}{k!} D^k f(z) \mathbf{X}^k . \qquad (4.13)$$

If we have another $\mathcal{C}_{\mathfrak{s}}^\gamma$ function g which is lifted to G in the same way we get

$$F\,G(z) = \sum_{|k|_{\mathfrak{s}} \leqslant \gamma} \frac{1}{k!} \sum_j \binom{k}{j} \left(D^j f(z) D^{k-j} g(z) \right) \mathbf{X}^k ,$$

which is nothing but Leibniz rule. Note, that here we have truncated the expansion to involve only those terms of order $\leqslant \gamma$. Indeed, the function FG is only of class $\mathcal{C}_{\mathfrak{s}}^\gamma$ and polynomials of order higher than γ give no information about the local behaviour of this function.

Example 4.7. — Now we can finally explain up to which order we need to expand Φ in order to solve the abstract fixed point problem for Φ_3^4. As in (4.6) we make the ansatz

$$\Phi(z) = \uparrow + \Phi_{\mathbf{1}}(z)\mathbf{1} - \curlyvee - 3\Phi_{\mathbf{1}}(z)\curlyvee + \langle \Phi_{\mathbf{X}}(z), \mathbf{X} \rangle . \qquad (4.14)$$

The term of lowest order in this description is the symbol $\uparrow$ which is of order $-\frac{1}{2} - \kappa$. Then we get

$$\Phi^3 := \curlyvee + 3\Phi_{\mathbf{1}}\curlyvee - 3\curlyvee + 3\Phi_{\mathbf{1}}^2\uparrow - 6\Phi_{\mathbf{1}}\curlyvee - 9\Phi_{\mathbf{1}}\curlyvee + 3\langle\Phi_{\mathbf{X}}, \curlyvee\mathbf{X}\rangle + \Phi_{\mathbf{1}}^3\mathbf{1} , \ (4.15)$$

where we included only terms of non-positive order. Using Theorem 2.13 we can conclude that for $\Phi \in \mathcal{D}_\alpha^\gamma$ we have $\Phi^3 \in \mathcal{D}_{3\alpha}^{\gamma - 2\alpha}$. This statement is always true for any γ, but in order to have a meaningful *reconstruction* of Φ^3 the exponent $\gamma + 2\alpha$ needs to be strictly positive. As $\alpha = -\frac{1}{2} - \kappa$ we need to describe Φ to order at least $\gamma > 1 + 2\kappa$.

4.4. Integration

At this stage, the only operation missing to define the fixed point operator (4.21) is the *integration map*. Recall that above in Theorem 2.8 we had stated that convolution with the heat kernel K can be defined for quite general distributions $f \in \mathcal{C}_{\mathfrak{s}}^\alpha$ and that (for $\alpha \notin \mathbf{Z}$) this operation improves the parabolic Hölder regularity by 2. This result is closely related to the fact that K is a singular kernel of order $-d_{\mathfrak{s}} + 2$, i.e. that K is a smooth function on $\mathbb{R} \times \mathbb{R}^d \setminus \{0\}$ with a well-controlled singularity at the origin. Our aim for this section is to define an analogue map $\mathcal{K}_\gamma$ that maps modelled distribution $F \in \mathcal{D}_\alpha^\gamma$ to $\mathcal{D}_{(\alpha+2)\wedge 0}^{\gamma+2}$. To make some expressions easier, we will from now on use a convention (slightly inconsistently with the previous sections but consistently with the notation used in [27]) to give a new interpretation

to the kernel K. We will replace the parabolic heat kernel by a kernel K which satisfies

(1) $|D^k K(z)| \lesssim \|z\|_{\mathfrak{s}}^{-d_{\mathfrak{s}}+2-|k|_{\mathfrak{s}}}$ for all multi-indices k (recall the definition of the parabolic dimension $d_{\mathfrak{s}} = 2 + d$).
(2) $K(z) = 0$ for all $z = (t, x)$ with $t < 0$.
(3) K has compact support in $\{z : \|z\|_{\mathfrak{s}} < 1\}$.
(4) $\int K(z) z^k \, \mathrm{d}z = 0$ for all multi-indices k with $|k|_{\mathfrak{s}} < \gamma$.

Of course, the Gaussian heat kernel satisfies assumptions (1) and (2) but not (3) and (4). However, for any $\gamma > 0$ it is possible to add a smooth function R_γ to the Gaussian kernel such that one obtains a kernel that also satisfies the assumptions (3) and (4). The convolution with a smooth function R_γ is an infinitely smoothing operation which can easily be dealt with separately. Therefore, for the rest of these notes we will assume that K satisfies all of these four assumptions, neglecting the extra terms that come from the convolution with R_γ.

Example 4.8. — Let us briefly discuss how to formulate a version of the classical Schauder estimate for $\mathcal{C}_{\mathfrak{s}}^\gamma$ functions (Theorem 2.8) using the regularity structure of abstract polynomials defined in Section 3.3. Let $f \colon \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be of regularity $\mathcal{C}_{\mathfrak{s}}^\gamma$ for some $\gamma \in (0, \infty) \setminus \mathbf{Z}$ and let F be its canonical lift to $\mathcal{D}^\gamma$. We want to define a map $\mathcal{K}_\gamma$ acting on $\mathcal{D}^\gamma$ which represents convolution with the kernel K. A natural definition would be

$$\mathcal{K}_\gamma F(z) = \sum_{|k|_{\mathfrak{s}} < \gamma+2} \frac{\mathbf{X}^k}{k!} \int_{\mathbb{R} \times \mathbb{R}^d} D^k K(z - y) \underbrace{\mathcal{R} F(y)}_{=f(y)} \, \mathrm{d}y \qquad (4.16)$$

but it is not obvious why this integral converges. Indeed, in general $|D^k K(z)| \sim |z|^{2 - d_{\mathfrak{s}} - |k|_{\mathfrak{s}}}$ which fails to be integrable for $|k|_{\mathfrak{s}} \geqslant 2$. However, if we use the regularity of F and replace $\mathcal{R} F(y)$ by $\mathcal{R} F(y) - \Pi_z F(y)$ (i.e. if we subtract the Taylor polynomial around z) we obtain a convergent expression

$$\mathcal{N}_\gamma F(z) := \sum_{|k|_{\mathfrak{s}} < \gamma+2} \frac{\mathbf{X}^k}{k!} \int_{\mathbb{R} \times \mathbb{R}^d} \underbrace{D^k K(z - y)}_{\lesssim \|z-y\|_{\mathfrak{s}}^{2-d_{\mathfrak{s}}-|k|_{\mathfrak{s}}}} \underbrace{(\mathcal{R} F(y) - \Pi_z F(y))}_{\|z-y\|_{\mathfrak{s}}^\gamma} \, \mathrm{d}y \, . \quad (4.17)$$

It remains to discuss, how to interpret the integrals of $D^k K(z - \cdot)$ against $\Pi_z F$. Note that this expression depends only *locally* on F at the point z and it is completely determined by $\int D^k K(z) z^\ell \mathrm{d}z$ for finitely many k and ℓ. But of course, (some of) these integrals still fail to converge absolutely. However, using the formal integration by part

$$\int_{\mathbb{R} \times \mathbb{R}^d} D^{(k)} K(z) z^\ell \, \mathrm{d}z = (-1)^{|k|} \int_{\mathbb{R} \times \mathbb{R}^d} K(z) D^{(k)} z^\ell \, \mathrm{d}z \, ,$$

it is not hard to show that these integrals converge if they are interpreted as principal values

$$\lim_{\varepsilon \to 0} \int_{|z|>\varepsilon} z^\ell D^k K(z) \, \mathrm{d}z \ , \tag{4.18}$$

and furthermore, using our convenient assumption that K integrates to zero against polynomials, one can see that these limits are zero. We can therefore define $\mathcal{K}_\gamma = \mathcal{N}_\gamma$ and the operator $\mathcal{K}_\gamma$ defined this way does indeed map $\mathcal{D}_0^\gamma$ to $\mathcal{D}_0^{\gamma+2}$.

Of course our main focus is on models and regularity structures that are larger than $\bar{\mathcal{T}}$ which allow us to work with singular distributions. In this case the integration map $\mathcal{K}_\gamma$ will in general consist of three different components

$$\mathcal{K}_\gamma = \mathcal{I} + \mathcal{J}_\gamma + \mathcal{N}_\gamma \ .$$

We have already encountered the operator $\mathcal{N}_\gamma$ in Example 4.8 above. The defining equation (4.17) remains meaningful if it is interpreted in a distributional sense. The operators $\mathcal{I}$ and $\mathcal{J}_\gamma$ correspond to the additional information provided by non-polynomial symbols. Both operators vanish when applied to abstract polynomials.

The operator $\mathcal{I}$ takes values in the span of the abstract symbols $\mathfrak{T}$ which are *not* polynomials. Like the multiplication, it is defined as a linear operator on T and its definition was essentially part of the recursive construction of the regularity structure $\mathcal{T}$. We have, for example

$$\mathcal{I}\Xi = \mathord{\uparrow} \ ,$$

etc. Note in particular, that (unlike a convolution operator) $\mathcal{I}$ acts *locally* on modelled distributions. One important property is that by definition $\mathcal{I}$ increases the order of every $a \in T_\alpha$ by 2. As for the product, we do not have to give a non-trivial interpretation for $\mathcal{I}\tau$ for all τ. Indeed, in order to describe our solution to a certain regularity it is sufficient to keep those basis elements of order $< \gamma$.

Of course, at this stage the definition of the *abstract integration map* $\mathcal{I}$ has nothing to do with the kernel K. The connection with K is established in the choice of the model. We had already discussed this issue in the context of the *canonical model* in (4.2). We will now turn the relevant property of the canonical model into a definition.

DEFINITION 4.9. — *A model is* admissible *if we have* $(\Pi_z \mathbf{X}^k)(y) = (y-z)^k$ *and*

$$(\Pi_z \mathcal{I}[\tau])(\bar{z}) := \int_{\mathbb{R}^{d+1}} \mathrm{d}y \; K(\bar{z} - y)(\Pi_z \tau)(y)$$

$$- \sum_{\substack{k \\ |k|_{\mathfrak{s}} < |\tau|+2}} \frac{(\bar{z} - z)^k}{k!} \int_{\mathbb{R}^{d+1}} \mathrm{d}y \; (D^k K)(z - y)(\Pi_z \tau)(y) \;.$$

We will denote by $\mathcal{M}_0 \subseteq \mathcal{M}$ *the space of admissible models.*

The construction of *canonical models* we explained above, automatically produces admissible models[28]. But it is non-trivial to perform the renormalization such that the models remain admissible.

The operator $\mathcal{J}_\gamma$ takes values in the abstract polynomials. It is the analog of the integrals (4.18) which in the case of a general regularity structure cannot be removed by a convenient choice of kernel. The operator is defined as

$$\mathcal{J}_\gamma F(z) := \sum_{|k|_{\mathfrak{s}} < \alpha+2} \frac{\mathbf{X}^k}{k!} \int D^{(k)} K(z - y) \, \Pi_z F(y) \, \mathrm{d}y \;, \qquad (4.19)$$

With these definitions in place we have the following result.

THEOREM 4.10. — *Let* $\mathcal{T}$ *be a regularity structure endowed with an admissible model and assume that* $\gamma \notin A$. *Then the operator* $\mathcal{K}_\gamma$ *is compatible with integration against the kernel* K *in the sense that*

$$\mathcal{R}\mathcal{K}_\gamma F = K * \mathcal{R}F \;. \qquad (4.20)$$

Furthermore $\mathcal{K}_\gamma$ *maps* $\mathcal{D}_\alpha^\gamma$ *into* $\mathcal{D}_{(\alpha+2)\wedge 0}^{\gamma+2}$ *and we have for every compact* $\mathfrak{K}$

$$\|\mathcal{K}_\gamma F\|_{\gamma+2, \mathfrak{K}} \lesssim \|F\|_{\gamma, \bar{\mathfrak{K}}} \;.$$

where $\bar{\mathfrak{K}} = \{z: \; \inf_{\bar{z} \in \mathfrak{K}} \|z - \bar{z}\| \leqslant 1\}$.

Example 4.11. — For Φ_2^4 we had made the ansatz $\Phi(z) = \mathord{\uparrow} + \Phi_{\mathbf{1}}\mathbf{1}$ and we had already seen above that this implies

$$\Phi^3 = \mathord{\Psi} + 3\Phi_{\mathbf{1}}\mathord{\vee} + 3\Phi_{\mathbf{1}}^2\mathord{\uparrow} + \Phi_{\mathbf{1}}^3\mathbf{1} \;.$$

Let us now give an explicit description of, $\mathcal{K}_\gamma \Xi - \mathcal{K}_\gamma[\Phi^3]$ because it is instructive. First of all, we have $\mathcal{K}_\gamma \Xi = \mathcal{I}[\Xi] = \mathord{\uparrow}$. Indeed, $\mathcal{I}[\Xi]$ has order $-\kappa < 0$ and

---

[28] Of course we discussed the construction of canonical models in the context of the Gaussian heat kernel, but the construction goes through unchanged if it is replaced by our modified kernel.

therefore, the sum (4.19) which defines the operator $\mathcal{J}$ is empty. Furthermore, the fact that $\Pi_z \Xi$ does not depend on z implies that $(\mathcal{R}\Xi)(z) = \Pi_z \Xi$ which in turn implies by (4.17) that $\mathcal{N}_\gamma \Xi$ vanishes as well.

On the other hand, we can take $-\mathcal{K}_\gamma \Phi^3 = -(\mathcal{J}_\gamma + \mathcal{N}_\gamma)\Phi^3$. Indeed, the symbols $\uparrow, \vee, \psi$ appearing in Φ^3 have order $-\kappa, -2\kappa, -3\kappa$ so that the abstract integration map $\mathcal{I}$ acting on these symbols would produce terms of order > 1. We do not require a description to such order, so these terms and the corresponding $\mathcal{J}$ can be dropped. We get

$$(\mathcal{J}_\gamma + \mathcal{N}_\gamma)(\Phi^3) = K * \mathcal{R}\Big(\psi + 3\Phi_1 \vee + 3\Phi_1^2 \uparrow + \Phi_1^3 \mathbf{1}\Big)\mathbf{1}\ .$$

The reconstruction operator gives $\mathcal{R}\psi = \psi$, $\mathcal{R}\Phi_1 \vee = \Phi_1 \vee$, $\mathcal{R}\Phi_1^2 \uparrow = \Phi_1^2 \uparrow$ and $\mathcal{R}\Phi_1^3 \mathbf{1} = \Phi_1^{3\,(29)}$. Hence, the equation for Φ_1 reduces to the equation for the remainder v_δ in the Da Prato–Debussche method. In this context, the continuity of the multiplication of modelled distribution, together with the existence and continuity of the reconstruction operator take the role of the multiplicative inequality, Theorem 2.13. This is actually a general fact. In [27, Thm 4.14] it is shown how Theorem 2.13 can be derived as a consequence of these two statements.

4.5. The fixed point argument

We now state a theorem guaranteeing the existence of a modelled distribution which solves the abstract fixed point problem (4.21). Our discussion will be informal; a precise version of such a theorem, stated in a quite general context, can be found in [27, §7.3, Thm. 7.8].

We aim to prove existence of solutions to the dynamic Φ_3^4 equation by solving a fixed point problem in a space $\mathcal{D}^\gamma$ of modelled distributions on $[0, T] \times \mathbb{R}^d$. The tools we developed in the previous sections will show that the non-linearity is locally Lipschitz continuous and for T small enough we can apply the contraction mapping theorem on some ball in $\mathcal{D}^\gamma$.

At this point it is important to remember that we have derived uniform bounds on models only locally in space-time. Indeed, going through the proof of the Kolmogorov Lemma, Theorem 2.7, the reader can easily convince himself that the constants explode over infinite space-time domains and the same phenomenon presents itself in the construction of various models discussed at the beginning of this section. This problem could be circumvented, by introducing weights in the norms that measure these models (as has been

---

$^{(29)}$ Actually, the processes $\uparrow, \vee$ and ψ should be constructed with the modified kernel.

implemented in [29, 38]) but this makes the deterministic analysis more difficult. Here we choose the simpler situation and compactify space by assuming that the noise is periodic.

Accordingly, we now assume that our space-time white noise ξ is defined on $\mathbf{T}^d \times \mathbb{R}$. When convenient we interpret ξ as a distribution on $\mathbb{R}^d \times \mathbb{R}$ which is periodic in space. We will again lift realizations of the noise to admissible models as before (see [27, §3.6] for the precise notion of periodicity for models). For any "periodic" model (Π, Γ) we define $\mathcal{D}^\gamma(\Pi, \Gamma, \Lambda_T)$ to be the family of modelled distributions $F : \mathbf{T}^d \times [0, T] \to T$ which satisfy condition (3.7). We define $\mathcal{M}_0 \ltimes \mathcal{D}^\gamma(\Lambda_T)$ to be the set of all triples (Π, Γ, F) with $(\Pi, \Gamma) \in \mathcal{M}_0$ "periodic" and $F \in \mathcal{D}^\gamma(\Pi, \Gamma, \Lambda_T)$. As before it will be important that this space can be equipped with a metric which behaves well with the machinery of the theory of regularity structures. With all of this notation in hand we can now state the following theorem.

THEOREM 4.12. — *Let $\mathcal{T}$ be a regularity structure for Φ_3^4, where the corresponding list of symbols $\mathfrak{T}$ was defined with a choice of $\gamma > 1 + 3\kappa$. Then for any admissible model (Π, Γ) there exists a strictly positive $T > 0$ and a unique modelled distribution $\Phi \in \mathcal{D}^\gamma(\Pi, \Gamma, \Lambda_T)$ that solves the fixed point equation*

$$\Phi = -\mathcal{K}_\gamma \Phi^3 + \mathcal{K}_\gamma \Xi . \tag{4.21}$$

Additionally, T is lower semicontinuous in (Π, Γ). Furthermore, if for some $(\bar{\Pi}, \bar{\Gamma})$ we have $T[(\bar{\Pi}, \bar{\Gamma})] > t$ then $\Phi : (\Pi, \Gamma) \mapsto (\Pi, \Gamma, \Phi) \in \mathcal{M}_0 \ltimes \mathcal{D}^\gamma(\Lambda_t)$ is continuous in a neighbourhood of $(\bar{\Pi}, \bar{\Gamma})$.

We make a few remarks about the contraction mapping argument used for the above theorem. The continuity of the mapping $\Phi \mapsto \Phi^3$ and the integration operator $\mathcal{K}_\gamma$ immediately imply that the mapping $\Phi \mapsto -\mathcal{K}_\gamma \Phi^3 + \mathcal{K}_\gamma \Xi$ is Lipschitz continuous on every ball of $\mathcal{D}^\gamma(\Lambda_t)$. The Lipschitz constant can be made arbitrarily small by choosing a slightly smaller γ (which for $t < 1$ produces a small power of t in front of the bounds) and then choosing T small enough.

Similar arguments yield the essential continuity statement promised by the last sentence of Theorem 4.12.

Remark 4.13. — Up to now we have always assumed that the initial data for our fixed point problem is zero. This is quite unsatisfactory for the solution theory, because it prevents us from *restarting* the solution at T to obtain a maximal solution. The theory of [27] does allow for the restarting of solutions but one must work with larger classes of modelled distributions. Since initial condition will typically not have a nice local description

the theory introduces spaces of "singular" modelled distributions where the given local expansions are allowed to blow up near the time zero hyperplane (see [27, §6]).

We now continue our discussion but will suppress the fact that we are actually working in a spatially periodic setting and finite time horizon. The particular form of the modelled distribution $\Phi \in \mathcal{D}^\gamma$ which solves the abstract fixed point problem can now be deduced by running a few steps of a Picard iteration. We run through this computation now which will end up justifying the ansatz (4.6). We start the iteration by setting $\Phi_0(z) = 0$. Applying the map $\Phi \mapsto -\mathcal{K}_\gamma \Phi^3 + \mathcal{K}_\gamma \Xi$ to Φ_0 gives

$$\Phi_1(z) = \mathcal{K}_\gamma \Xi = \mathcal{I}\Xi = \uparrow \,.$$

[30] Here we used that $\mathcal{J}\Xi = \mathcal{N}_\gamma \Xi = 0$ as was explained in Example 4.11. Applying this map again then gives

$$\Phi_2(z) = \uparrow - \big[\mathcal{K}_\gamma (\Phi_1)^3 \big](z) = \uparrow - (\mathcal{I} + \mathcal{J} + \mathcal{N}_\gamma)\mathbf{\Psi} = \uparrow - \mathbf{Y} - \mathbf{Y}(z)\mathbf{1} \,.$$

Observe that the appearance of the positive order symbol $\mathbf{Y}$ automatically produces the first "polynomial" (the symbol $\mathbf{1}$). Here the notation $\mathbf{Y}(\bullet)$ (which is colored black) refers to the corresponding concrete space-time distribution which was introduced in Section 3.1. In going to our expression Φ_2 we used that $\mathcal{N}_\gamma \mathbf{\Psi} = 0$, this is because $\Pi_z \mathbf{\Psi}$ does not depend on z and hence $\mathcal{R}\mathbf{\Psi} = \Pi_z \mathbf{\Psi}$.

Going one step further in the Picard iteration gives

$$\Phi_3(z) = \uparrow - \big[\mathcal{K}_\gamma (\Phi_2)^3 \big](z)$$
$$= \uparrow - \mathbf{Y} - \mathbf{Y}(z)\mathbf{1} + \mathcal{K}_\gamma \Big(3\mathbf{\Psi} + 3\mathbf{Y}(z)\mathbf{v} - 6\mathbf{Y}(z)\mathbf{Y} - 3\mathbf{Y}(z)^2\uparrow \Big) \,.$$

Here we have dropped all terms of order > 0 under the operator $\mathcal{K}_\gamma$ because we do not need them. Indeed, the two requirements that determine the degree to which we have to expand each quantity are:

- The solution Φ should solve a fixed point problem in $\mathcal{D}_\alpha^\gamma$ for $\gamma \approx 1 + 3\kappa$ and $\alpha = -\frac{1}{2} - \kappa$. Therefore, we need to keep all symbols with order less than or equal to γ in the expansion Φ.
- Below we will apply the reconstruction operator to the right hand side of the fixed point problem in order to identify the equation the reconstruction of Φ solves. In order to be able to do that we need to ensure that quantities under the integral operator are described to strictly positive order.

---

[30] Note that the Φ_1 is different from $\Phi_{\mathbf{1}}$ with a blue bold subscript.

It is now clear the fixed point $\Phi \in \mathcal{D}^\gamma$ for the map $\Phi \mapsto -\mathcal{K}_\gamma \Phi^3 + \mathcal{K}_\gamma \Xi$ will have the property that the symbols $\mathord{\uparrow}$ and $\mathord{\Psi}$ enter with z-independent coefficients. Indeed, both symbols only ever arise after integrating the terms Ξ and $\mathord{\vee}$ from the previous step, both of which cannot have a non-constant prefactor. Furthermore, it is clear why the pre-factor of $\mathord{\Upsilon}$ has to be $\Phi_{\mathbf{1}}$. Indeed, this symbol only arises after applying $\mathcal{I}$ to $\mathord{\vee}$ which in turn only appears from the multiplication of $\mathord{\vee}$ with $\Phi_{\mathbf{1}}\mathbf{1}$.

As mentioned above, we will now apply the reconstruction operators to Φ to get concrete space-time distributions and then show that these objects satisfy certain concrete PDE. For $\delta > 0$ let $(\Pi^\delta, \Gamma^\delta)$ be the canonical model built from the smoothed noise ξ_δ. Let $\mathcal{R}^\delta$ be the associated reconstruction operator on the $\mathcal{D}^\gamma$ space built from the canonical model with γ slightly larger than 1 as in Theorem 4.12. We denote by Φ_δ the modelled distribution which is the solution to the corresponding abstract fixed point problem. It follows that

$$
\begin{aligned}
\left(\mathcal{R}^\delta \Phi_\delta\right)(z) &= \mathcal{R}^\delta \left[\mathcal{K}\left(\Xi - \Phi_\delta^3\right)\right](z) \\
&= \int \mathrm{d}y \; K(z - y) \left[\mathcal{R}^\delta \left(\Xi - \Phi_\delta^3\right)\right](y) \\
&= \int \mathrm{d}y \; K(z - y) \left[\left(\Pi_y^\delta \Xi\right)(y) - \left(\Pi_y^\delta \Phi_\delta^3(y)\right)(y)\right] \\
&= \int \mathrm{d}y \; K(z - y) \left[\left(\Pi_y^\delta \Xi\right)(y) - \left(\Pi_y^\delta \Phi_\delta(y)\right)^3(y)\right] \\
&= \int \mathrm{d}y \; K(z - y) \left[\xi_\delta(y) - \left(\mathcal{R}^\delta \Phi_\delta\right)^3(y)\right].
\end{aligned}
$$

The first equality above is just the fixed point relation. The second equality is (4.20), the third is (4.8), the fourth is a consequence of the product property of the canonical model, and in the final equality we again use (4.8). It follows that $\mathcal{R}^\delta \Phi_\delta$ is the mild solution to the equation

$$
\partial_t \varphi_\delta = \Delta \varphi_\delta - \varphi_\delta^3 + \xi_\delta .
$$

Now let $(\hat{\Pi}^\delta, \hat{\Gamma}^\delta)$ be the renormalized model we introduced earlier, and let $\hat{\mathcal{R}}^\delta$ be the reconstruction operator on the associated $\hat{\mathcal{D}}^\gamma$ space (with γ as before). Let $\hat{\Phi}_\delta$ be the solution to the abstract fixed point problem with model $(\hat{\Pi}^\delta, \hat{\Gamma}^\delta)$. We then have

$$
\begin{aligned}
\left(\hat{\mathcal{R}}^\delta \hat{\Phi}_\delta\right)(z) &= \int \mathrm{d}y \; K(z - y) \left[\left(\hat{\Pi}_y^\delta \Xi\right)(y) - \left(\hat{\Pi}_y^\delta \hat{\Phi}_\delta^3(y)\right)(y)\right] \\
&= \int \mathrm{d}y \; K(z - y) \left[\xi_\delta(y) - \left(\Pi_y^\delta M \hat{\Phi}_\delta^3(y)\right)(y)\right].
\end{aligned}
\tag{4.22}
$$

In going to the last line we have used (4.5).

We know that $\hat{\Phi}_\delta$ has an expansion of the form (4.14) where the spatially varying coefficients appearing in the expansion are unknown. Since $\left(\Pi_y^\delta \tau\right)(y) = 0$ for any homogenous τ with $|\tau| > 0$ we can replace $M^\delta \hat{\Phi}_\delta^3(y)$ with something equivalent modulo symbols of strictly positive order. Denoting this approximate equivalence by $\approx$, the reader is encouraged to apply M^δ to the formula (4.15) to see that

$$M^\delta \hat{\Phi}_\delta^3(y) \approx (M^\delta \hat{\Phi}_\delta(y))^3 - (3C_\delta + 9\tilde{C}_\delta)M^\delta \hat{\Phi}_\delta(y) \,.$$

Applying the canonical model Π_y^δ to both sides then gives

$$\left(\Pi_y^\delta M\hat{\Phi}_\delta^3(y)\right)(y) = \left(\Pi_y^\delta (M^\delta \hat{\Phi}_\delta(y))^3\right)(y) - (3C_\delta + 9\tilde{C}_\delta)\left(\Pi_y^\delta M^\delta \hat{\Phi}_\delta(y)\right)(y)$$

$$= \left(\Pi_y^\delta M^\delta \hat{\Phi}_\delta(y)\right)^3(y) - (3C_\delta + 9\tilde{C}_\delta)\left(\Pi_y^\delta M^\delta \hat{\Phi}_\delta(y)\right)(y)$$

$$= \left(\hat{\Pi}_y^\delta \hat{\Phi}_\delta(y)\right)^3(y) - (3C_\delta + 9\tilde{C}_\delta)\left(\hat{\Pi}_y^\delta \hat{\Phi}_\delta(y)\right)(y)$$

$$= \left(\hat{\mathcal{R}}^\delta \hat{\Phi}_\delta\right)^3(y) - (3C_\delta + 9\tilde{C}_\delta)\left(\hat{\mathcal{R}}^\delta \hat{\Phi}_\delta\right)(y) \,.$$

Inserting this into (4.22) immediately yields that $\left(\hat{\mathcal{R}}^\delta \hat{\Phi}_\delta\right)(z)$ is the mild solution to the PDE

$$\partial_t \varphi_\delta = \Delta\varphi_\delta - \varphi_\delta^3 + (3C_\delta + 9\tilde{C}_\delta)\varphi_\delta + \xi_\delta \,. \tag{4.23}$$

We now take advantage of the fact that all of the abstract machinery introduced in this lecture has good continuity properties with respect to the convergence of models. If the models $(\hat{\Pi}^\delta, \hat{\Gamma}^\delta)$ converge in probability to a limiting model $(\hat{\Pi}, \hat{\Gamma})$ as $\delta \downarrow 0$ then from Theorem 4.12 it follows that the triples $(\hat{\Pi}^\delta, \hat{\Gamma}^\delta, \hat{\Phi}_\delta)$, viewed as random elements of the space $\mathcal{M} \ltimes \mathcal{D}^\gamma$, converge in probability to a limiting triple $(\hat{\Pi}, \hat{\Gamma}, \hat{\Phi})$ as $\delta \downarrow 0$.

Theorem 4.2 then implies that the solutions of (4.23), given by $\hat{\mathcal{R}}^\delta \hat{\Phi}^\delta$, converge in probability to a limiting space-time distribution we will call φ. Here convergence in probability takes place on the space $\mathcal{C}_{\mathfrak{s}}^\alpha$. While one may not be able to write down an explicit SPDE that the φ satisfies, we can say φ solves the abstract formulation of the given SPDE since $\varphi = \tilde{\mathcal{R}}[(\hat{\Pi}, \hat{\Gamma}, \hat{\Phi})]$ and the triple $(\hat{\Pi}, \hat{\Gamma}, \hat{\Phi})$ is a solution to our abstract fixed point problem.

Appendix A. A primer on iterated stochastic integrals

In this appendix we collect some facts about iterated stochastic integrals used in Lecture 2. Our discussion is brief and somewhat formal, a detailed exposition can be found in [40, Chap. 1]. Throughout the appendix we adopt

a slightly more general framework than in Lecture 2 and replace $\mathbb{R} \times \mathbb{R}^d$ or $\mathbb{R} \times \mathbf{T}^d$ by an arbitrary measure space $(E, \mathcal{E})$ endowed with a sigma-finite non-atomic measure μ. Extending the definition presented in Section 1.1.1 a *white noise* is then defined as a centred Gaussian family of random variables (ξ, φ) indexed by $\varphi \in L^2(E, \mu)$ which satisfy

$$\mathbb{E}(\xi, \varphi_1)(\xi, \varphi_2) = \int_E \varphi_1(z)\, \varphi_2(z)\, \mu(\mathrm{d}z)\ . \tag{A.1}$$

It is particularly interesting to evaluate ξ at indicator functions $\mathbf{1}_A$ of measurable sets in $A \in \mathcal{E}$ with $\mu(A) < \infty$ and we write $\xi(A)$ as a shorthand for $\xi(\mathbf{1}_A)$. The following properties follow

- $\mathbb{E}\xi(A) = 0$ and $\mathbb{E}\xi(A)^2 = \mu(A)$.
- If A_1 and A_2 are disjoint, then $\xi(A_1)$ and $\xi(A_2)$ are independent.
- If $(A_j)_{j \in \mathbf{N}} \ldots$ are pairwise disjoint and of finite measure, then $\xi(\cup_j A_j) = \sum_j \xi(A_j)$, where the convergence holds in $L^2(\Omega, \mathbb{P})$.

Although the last identity suggests to interpret $A \mapsto \xi(A)$ as a random signed measure, it is important to note that in general the ξ does not have a modification as a random signed measure (cf. the regularity discussion in Besov spaces above).

We now discuss, how to construct an iterated stochastic integrals of the type "$\int_{E^n} f(z_1, z_2, \ldots, z_n)\xi(\mathrm{d}z_1) \ldots \xi(\mathrm{d}z_n)$" for $f \in L^2(E^n, \mu^{\otimes n})$. For simplicity we will only treat the case $n = 2$, the general case of n-fold iterated integrals following in a similar way. In this case, we will call *elementary* any $f \colon E \times E \to \mathbb{R}$ of the form

$$f = \sum_{\substack{j,k=1 \\ j \neq k}}^{N} \alpha_{j,k} \mathbf{1}_{A_j \times A_k}\ ,$$

for pairwise disjoint sets $A_1, \ldots, A_N$ with finite measure and real coefficients $\alpha_{j,k}$. Note that such a function f is necessarily zero on the diagonal $f(z, z) = 0$ for $z \in E$. We define for such an f

$$\int_{E \times E} f(z_1, z_2)\xi(\mathrm{d}z_1)\, \xi(\mathrm{d}z_2) = \sum_{\substack{j,k=1 \\ j \neq k}}^{N} \alpha_{j,k}\xi(A_j)\xi(A_k)\ .$$

We then get the following identity which resembles the Itô isometry:

$$
\begin{aligned}
\mathbb{E}\bigg(\int_{E \times E} & f(z_1, z_2) \xi(\mathrm{d}z_1)\, \xi(\mathrm{d}z_2) \bigg)^2 \\
&= \sum_{j_1 < k_1} \sum_{j_2 < k_2} (\alpha_{j_1, k_1} + \alpha_{k_1, j_1})(\alpha_{j_2, k_2} + \alpha_{k_2, j_2})\, \mathbb{E}(\xi(A_{j_1})\xi(A_{k_1})\xi(A_{j_2})\xi(A_{k_2})) \\
&= \sum_{j < k} (\alpha_{j,k} + \alpha_{k,j})^2 \mu(A_k)\mu(A_j) \\
&= 2 \int_{E \times E} \left(\frac{1}{2}\big(f(z_1, z_2) + f(z_2, z_1)\big) \right)^2 \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) \\
&\leqslant 2 \int_{E \times E} \big(f(z_1, z_2)\big)^2 \mu(\mathrm{d}z_1)\mu(\mathrm{d}z_2) \,. \tag{A.2}
\end{aligned}
$$

Note that we have crucially used the fact that no "diagonal terms" appear when passing from the second to the third line. It is relatively easy to show that the elementary functions are dense in $L^2(E \times E, \mu \otimes \mu)$ (due to the off-diagonal assumption this is only true for non-atomic measures) and hence we can extend the definition of $\int_{E \times E} f(z_1, z_2)\, \xi(\mathrm{d}z_1)\, \xi(\mathrm{d}z_2)$ to all of $L^2(E \times E, \mu \otimes \mu)$.

However, the fact that we have defined the iterated integral as a limit of approximations that "cut out" the diagonal has an effect, when treating non-linear functions of iterated stochastic integrals. Formally, for $f \in L^2(E, \mu)$ one might expect the identity

$$
\left(\int_E f(z)\xi(\mathrm{d}z) \right)^2 = \int_{E \times E} f(z_1)f(z_2)\xi(\mathrm{d}z_1)\xi(\mathrm{d}z_2) \,, \tag{A.3}
$$

which "follows" by formally expanding the integral. But at this point it becomes relevant that as mentioned above ξ is typically *not* a random measure, so that this operation is not admissible. In order to get the right answer, we have to approximate f by simple functions $f \approx \sum_{j=1}^{N} \alpha_j A_j$. Mimicking the construction of the iterated integrals above we write

$$
\left(\int f(z)\xi(\mathrm{d}z) \right)^2 \approx \sum_{j \neq k} \alpha_j \alpha_k \xi(A_j)\xi(A_k) + \sum_j \alpha_j^2 \xi(A_j)^2 \,.
$$

As expected, the first sum involving only off-diagonal entries will converge to $\int f(z_1)f(z_2)\xi(\mathrm{d}z_1)\xi(\mathrm{d}z_2)$ as the partition (A_j) gets finer. However, differing from the case where ξ is a measure, the "diagonal" term does not vanish in the limit. Indeed, its expectation is given by

$$
\sum_j \alpha_j^2\, \mathbb{E}\xi(A_j)^2 = \sum_j \alpha_j^2\, \mu(A_j) \approx \int_E f(z)\mu(\mathrm{d}z) \,,
$$

but the variance of this term will go to zero as the partition gets finer. This suggests that instead of (A.3) we should get

$$\left(\int_E f(z)\xi(\mathrm{d}z) \right)^2 = \int_{E \times E} f(z_1)f(z_2)\xi(\mathrm{d}z_1)\xi(\mathrm{d}z_2) + \int_E f(z)^2 \, \mathrm{d}z \;, \quad \text{(A.4)}$$

and this formula is indeed true.

Remark A.1. — In the one-dimensional case, i.e. when $E = \mathbb{R}$ and μ is the Lebesgue measure our construction yields iterated *Itô* integrals. In the case where $f = \mathbf{1}_{[0,t]}$ the formula (A.4) reduces to

$$\left(\int \mathbf{1}_{[0,t]}(s)\xi(\mathrm{d}s) \right)^2 = \int_{\mathbb{R} \times \mathbb{R}} \mathbf{1}_{[0,t]}(s_1)\mathbf{1}_{[0,t]}(s_2) \, \xi(\mathrm{d}s_2) \, \xi(\mathrm{d}s_1) + t$$

which in the more common notation of stochastic calculus reduces to the Itô formula

$$W_t^2 = 2 \int_0^t \left(\int_0^{s_1} \mathrm{d}W_{s_2} \right) \mathrm{d}W_{s_1} + t \;.$$

Remark A.2. — The generalisation to iterated integrals of arbitrary order follows a similar pattern. We leave it as an exercise to the reader to convince himself that for $n = 3$ formula (A.4) becomes

$$\left(\int_E f(z)\xi(\mathrm{d}z) \right)^3 = \int_{E \times E} f(z_1)f(z_2)f(z_3)\xi(\mathrm{d}z_1)\xi(\mathrm{d}z_2)\xi(\mathrm{d}z_3)$$

$$+ 3 \int_E f(z)^2 \, \mathrm{d}z \int_E f(z)\xi(\mathrm{d}z) \;. \quad \text{(A.5)}$$

For larger n such identities are expressed most conveniently in terms of *Hermite* polynomials H_n. For example, one gets

$$H_n\left(\int f(z)\xi(\mathrm{d}z), \|f\|_{L^2(\mu)} \right) = \int \ldots \int \prod_{j=1}^n f(z_j)\xi(\mathrm{d}z_1)\ldots\xi(\mathrm{d}z_n) \;,$$

where $H_0(Z,\sigma) = 1$, $H_1(Z,\sigma) = Z$, $H_2(Z,\sigma) = Z^2 - \sigma^2$, $H_3(Z,\sigma) = Z^3 - 3\sigma^2 Z$ etc. (see [40, Prop. 1.1.4]).

At this point almost all the tools we need in the analysis of non-linear functionals of Gaussian processes are in place. For example, we use expression (A.5) to decompose $\mathchar"013_\delta^3$ into two parts, the variances of each can be evaluated explicitly by the L^2 isometry (A.2). However, if we wanted to feed these bounds directly into the Kolmogorov Theorem 2.7 we would lose too much regularity. As in the case of white noise, above we have to replace the L^2-type bounds by L^p bounds for p large enough. In the Gaussian case we used the fact that for centred Gaussian random variables all moments are controlled by the variance. Fortunately iterated stochastic integrals satisfy

a similar property. This is the content of the famous *Nelson estimate* which states in our context that

$$\mathbb{E}\left(\int f(z_1, \ldots, z_n)\xi(\mathrm{d}z_1) \ldots \xi(\mathrm{d}z_n) \right)^p$$

$$\leqslant C_{n,p}\left(\int f(z_1, \ldots, z_n)^2 \mu(\mathrm{d}z_1) \ldots \mu(\mathrm{d}z_n) \right)^{\frac{p}{2}}. \quad \text{(A.6)}$$

This estimate is an immediate consequence of the hypercontractivity of the Ornstein–Uhlenbeck semigroup, see [40, Thm. 1.4.1].

Bibliography

[1] M. AIZENMAN, "Geometric analysis of φ^4 fields and Ising models. I, II", *Commun. Math. Phys.* **86** (1982), no. 1, p. 1-48.

[2] G. AMIR, I. CORWIN & J. QUASTEL, "Probability distribution of the free energy of the continuum directed random polymer in $1 + 1$ dimensions", *Commun. Pure Appl. Math.* **64** (2011), no. 4, p. 466-537.

[3] H. BAHOURI, J.-Y. CHEMIN & R. DANCHIN, *Fourier analysis and nonlinear partial differential equations*, Grundlehren der Mathematischen Wissenschaften, vol. 343, Springer, 2011, xvi+523 pages.

[4] L. BERTINI & G. GIACOMIN, "Stochastic Burgers and KPZ equations from particle systems", *Commun. Math. Phys.* **183** (1997), no. 3, p. 571-607.

[5] L. BERTINI, E. PRESUTTI, B. RÜDIGER & E. SAADA, "Dynamical fluctuations at the critical point: convergence to a nonlinear stochastic PDE", *Teor. Veroyatnost. i Primenen.* **38** (1993), no. 4, p. 689-741.

[6] J. BOURGAIN, "Invariant measures for the 2D-defocusing nonlinear Schrödinger equation", *Commun. Math. Phys.* **176** (1996), no. 2, p. 421-445.

[7] D. C. BRYDGES, J. FRÖHLICH & A. D. SOKAL, "A new proof of the existence and nontriviality of the continuum φ_2^4 and φ_3^4 quantum field theories", *Commun. Math. Phys.* **91** (1983), no. 2, p. 141-186.

[8] R. CATELLIER & K. CHOUK, "Paracontrolled distributions and the 3-dimensional stochastic quantization equation", `https://arxiv.org/abs/1310.6869v1`, 2013.

[9] I. CORWIN & J. QUASTEL, "Renormalization fixed point of the KPZ universality class", *J. Stat. Phys.* **160** (2015), no. 4, p. 815-834.

[10] L. COUTIN & Z. QIAN, "Stochastic analysis, rough path analysis and fractional Brownian motions", *Probab. Theory Relat. Fields* **122** (2002), no. 1, p. 108-140.

[11] G. DA PRATO & A. DEBUSSCHE, "Strong solutions to the stochastic quantization equations", *Ann. Probab.* **31** (2003), no. 4, p. 1900-1916.

[12] G. DA PRATO & J. ZABCZYK, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992, xviii+454 pages.

[13] J. FELDMAN, "The $\lambda\varphi_3^4$ field theory in a finite volume", *Commun. Math. Phys.* **37** (1974), no. 2, p. 93-120.

[14] J. FRITZ & B. RÜDIGER, "Time dependent critical fluctuations of a one-dimensional local mean field model", *Probab. Theory Relat. Fields* **103** (1995), no. 3, p. 381-407.

[15] P. K. FRIZ & M. HAIRER, *A course on rough paths*, Universitext, Springer, 2014, xiv+251 pages.

[16] J. FRÖHLICH, "On the triviality of $\Lambda\Phi^4$ theories and the approach to the critical-point in D$\geqslant$ 4-dimensions", *Nuclear Physics B* **200** (1982), no. 2, p. 281-296.

[17] T. FUNAKI, "Random motion of strings and related stochastic evolution equations", *Nagoya Math. J.* **89** (1983), p. 129-193.

[18] G. GIACOMIN, J. L. LEBOWITZ & E. PRESUTTI, "Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems", in *Stochastic partial differential equations: six perspectives*, Mathematical Surveys and Monographs, vol. 64, American Mathematical Society, 1999, p. 107-152.

[19] J. GLIMM & A. JAFFE, *Quantum physics. A functional integral point of view*, Springer, 1981, xx+417 pages.

[20] J. GLIMM, A. JAFFE & T. SPENCER, "The Wightman axioms and particle structure in the P$(\varphi)_2$ quantum field model", *Ann. Math.* **100** (1974), no. 3, p. 585-632.

[21] ———, "Phase transitions for φ_2^4 quantum fields", *Commun. Math. Phys.* **45** (1975), no. 3, p. 203-216.

[22] M. GUBINELLI, "Controlling rough paths", *J. Funct. Anal.* **216** (2004), no. 1, p. 86-140.

[23] M. GUBINELLI, P. IMKELLER & N. PERKOWSKI, "Paracontrolled distributions and singular PDEs", *Forum Math. Pi* **3** (2015), Article ID e6, 75 p.

[24] M. HAIRER, "An Introduction to Stochastic PDEs", `https://arxiv.org/abs/0907.4178`, 2009.

[25] ———, "Solving the KPZ equation", *Ann. Math.* **178** (2013), no. 2, p. 559-664.

[26] ———, "Singular Stochastic PDES", `https://arxiv.org/abs/1403.6353`, 2014.

[27] ———, "A theory of regularity structures", *Invent. Math.* **198** (2014), no. 2, p. 269-504.

[28] ———, "Introduction to regularity structures", *Braz. J. Probab. Stat.* **29** (2015), no. 2, p. 175-210.

[29] M. HAIRER & C. LABBÉ, "Multiplicative stochastic heat equations on the whole space", `https://arxiv.org/abs/1504.07162`, to appear in *J. Eur. Math. Soc.*, 2015.

[30] M. HAIRER & É. PARDOUX, "A Wong-Zakai theorem for Stochastic PDEs", *J. Math. Soc. Japan* **67** (2015), no. 4, p. 1551-1604.

[31] M. HAIRER, M. D. RYSER & H. WEBER, "Triviality of the 2D stochastic Allen-Cahn equation", *Electron. J. Probab* **17** (2012), no. 39, p. 1-14.

[32] M. KARDAR, G. PARISI & Y.-C. ZHANG, "Dynamic Scaling of Growing Interfaces", *Phys. Rev. Lett.* **56** (1986), no. 9, p. 889-892.

[33] P. E. KLOEDEN & E. PLATEN, *Numerical solution of stochastic differential equations*, Applications of Mathematics, vol. 23, Springer, 1992, xxxvi+632 pages.

[34] N. V. KRYLOV, *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, Graduate Studies in Mathematics, vol. 12, American Mathematical Society, 1996, xii+164 pages.

[35] A. KUPIAINEN, "Renormalization Group and Stochastic PDEs", *Ann. Henri Poincaré* **17** (2015), no. 3, p. 1-39.

[36] T. J. LYONS, "Differential equations driven by rough signals", *Rev. Mat. Iberoam.* **14** (1998), no. 2, p. 215-310.

[37] J.-C. MOURRAT & H. WEBER, "Convergence of the two-dimensional dynamic Ising-Kac model to Φ_2^4", *Commun. Pure Appl. Math.* **70** (2017), no. 4, p. 717-812.

[38] ———, "Global well-posedness of the dynamic Φ^4 model in the plane", *Ann. Probab.* **45** (2017), no. 4, p. 2398-2476.

[39] A. R. NAHMOD & G. STAFFILANI, "Almost sure well-posedness for the periodic 3D quintic nonlinear Schrödinger equation below the energy space", *J. Eur. Math. Soc. (JEMS)* **17** (2015), no. 7, p. 1687-1759.

[40] D. Nualart, *The Malliavin calculus and related topics*, 2nd ed., Probability and Its Applications, Springer, 2006, xiv+382 pages.

[41] R. Peierls, "On Ising's model of ferromagnetism", *Proc. Camb. Philos. Soc.* **32** (1936), p. 477-481.

[42] C. Prévôt & M. Röckner, *A concise course on stochastic partial differential equations*, Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007, vi+144 pages.

[43] J. Quastel & H. Spohn, "The one-dimensional KPZ equation and its universality class", *J. Stat. Phys.* **160** (2015), no. 4, p. 965-984.

[44] T. Sasamoto & H. Spohn, "One-dimensional Kardar-Parisi-Zhang equation: an exact solution and its universality", *Phys. Rev. Lett.* **104** (2010), no. 23, Article ID 230602.

[45] K. A. Takeuchi & M. Sano, "Universal Fluctuations of Growing Interfaces: Evidence in Turbulent Liquid Crystals", *Phys. Rev. Lett.* **104** (2010), no. 23, Article ID 230601.

[46] E. Wong & M. Zakai, "On the convergence of ordinary integrals to stochastic integrals", *Ann. Math. Stat.* **36** (1965), p. 1560-1564.

[47] ———, "On the relation between ordinary and stochastic differential equations", *Int. J. Eng. Sci.* **3** (1965), p. 213-229.