Paracontrolled distributions and singular SPDEs

by Nicolas Perkowski

March 30, 2020

Abstract

These are the lecture notes for a course at the Max Planck Institute for Mathematics in the Sciences in the winter semester 2018/19.

1 Introduction

Consider a domain $\mathcal{O} \subset \mathbb{R}^d$ and the potential

$$V(\phi) = \int_{\mathcal{O}} |\nabla \phi(x)|^2 \mathrm{d}x + \int_{\mathcal{O}} \left(\frac{1}{2} \phi(x)^4 - \phi(x)^2 \right) \mathrm{d}x.$$

We are interested in the measure

$$\mu(\mathrm{d}\phi) = \frac{1}{Z} \exp(-V(\phi)) \prod_{x \in \mathcal{O}} \mathrm{d}\phi(x), \tag{1.1}$$

where $\prod_{x \in \mathcal{O}} \mathrm{d}\phi(x)$ should be interpreted as the "Lebesgue measure" on the space of real-valued functions on $\mathcal{O}$ (which of course does not exist). This measure is called the Φ^4_d measure (4 because of the power $\phi(x)^4$ in the potential, d because of the dimension of the base space), and it is important in quantum field theory, see for example [15].

It is also a sort of continuum version of the Ising model: Recall that the Ising model on the finite lattice $\Lambda \subset \mathbb{Z}^d$ is the measure on $\{-1, 1\}^{\Lambda}$ formally given by the density (the value of Z will change throughout, it is always taken as the positive constant for which the total mass of the measure becomes 1)

$$\frac{1}{Z} \exp\left(\sum_{i,j:|i-j|=1} \sigma_i \sigma_j \right) \prod_{i \in \Lambda} \mathrm{d}\sigma_i = \frac{1}{Z} \exp\left(-\frac{1}{2} \sum_{i,j:|i-j|=1} (\sigma_i - \sigma_j)^2 \right) \prod_{i \in \Lambda} \mathrm{d}\sigma_i,$$

where $\prod_{i \in \Lambda} \mathrm{d}\sigma_i$ should be interpreted now as the counting measure on $\{-1, 1\}^{\Lambda}$. This is a model for a ferromagnet, and the spin σ_i is the magnetization at point i (positive or negative). The Ising model has the tendency of favoring configurations $\sigma \in \{-1, 1\}^{\Lambda}$ for which neighboring spins are aligned, i.e. for which $\sigma_i = \sigma_j$ if $|i - j| = 1$. On the other hand the counting measure introduces randomness.

In our case the potential V is also small for those ϕ with small gradient (i.e. for which $\int |\nabla \phi(x)|^2 \mathrm{d}x$ is small), and the "Lebesgue measure" introduces randomness. But now there is an additional contribution to the potential, which can be motivated as follows: For the Ising model the spins σ_i only take the values ± 1, while the functions ϕ take arbitrary values in $\mathbb{R}$ and shifting ϕ by a constant does not change the value of $\int |\nabla \phi(x)|^2 \mathrm{d}x$. So the second contribution to the potential in a way anchors ϕ around the values ± 1. Indeed, the double well potential is minimized exactly in the points ± 1:

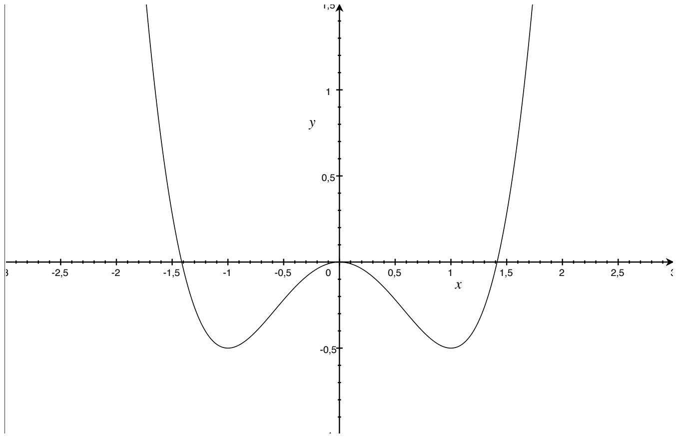


Figure 1.1.

Of course all of the discussion above is purely formal because the Lebesgue measure $\prod_{x\in\mathcal{O}} \mathrm{d}\phi(x)$ does not exist. An attempt to make it rigorous could be to note that, at least if we ignore boundary terms from the integration by parts (this is for example ok if $\mathcal{O}=\mathbb{T}^d=(\mathbb{R}/\mathbb{Z})^d$ is the torus or if $\mathcal{O}=\mathbb{R}^d$ and we consider functions with sufficient decay at ∞)

$$\exp\left(-\int |\nabla\phi(x)|^2 \mathrm{d}x\right) = \exp\left(\int \phi(x)\Delta\phi(x)\mathrm{d}x\right) = \exp\left(-\int \phi(x)(-\Delta)\phi(x)\mathrm{d}x\right).$$

The operator $-\Delta$ is symmetric and positive, so this expression looks very much like a Gaussian density: If C is a symmetric and strictly positive definite matrix on $\mathbb{R}^n$, then the measure

$$\pi(\mathrm{d}x) = \frac{1}{Z}\exp(-\langle x, Cx\rangle)\prod_i \mathrm{d}x_i,$$

where $\langle\cdot,\cdot\rangle$ is the inner product on $\mathbb{R}^n$, is centered Gaussian with covariance $\frac{1}{2}C^{-1}$, i.e. for all a, b we have

$$\int \langle x, a\rangle\langle x, b\rangle \pi(\mathrm{d}x) = \frac{1}{2}\langle a, C^{-1}b\rangle.$$

So in our case we would expect

$$\nu(\mathrm{d}\phi) = \frac{1}{Z}\exp\left(-\int |\nabla\phi(x)|^2\mathrm{d}x\right)\prod_{x\in\mathcal{O}} \mathrm{d}\phi(x)$$

to be a Gaussian measure with

$$\int \langle\phi, f\rangle\langle\phi, g\rangle\nu(\mathrm{d}\phi) = \frac{1}{2}\langle f, (-\Delta)^{-1}g\rangle = \frac{1}{2}\langle(-\Delta)^{-1/2}f, (-\Delta)^{-1/2}g\rangle,$$

where now $\langle\cdot,\cdot\rangle$ is the inner product on $L^2(\mathcal{O})$. Modulo technicalities (it is for example not always possible to invert $-\Delta$, and it may be better to consider $1-\Delta$ instead), such a measure exists, and it is called the *Gaussian free field*.

This brings us a good step closer towards making sense of the Φ_d^4 measure that we formally defined in (1.1). A "rigorous candidate" seems to be

$$\mu(\mathrm{d}\phi) = \frac{1}{Z}\exp\left(-\int\left(\frac{1}{2}\phi(x)^4 - \phi(x)^2\right)\mathrm{d}x\right)\nu(\mathrm{d}\phi).$$

And indeed this measure is well defined if $d=1$. Unfortunately for $d>1$ the measure ν is only supported on generalized functions (Schwartz distributions), and $\nu(L^p(\mathcal{O}))=0$ for all $p\in[1,\infty]$.

Exercise. (Suggested by Scott Smith):

Consider the "massive Gaussian free field on $\mathbb{R}^d$", i.e. the centered Gaussian process $(\eta(\varphi))_{\varphi\in\mathscr{S}}$ with covariance $\mathbb{E}[\eta(\varphi)\eta(\psi)]=\frac{1}{2}\langle\varphi,(m-\Delta)^{-1}\psi\rangle_{L^2(\mathbb{R}^d)}$ for mass $m>0$. The covariance is well defined, and therefore η exists. Show that if $(\rho_n)\subset C_c^\infty$ is an approximation of the identity with $\mathrm{supp}(\rho_n)\subset B(0,1/n)$, then $\mathbb{E}[\eta(\rho_n(x-\cdot))^2]\to\infty$ if and only if $d>1$. This shows that the formal expression $\eta(x)=\eta(\delta(x-\cdot))$ for the Dirac delta does not exist (a sequence of Gaussian random variables with diverging variance cannot even converge in distribution).

Solution. We can represent

$$(m-\Delta)^{-1}f(x)=\int_0^\infty e^{-mt}\int_{\mathbb{R}^d}p(t,x-y)f(y)\mathrm{d}y\mathrm{d}t$$

for the centered Gaussian density with covariance $2tI$, i.e.

$$p(t,x)=(4\pi t)^{-d/2}\exp\left(-\frac{|x|^2}{4t}\right).$$

Thus we get

$$\begin{aligned}&\langle\rho_n(x-\cdot),((m-\Delta)^{-1}\rho_n(x-\cdot))\rangle\\&=\int_{\mathbb{R}^d}\rho_n(x-z)\int_0^\infty e^{-mt}\int_{\mathbb{R}^d}p(t,x-z-y)\rho_n(y)\mathrm{d}y\mathrm{d}t\mathrm{d}z\\&=\int_{\mathbb{R}^d}\rho_n(z)\int_0^\infty e^{-mt}\int_{\mathbb{R}^d}p(t,z-y)\rho_n(y)\mathrm{d}y\mathrm{d}t\mathrm{d}z.\end{aligned}$$

We can restrict the integration to $|z|,|y|\leqslant 1/n$ because otherwise ρ_n vanishes. But then $|z-y|\leqslant 2/n$ and thus

$$p(t,z-y)\gtrsim t^{-d/2}\exp\left(-\frac{4}{4tn^2}\right),$$

i.e.

$$\begin{aligned}\langle\rho_n(x-\cdot),((m-\Delta)^{-1}\rho_n(x-\cdot))\rangle&\gtrsim\int_{\mathbb{R}^d}\rho_n(z)\int_0^\infty e^{-mt}\int_{\mathbb{R}^d}t^{-d/2}\exp\left(-\frac{1}{tn^2}\right)\rho_n(y)\mathrm{d}y\mathrm{d}t\mathrm{d}z\\&=\int_0^\infty e^{-mt}t^{-d/2}\exp\left(-\frac{1}{tn^2}\right)\mathrm{d}t\\&\geqslant\int_{n^{-2}}^\infty e^{-mt}t^{-d/2}\exp(-1)\mathrm{d}t\gtrsim\int_{n^{-2}}^1 t^{-d/2}\mathrm{d}t,\end{aligned}$$

which diverges as $n\to\infty$ if $d\geqslant 2$. On the other hand the uniform bound $p(t,x)\lesssim t^{-1/2}$ shows that for $d=1$ the variance converges.

Since ν is only supported on distributions, it is unclear how to interpret the expression

$$\int\left(\frac{1}{2}\phi(x)^4-\phi(x)^2\right)\mathrm{d}x$$

for typical $\phi\in\mathrm{supp}(\nu)$, and we are still stuck with the construction of ϕ. This problem can be overcome for $d<4$ with the help of renormalization group techniques from quantum field theory [15], and indeed one has to suitably renormalize the potential $\frac{1}{2}\phi(x)^4-\phi(x)^2$ by subtracting infinite counterterms. Let us stress that d is the dimension of space-time so we would like to take $d=4$ for physical applications. Unfortunately this case is just out of reach of the existing theories!

Here we will follow a different route towards constructing μ. Consider a smooth potential function $V:\mathbb{R}^n\to\mathbb{R}$, such that

$$\int \exp(-V(x))\prod_{i=1}^n \mathrm{d}x_i<\infty,$$

for example $V(x)=x^2$. Then $\mu(\mathrm{d}x)=\frac{1}{Z}\exp(V(x))\prod_i \mathrm{d}x_i$ defines a probability measure, and it is a classical result that the following n-dimensional SDE, sometimes called *overdamped Langevin dynamics*, has μ as a reversible measure:

$$\mathrm{d}X_t=-\frac{1}{2}\nabla V(X_t)\mathrm{d}t+\mathrm{d}W_t.$$

To see this note that the generator of X is $\mathcal{L}=-\frac{1}{2}\nabla V\cdot\nabla+\frac{1}{2}\Delta$ and show that $\mathcal{L}$ is self-adjoint in $L^2(\mu)$ (which is not a proof, but at least explains intuitively why X should be reversible). In coordinates and formal notation we can also write

$$\partial_t X(t,i)=-\frac{1}{2}\partial_i V(X(t,\cdot))+\xi(t,i),$$

where we set $\xi(f)=\sum_i\int_{\mathbb{R}_+}f(t,i)\mathrm{d}W_t^i$ and thus we can interpret ξ is a "space-time white noise on $\mathbb{R}_+\times\{1,\ldots,n\}$", i.e. a centered Gaussian process indexed by $L^2(\mathbb{R}_+\times\{1,\ldots,n\})$ (with product measure of Lebesgue and counting measure) such that

$$\mathbb{E}[\xi(f)\xi(g)]=\langle f,g\rangle=\sum_{i=1}^n\int_{\mathbb{R}_+}f(t,i)g(t,i)\mathrm{d}t.$$

This provides a way of constructing μ: simply let the dynamics of X run and apply the ergodic theorem (this still requires ergodicity of X, which we take for granted here) to get

$$\int f(x)\mu(\mathrm{d}x)=\lim_{T\to\infty}\frac{1}{T}\int_0^T f(X_s)\mathrm{d}s.$$

If we apply this philosophy to the Φ_d^4 measure, we would formally get a dynamic process that satisfies

$$\partial_t\phi(t,x)=-\frac{1}{2}\delta_x V(\phi(t,\cdot))+\xi(t,x)=\Delta\phi(t,x)-\phi(t,x)^3+\phi(t,x)+\xi(t,x), \qquad (1.2)$$

where ξ is now a space-time white noise on $\mathbb{R}_+\times\mathcal{O}$, i.e. ξ is centered Gaussian and

$$\mathbb{E}[\xi(f)\xi(g)]=\langle f,g\rangle=\int_{\mathbb{R}_+\times\mathcal{O}}f(t,x)g(t,x)\mathrm{d}t\mathrm{d}x$$

for all $f,g\in L^2(\mathbb{R}_+\times\mathcal{O})$. The functional derivative $\delta_x V(\phi)$ is a bit subtle and can be computed by taking first the functional derivative in the direction of some function ψ, and then formally setting $\psi=\delta(x-\cdot)$ for the Dirac delta in x.

The equation (1.2) is called the Φ_d^4 equation, or the *stochastic quantization equation* for the Φ_d^4 measure [33]. We should not expect a free lunch though, and we cannot overcome the essential difficulties in the construction of the Φ_d^4 measure by simply writing the problem in the language of stochastic differential equations. In fact (1.2) is still badly ill posed, because as we will see the solution $\phi(t,\cdot)$ is at fixed times t only a distribution in the space variable, and therefore we have again a problem with the nonlinearity, this time given by ϕ^3. In fact at least for $d=3$ the solution theory for the Φ_3^4 *equation* (1.2) is much younger than the construction of the Φ_d^4 measure, and it was one of the first big successes of Hairer's regularity structures [22] respectively paracontrolled distributions [16, 10]; the Φ_2^4 equation had been previously solved with Dirichlet forms [1] or with the "Da Prato-Debussche method" [13].

Regularity structures and paracontrolled distributions take a new point of view on regularity, and point out that from the right perspective the solution ϕ to (1.2) is actually quite "smooth", and working with this new notion of regularity allows us to derive good estimates for ϕ and to solve the equation. More precisely, ϕ can be interpreted as a perturbation of a Gaussian process X, where the perturbation becomes negligible on small scales. Indeed, consider the rescaling operation $\mathcal{S}_\lambda\phi(t,x)=\phi(\lambda^2 t,\lambda x)$ and set $\phi_\lambda=\lambda^{d/2-1}\mathcal{S}_\lambda\phi$. Then

$$\begin{aligned}\partial_t\phi_\lambda(t,x) =&\lambda^{2+d/2-1}\mathcal{S}_\lambda(\Delta\phi-\phi^3+\phi+\xi)(t,x)\\ =&(\Delta\phi_\lambda-\lambda^{4-d}\phi_\lambda^3+\lambda^2\phi_\lambda+\lambda^{1+d/2}\mathcal{S}_\lambda\xi)(t,x).\end{aligned}$$

Exercise. Let ξ be a space-time white noise and set $\xi_\lambda:=\lambda^{1+d/2}\mathcal{S}_\lambda\xi$, interpreted rigorously as $\xi_\lambda(f)=\lambda^{1+d/2}\xi(\lambda^{-2-d}\mathcal{S}_{\lambda^{-1}}f)$. Show that ξ_λ is a new space-time white noise.

Hint: It suffices to show that ξ_λ is centered Gaussian with the right covariance.

By the exercise we have

$$\partial_t\phi_\lambda=\Delta\phi_\lambda-\lambda^{4-d}\phi_\lambda^3+\lambda^2\phi_\lambda+\xi_\lambda, \tag{1.3}$$

and since we are interested in small scales we want to take $\lambda\ll 1$. If $d<4$ the nonlinearity formally drops out and we remain with the equation

$$\partial_t\phi=\Delta\phi+\xi,$$

whose solution we will denote by X. Note that X is Gaussian because it depends linearly on the Gaussian process ξ and because linear functionals of Gaussian processes are Gaussian. There is a small subtlety here because actually the white noise ξ_λ should depend on λ, but let us ignore this. By the above discussion we expect that for $d<4$ and on small scales ϕ should resemble X, and as a first step we can then try to make sense of X^3, which is indeed possible (at least modulo renormalization). Since the difficulty we have in making sense of ϕ^3 comes from its irregularity, and irregularity is a small scale property and on small scales ϕ is essentially X, we can thus hope to make sense of ϕ^3. Regularity structures and paracontrolled distributions provide different tools for implementing this intution. Regularity structures are based on controlling increments, while paracontrolled distributions are based on Fourier descriptions of regularity.

Paracontrolled distributions are less general than regularity structures. To get a feeling of the range of applicability of regularity structures, note that above we ignored the case $d\geqslant 4$ until now. Looking at (1.3), we see that for $d=4$ the nonlinear term does not change with λ, and therefore this case is called *(locally) critical* in the language of Hairer [22]. For $d>4$ things are even worse and the nonlinearity blows up as we zoom into the small scales; this case is called *supercritical*. Accordingly the case $d<4$ is called *subcritical*. Regularity structures provide a general theory to treat subcritical equations, and by now there exist black box type results that give a (local-in-time) renormalization and existence-uniqueness theory for large classes of subcritical equations [22, 8, 11, 7].

The following intuitive description might or might not be useful: Paracontrolled distributions are at the moment restricted to equations where the scaling exponent α in the factor λ^α that we pick up in front of the nonlinearity by a scaling argument as above is bounded from below by some $\alpha\geqslant\alpha_0>0$ (with α_0 depending on the specific equation), while in regularity structures it suffices if $\alpha>0$. On the other hand paracontrolled distributions are based on classical tools and function spaces from PDE theory, which might make them easier to learn and easier to implement in some applications. We may also consider regularity structures as generalized Taylor expansions, and then the restriction of paracontrolled distributions is that we can only deal with first order expansions while regularity structures allow expansions of arbitrary order; see however [4, 27] for some progress towards generalizations of paracontrolled distributions.

In the following we will learn the basics of paracontrolled distributions on the guiding example of the Φ^4_d equations ($d=2,3$) and then see some applications to these and other equations.

Bibliography: The discussion above is partially inspired by the nice survey papers [12, 30].

2 Besov spaces and paraproducts

Here we introduce some basic tools from harmonic analysis to measure the regularity of tempered distributions and to control the product of a distribution and a function in the case of compatible regularities. An excellent reference is [3], where much of the material here is taken from.

Recap: Tempered distributions

We will work with tempered distributions on $\mathbb{R}^d$. Recall that the *Schwartz functions* are

$$\mathscr{S} = \{\varphi \in C^{\infty}(\mathbb{R}^d, \mathbb{C}) : \|\varphi\|_{k,\mathscr{S}} < \infty \forall k \in \mathbb{N}_0\},$$

where

$$\|\varphi\|_{k,\mathscr{S}} = \sup_{|\mu| \leqslant k} \|(1+|\cdot|^k)\partial^{\mu}\varphi\|_{L^{\infty}}.$$

The *Schwartz distributions* are the linear maps $u: \mathscr{S} \to \mathbb{C}$ which satisfy

$$|u(\varphi)| \leqslant C\|\varphi\|_{k,\mathscr{S}}$$

for some $C>0$ and $k \in \mathbb{N}_0$. In that case we write $u \in \mathscr{S}'$.

Example. Clearly $L^p = L^p(\mathbb{R}^d) \subset \mathscr{S}'$ for all $p \in [1,\infty]$ if we identify $u \in L^p$ with the map $\varphi \mapsto \int_{\mathbb{R}^d} u(x)\varphi(x)\mathrm{d}x$, and more generally the space of finite signed measures on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ is contained in $\mathscr{S}'$. Another example of a tempered distribution is $\varphi \mapsto \partial^{\mu}\varphi(x)$ for $\mu \in \mathbb{N}_0^d$ and $x \in \mathbb{R}^d$. A continuous function u is in $\mathscr{S}'$ if and only if it has at most polynomial growth at infinity.

Many maps on $\mathscr{S}'$ can be defined by duality: Let $A: \mathscr{S} \to \mathscr{S}$ be such that there exists a linear map $A^t: \mathscr{S} \to \mathscr{S}$ which satisfies for all $\varphi, \psi \in \mathscr{S}$

$$\int_{\mathbb{R}^d} (A\varphi)(x)\psi(x)\mathrm{d}x = \int_{\mathbb{R}^d} \varphi(x)({}^tA\psi)(x)\mathrm{d}x$$

and also for all $m \in \mathbb{N}_0$ there exist $k_m \in \mathbb{N}_0$, $C_m > 0$ with $\|{}^tA\varphi\|_{m,\mathscr{S}} \leqslant C_m\|\varphi\|_{k_m,\mathscr{S}}$. Then we define for $u \in \mathscr{S}'$

$$(Au)(\varphi) := u({}^tA\varphi).$$

Example.

i. For $\mu \in \mathbb{N}_0^d$ and $A = \partial^{\mu}$ we have ${}^tA = (-1)^{|\mu|}\partial^{\mu}$.

ii. For $f \in C^{\infty}$ with all partial derivatives of at most polynomial growth and $A\varphi = f\varphi$ we have ${}^tA = A$.

iii. For the Fourier transform

$$A\varphi(z) = \mathscr{F}\varphi(z) := \hat{\varphi}(z) := \int_{\mathbb{R}^d} e^{-2\pi i x z}\varphi(x)\mathrm{d}x$$

we have ${}^tA = A$.

iv. For the inverse Fourier transform

$$A\varphi(z) = \mathscr{F}^{-1}\varphi(z) = \int_{\mathbb{R}^d} e^{2\pi i x z}\varphi(x)\mathrm{d}x$$

we have ${}^tA = A$.

v. For $\chi \in \mathscr{S}$ and the convolution

$$A\varphi = \chi * \varphi = \int_{\mathbb{R}^d} \chi(\cdot - y)\varphi(y)\mathrm{d}y$$

we have ${}^tA\varphi = (\chi(-\cdot)) * \varphi$. In this case one can show that $\chi * u \in C^\infty \cap \mathscr{S}$ for all $u \in \mathscr{S}'$.

The main reason for considering test functions in $\mathscr{S}$ rather than in the simpler space C_c^∞ is that for elements of $\mathscr{S}'$ we can define the Fourier transform by duality, which is not true for elements of $(C_c^\infty)'$ because C_c^∞ is not closed under Fourier transformation.

Example. Let $u \in \mathscr{S}'$ and $\varphi, \psi \in \mathscr{S}$. The following relations will be used all the time:

- $\mathscr{F}^{-1}\mathscr{F}u = \mathscr{F}\mathscr{F}^{-1}u = u$ for all $u \in \mathscr{S}'$;
- Parseval's identity:

  $$\int_{\mathbb{R}^d} \varphi(x)\psi(x)^*\mathrm{d}x = \int_{\mathbb{R}^d} \hat{\varphi}(x)\hat{\psi}(x)^*\mathrm{d}x$$

  and by extension $u(\varphi^*) = \hat{u}(\hat{\varphi}^*)$;
- $\widehat{\partial^\mu u} = (-2\pi i x)^\mu u$;
- $\widehat{u\varphi} = \hat{u} * \hat{\varphi}$;
- $\widehat{u * \varphi} = \hat{u}\hat{\varphi}$;
- $\mathrm{supp}(\varphi * \psi) \subset \overline{\mathrm{supp}(\varphi) + \mathrm{supp}(\psi)} = \overline{\{x + y : x \in \mathrm{supp}(\varphi), y \in \mathrm{supp}(\psi)\}}$.

Recall also the following fundamental inequalities that we will constantly use:

Lemma. (Hölder's inequality) *Let $p, q, r \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. Then*

$$\|uv\|_{L^r} \leqslant \|u\|_{L^p}\|v\|_{L^q}.$$

Lemma. (Young's inequality for convolutions) *Let $p, q, r \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$. Then*

$$\|u * v\|_{L^r} \leqslant \|u\|_{L^p}\|v\|_{L^q}.$$

We will refer to this as "Young's inequality", omitting "for convolutions". If we need Young's inequality for products we will distinguish this by explicitly mentioning "for products".

2.1 Besov spaces

The main difficulty we encountered in the introduction was that we had to multiply distributions. Note that for $u\in\mathscr{S}'$ and $\varphi\in C^\infty$ with partial derivatives of polynomial growth we can define the product $u\varphi$ by duality. But if φ is a non-smooth function or even a distribution, then the duality approach breaks down completely and we need other arguments. If say $u,\varphi\in L^2$, then $u\varphi\in L^1$ is of course also well defined, so we might hope to find another approach that makes sense of $u\varphi$ for all $u,\varphi\in\mathscr{S}'$. But this is not possible:

Example 2.1. (Schwartz) In $d=1$ we can turn $\frac{1}{x}$ into a tempered distribution via the so called *principal value*. The details of that construction are not important for us, but with it we obtain for the Dirac delta δ (i.e. $\delta(\varphi)=\varphi(0)$)

$$0=(\delta\times x)=(\delta\times x)\times\frac{1}{x}\neq\delta\times\left(x\times\frac{1}{x}\right)=\delta\times 1=\delta.$$

This example shows that a general extension of the product $u\varphi$ to distributions or non-smooth functions φ is not possible. Our way of overcoming this difficulty is to restrict both u and φ to suitable subspaces of $\mathscr{S}'$. Of course the example $u,\varphi\in L^2$ from above works, but we are interested in situations where at least one of u,φ is a bona fide distribution and not a function. The simplest solution is to require u and φ to have compatible regularity. For that purpose we need to introduce regularities on distribution spaces.

To measure the regularity of distributions we first note that if $u\in\mathscr{S}'$ with $\operatorname{supp}(\hat{u})\subset K$, where $K\subset\mathbb{R}^d$ is a compact set, then there exists $\varphi\in C_c^\infty$ with $\varphi|_K\equiv 1$ and therefore

$$u=\mathscr{F}^{-1}(\hat{u})=\mathscr{F}^{-1}(\varphi\hat{u})=(\mathscr{F}^{-1}\varphi)\ast u.$$

Since $\mathscr{F}^{-1}\varphi\in\mathscr{S}$ we get $u\in C^\infty$. Moreover, if $|z|\simeq\lambda$ for all $z\in\operatorname{supp}(\hat{u})$, then essentially we can picture u as a sine-function with period $(2\pi\lambda)^{-1}$. So if λ is small, u is smooth and oscillating very slowly but if $\lambda\gg 1$, then u is very wild. This suggests that smooth functions have some decay in their Fourier transform. It turns out that measuring the size of single Fourier coefficients does not provide enough information and instead it is more useful to group the different frequency ranges into blocks. More precisely, we would like to decompose

$$u=\mathscr{F}^{-1}(\hat{u})=\mathscr{F}^{-1}(\mathbb{I}_{[0,1)}(|\cdot|)\hat{u})+\sum_{j\geqslant 0}^{\infty}\mathscr{F}^{-1}(\mathbb{I}_{[2^j,2^{j+1})}(|\cdot|)\hat{u})=\Delta_{-1}u+\sum_{j=0}^{\infty}\Delta_j u.$$

Then $\Delta_j u$ is the projection of u onto its frequencies of order $\sim 2^j$. Since frequencies of order 2^j correspond to spatial scales of order 2^{-j}, the sum $\sum_{i\leqslant j}\Delta_i u$ provides a description of u up to the spatial scale 2^{-j}. For a smooth function u this should already give a very accurate picture of u, and therefore we expect $\Delta_j u$ to rapidly decay as $j\to\infty$. Measuring the strength of that decay will provide us with a notion of regularity. But there are two problems with the above formal decomposition: First of all it is not even well defined, because we are multiplying $\hat{u}\in\mathscr{S}'$ with non-smooth indicator functions. And even in situations where we can make sense of this product, it still turns out that the operation $u\mapsto\Delta_j u$ is quite badly behaved. For example, we would like to estimate $\|\Delta_j u\|_{L^p}\leqslant\|\mathscr{F}^{-1}(\mathbb{I}_{[2^j,2^{j+1})}(|\cdot|))\|_{L^1}\|u\|_{L^p}$ via Young's inequality, but the L^1 norm on the right hand side is infinite because while $\mathscr{F}^{-1}(\mathbb{I}_{[2^j,2^{j+1})}(|\cdot|))\in C^\infty$, it is not in L^1.

Definition 2.2. *For $u \in \mathscr{S}'$ and $j \geqslant -1$ we define the Littlewood-Paley blocks of u as*

$$\Delta_j u = \mathscr{F}^{-1}(\rho_j \hat{u}),$$

where $(\rho_j) \subset C_c^\infty$ is a (smooth, dyadic) partition of unity with

$$\rho_{-1} \simeq \mathbb{I}_{[0,1)}(|\cdot|), \qquad \rho_j \simeq \mathbb{I}_{[2^j, 2^{j+1})}(|\cdot|), j \geqslant 0.$$

Here we choose the ρ as radial functions that sum up to 1 everywhere ("unity") and such that the support of ρ_j only overlaps with the supports of ρ_{j-1} and ρ_{j+1} ("smooth partition"). We also take the ρ such that $\rho_j = \rho_0(2^{-j}\cdot)$ for $j \geqslant 0$. We write $K_j = \mathscr{F}^{-1}\rho_j$, so that $\Delta_j u = K_j * u$. We also use the notation

$$\Delta_{\leqslant j} u = \sum_{i \leqslant j} \Delta_i u, \qquad \Delta_{<j} u = \sum_{i<j} \Delta_i u, \qquad \Delta_{\geqslant j} u = \sum_{i \geqslant j} \Delta_i u, \qquad \Delta_{>j} u = \sum_{i>j} \Delta_i u,$$

$$K_{\leqslant j} = \sum_{i \leqslant j} K_i, \qquad K_{<j} = \sum_{i<j} K_i.$$

The kernels $K_j, K_{<j}, K_{\leqslant j}$ are all bounded in L^1, uniformly in j. Moreover, for $j \geqslant 0$ we have the scaling relation $K_j = 2^{jd} K_0(2^j \cdot)$.

Note that no nice kernel exists for $\Delta_{>j}$, because we would need to take $K_{>j} = \delta - K_{\leqslant j}$, where δ is the Dirac delta and $K_{\leqslant j}$ is a smooth kernel.

It is easy to see that $u = \sum_{j \geqslant -1} \Delta_j u = \lim_{j \to \infty} \Delta_{\leqslant j} u$ for all $u \in \mathscr{S}'$ and with Young's inequality we get uniformly in j:

$$\|\Delta_j u\|_{L^p} \leqslant \|K_j\|_{L^1} \|u\|_{L^p} \lesssim \|u\|_{L^p}.$$

As discussed above, we want to describe the regularity of $u \in \mathscr{S}'$ by the decay (or growth) of $\Delta_j u$. For that purpose we first have to decide how to measure the size of $\Delta_j u$. A canonical choice is to consider the L^p norm for $p \in [1, \infty]$.

Definition 2.3. *For $\alpha \in \mathbb{R}$ and $p, q \in [1, \infty]$ the Besov space $B_{p,q}^\alpha$ is defined as*

$$B_{p,q}^\alpha = \left\{ u \in \mathscr{S}' : \|u\|_{B_{p,q}^\alpha} = \|(2^{j\alpha} \|\Delta_j u\|_{L^p})_{j \geqslant -1}\|_{\ell_j^q} < \infty \right\}.$$

So the index p describes the integrability and the index α the "regularity" (i.e. the decay of the blocks). The index q provides some fine-tuning and is not very important: Indeed we have for $q_1 \leqslant q_2$ and $\alpha \in \mathbb{R}$ the inclusions

$$B_{p,q_1}^\alpha \subset B_{p,q_2}^\alpha \subset B_{p,q_1}^{\alpha'}$$

whenever $\alpha' < \alpha$. $B_{p,q}^\alpha$ is a Banach space for all α, p, q. The Gaussian noises that we will consider have the same regularity index α in any Besov space $B_{p,q}^\alpha$, which, as we will see, is a consequence of the comparability of moments for Gaussian random variables. Therefore we mainly work in the easiest setting $p = q = \infty$, for which we introduce a special notation:

$$\mathscr{C}^\alpha = B_{\infty,\infty}^\alpha, \qquad \|\cdot\|_\alpha = \|\cdot\|_{B_{\infty,\infty}^\alpha}.$$

Exercise 2.1. Let δ denote the Dirac delta, $\delta(\varphi) = \varphi(0)$. Show that $\delta_0 \in B_{p,\infty}^{-d(1-1/p)}$ for all p. So when dealing with equations involving the Dirac delta (say as initial condition), it may be advantageous to work in Besov spaces with finite integrability index.

Exercise 2.2. Show that $\|u\|_\alpha \leqslant \|u\|_\beta$ for $\alpha \leqslant \beta$, that $\|u\|_{L^\infty} \lesssim \|u\|_\alpha$ for $\alpha > 0$, that $\|u\|_\alpha \lesssim \|u\|_{L^\infty}$ for $\alpha \leqslant 0$, and that $\|\Delta_{\leqslant j} u\|_{L^\infty} \lesssim 2^{-j\alpha}\|u\|_\alpha$ for $\alpha < 0$ and $\|\Delta_{>j} u\|_{L^\infty} \lesssim 2^{-j\alpha}\|u\|_{\mathscr{C}^\alpha}$ for $\alpha > 0$. We will often use these inequalities without explicitly mentioning it.

If $\alpha \in (0,\infty) \setminus \mathbb{N}$, then $\mathscr{C}^\alpha$ is the space of $\lfloor \alpha \rfloor$ times differentiable functions whose partial derivatives of order $\lfloor \alpha \rfloor$ are $(\alpha - \lfloor \alpha \rfloor)$-Hölder continuous, with norm equivalent to

$$\|u\|_\alpha \simeq \|u\|_{C_b^\alpha} := \sum_{|\mu| \leqslant \lfloor \alpha \rfloor} \|\partial^\mu u\|_{L^\infty} + \sum_{|\mu| = \lfloor \alpha \rfloor} \sup_{x \neq y} \frac{|\partial^\mu u(x) - \partial^\mu u(y)|}{|x-y|^{\alpha - \lfloor \alpha \rfloor}}.$$

But for $k \in \mathbb{N}$ the space $\mathscr{C}^k$ is strictly larger than C_b^k, the space of k times continuously differentiable functions with bounded partial derivatives of all order. We will see the equivalence for $\alpha \in (0,1)$ as an exercise below, but before that we need the following Bernstein inequality, which is very useful when dealing with functions with compactly supported Fourier transform.

Lemma 2.4. (Bernstein inequality) *Let $\mathscr{B}$ be a ball, $k \in \mathbb{N}_0$, and $1 \leqslant p \leqslant q \leqslant \infty$. There exists a constant $C > 0$ that depends only on k, $\mathscr{B}$, p and q such that for all $\lambda > 0$ and $u \in L^p$ with* $\operatorname{supp}(\mathscr{F}u) \subseteq \lambda \mathscr{B}$ *we have*

$$\max_{\mu \in \mathbb{N}^d : |\mu| = k} \|\partial^\mu u\|_{L^q} \leqslant C \lambda^{k + d\left(\frac{1}{p} - \frac{1}{q}\right)} \|u\|_{L^p}.$$

Proof. Let $\psi \in C_c^\infty$ with $\psi \equiv 1$ on $\mathscr{B}$ and write $\psi_\lambda(x) = \psi(\lambda^{-1} x)$. Young's inequality gives

$$\|\partial^\mu u\|_{L^q} = \|\partial^\mu \mathscr{F}^{-1}(\psi_\lambda \hat{u})\|_{L^q} = \|(\partial^\mu \mathscr{F}^{-1}(\psi_\lambda)) * u\|_{L^q} \leqslant \|\partial^\mu \mathscr{F}^{-1}(\psi_\lambda)\|_{L^r} \|u\|_{L^p},$$

where $1 + \frac{1}{q} = \frac{1}{p} + \frac{1}{r}$. Now it suffices to note that

$$\begin{aligned}
\|\partial^\mu \mathscr{F}^{-1}(\psi_\lambda)\|_{L^r} &= \left(\int_{\mathbb{R}^d} |\partial^\mu (\lambda^d \mathscr{F}^{-1}\psi(\lambda x))|^r \mathrm{d}x \right)^{1/r} \\
&= \left(\lambda^{(|\mu|+d)r} \int_{\mathbb{R}^d} |(\partial^\mu \mathscr{F}^{-1}\psi)(\lambda x)|^r \mathrm{d}x \right)^{1/r} \\
&= \left(\lambda^{(|\mu|+d)r - d} \int_{\mathbb{R}^d} |\partial^\mu \mathscr{F}^{-1}\psi(x)|^r \mathrm{d}x \right)^{1/r} \\
&= \lambda^{|\mu| + d\left(1 - \frac{1}{r}\right)} \|\partial^\mu \mathscr{F}^{-1}\psi\|_{L^r}.
\end{aligned}$$

The claim follows by plugging in the equality $1 - \frac{1}{r} = \frac{1}{p} - \frac{1}{q}$. □

It follows immediately that for $\alpha \in \mathbb{R}$, $u \in \mathscr{C}^\alpha$, and $\mu \in \mathbb{N}_0^d$, we have

$$\|\partial^\mu u\|_{\alpha - |\mu|} \lesssim \|u\|_\alpha.$$

We also use the Bernstein inequality to see the claimed equivalence of $\mathscr{C}^\alpha$ and C_b^α, at least in the case $\alpha \in (0,1)$ (the case $\alpha > 1$ is similar but more technical). The next exercise is extremely instructive, because it is based on many arguments we will often encounter later (convergent and divergent geometric series, smuggling in constant terms into integrals against K_j, treating small and large scales separately).

Exercise 2.3. Let $\alpha \in (0,1)$. Then $\mathscr{C}^\alpha = C_b^\alpha$, the space of bounded α-Hölder continuous functions, and

$$\|u\|_\alpha \simeq \|u\|_{C_b^\alpha}.$$

Solution. Let $u \in C_b^\alpha$. Then

$$\|\Delta_{-1} u\|_{L^\infty} \lesssim \|u\|_{L^\infty} \leqslant \|u\|_{C_b^\alpha}$$

and for $j \geqslant 0$ we have $\int K_j(x-y)\mathrm{d}y = 0$ and thus

$$\begin{aligned}|\Delta_j u(x)| &= \left|\int K_j(x-y)u(y)\mathrm{d}y\right| = \left|\int K_j(x-y)(u(y)-u(x))\mathrm{d}y\right| \\ &\leqslant \int |K_j(x-y)| \times |y-x|^\alpha \mathrm{d}y \|u\|_{C_b^\alpha} \\ &= 2^{jd}\int |K_0(2^j(x-y))| \times |2^j(y-x)|^\alpha \mathrm{d}y 2^{-j\alpha}\|u\|_{C_b^\alpha} \\ &= \int |K_0(y)| \times |y|^\alpha \mathrm{d}y 2^{-j\alpha}\|u\|_{C_b^\alpha} \simeq 2^{-j\alpha}\|u\|_{C_b^\alpha},\end{aligned}$$

where in the last step we used that K_0 is a Schwartz function and thus $|K_0(y)| \times |y|^\alpha$ is integrable. Thus, we showed that

$$\|u\|_\alpha = \sup_{j \geqslant -1} 2^{j\alpha}\|\Delta_j u\|_{L^\infty} \lesssim \|u\|_{C_b^\alpha}.$$

Conversely, let $u \in \mathscr{C}^\alpha$. Since $\alpha > 0$ the following geometric series converges and we have

$$\|u\|_{L^\infty} \leqslant \sum_j \|\Delta_j u\|_{L^\infty} \leqslant \sum_j 2^{-j\alpha}\|u\|_\alpha \simeq \|u\|_\alpha.$$

To control $|u(x)-u(y)|$ it suffices to assume that $|x-y| \leqslant 1$. Indeed, for $|x-y| > 1$ we can simply estimate $|u(x)-u(y)| \leqslant 2\|u\|_{L^\infty} \lesssim \|u\|_\alpha \leqslant \|u\|_\alpha |x-y|^\alpha$. So consider x, y with $|x-y| \leqslant 1$ and let $j_0 \geqslant -1$ be such that $2^{-j_0} \simeq |x-y|$. We estimate for $j \leqslant j_0$ with the help of the Bernstein inequality

$$|\Delta_j u(x) - \Delta_j u(y)| \lesssim \sup_{|\mu|=1} \|\partial^\mu \Delta_j u\|_{L^\infty}|x-y| \lesssim 2^{j(1-\alpha)}\|u\|_\alpha |x-y|,$$

and thus, since $1-\alpha > 0$,

$$\sum_{j \leqslant j_0} |\Delta_j u(x) - \Delta_j u(y)| \lesssim \sum_{j \leqslant j_0} \|u\|_\alpha 2^{j(1-\alpha)}|x-y| \lesssim \|u\|_\alpha 2^{j_0(1-\alpha)}|x-y| \simeq \|u\|_\alpha |x-y|^\alpha.$$

For $j > j_0$ we estimate

$$|\Delta_j u(x) - \Delta_j u(y)| \leqslant 2\|\Delta_j u\|_{L^\infty} \lesssim 2^{-j\alpha}\|u\|_\alpha$$

and thus, since $\alpha > 0$,

$$\sum_{j > j_0} |\Delta_j u(x) - \Delta_j u(y)| \lesssim \sum_{j > j_0} 2^{-j\alpha}\|u\|_\alpha \lesssim 2^{-j_0\alpha}\|u\|_\alpha \simeq |x-y|^\alpha \|u\|_\alpha.$$

Another simple application of the Bernstein inequality is the Besov embedding theorem, whose proof we leave as an exercise.

Lemma 2.5. (Besov embedding) *Let $1 \leqslant p_1 \leqslant p_2 \leqslant \infty$ and $1 \leqslant q_1 \leqslant q_2 \leqslant \infty$, and let $\alpha \in \mathbb{R}$. Then for all $u \in \mathscr{S}'$*

$$\|u\|_{B_{p_2,q_2}^{\alpha - d(1/p_1 - 1/p_2)}} \lesssim \|u\|_{B_{p_1,q_1}^\alpha}.$$

The next lemma, a characterization of Besov regularity for functions that can be decomposed into pieces which are localized in Fourier space, will also be immensely useful. Recall that an *annulus* is a set $\mathscr{A} = \{x \in \mathbb{R}^d : a \leqslant |x| \leqslant b\}$ for some $0 < a < b$, and a *ball* is a set $\mathscr{B} = \{x \in \mathbb{R}^d : |x| \leqslant b\}$.

Lemma 2.6.

1. *Let* $\mathscr{A} \subset \mathbb{R}^d$ *be an annulus, let* $\alpha \in \mathbb{R}$*, and let* (u_j) *be a sequence of smooth functions with* $\operatorname{supp}(\mathscr{F} u_j) \subset 2^j \mathscr{A}$ *and such that* $\|u_j\|_{L^\infty} \lesssim 2^{-j\alpha}$ *for all* j*. Then*
$$u = \sum_{j \geqslant -1} u_j \in \mathscr{C}^\alpha \qquad \text{and} \qquad \|u\|_\alpha \lesssim \sup_{j \geqslant -1} \{2^{j\alpha} \|u_j\|_{L^\infty}\}.$$

2. *Let* $\mathscr{B} \subset \mathbb{R}^d$ *be a ball, let* $\alpha > 0$*, and let* (u_j) *be a sequence of smooth functions with* $\operatorname{supp}(\mathscr{F} u_j) \subset 2^j \mathscr{B}$ *and such that* $\|u_j\|_{L^\infty} \lesssim 2^{-j\alpha}$ *for all* j*. Then*
$$u = \sum_{j \geqslant -1} u_j \in \mathscr{C}^\alpha \qquad \text{and} \qquad \|u\|_\alpha \lesssim \sup_{j \geqslant -1} \{2^{j\alpha} \|u_j\|_{L^\infty}\}.$$

Proof. If $\mathscr{F} u_j$ is supported in $2^j \mathscr{A}$, then $\Delta_i u_j \neq 0$ only if $2^i \simeq 2^j$ and therefore
$$\|\Delta_i u\|_{L^\infty} \leqslant \sum_{j : 2^j \simeq 2^i} \|\Delta_i u_j\|_{L^\infty} \leqslant \sup_{k \geqslant -1} \{2^{k\alpha} \|u_k\|_{L^\infty}\} \sum_{j : 2^j \simeq 2^i} 2^{-j\alpha} \simeq \sup_{k \geqslant -1} \{2^{k\alpha} \|u_k\|_{L^\infty}\} \, 2^{-i\alpha}.$$
If $\mathscr{F} u_j$ is supported in $2^j \mathscr{B}$, then $\Delta_i u_j \neq 0$ only if $2^i \lesssim 2^j$. Therefore,
$$\|\Delta_i u\|_{L^\infty} \leqslant \sum_{j : 2^j \gtrsim 2^i} \|\Delta_i u_j\|_{L^\infty} \leqslant \sup_{k \geqslant -1} \{2^{k\alpha} \|u_k\|_{L^\infty}\} \sum_{j : 2^j \gtrsim 2^i} 2^{-j\alpha} \simeq \sup_{k \geqslant -1} \{2^{k\alpha} \|u_k\|_{L^\infty}\} \, 2^{-i\alpha},$$
where in the last step we used that $\alpha > 0$. □

A similar result also holds for Besov spaces $B^\alpha_{p,q}$ with general $p, q \in [1, \infty]$. As a first application, one can use this lemma to show that while the norm $\|\cdot\|_{B^\alpha_{p,q}}$ depends on the specific partition of unity used to define it, the space $B^\alpha_{p,q}$ does not and every other partition of unity induces an equivalent norm.

2.2 The paraproduct and the resonant term

Now that we know how to measure the regularity of distributions, let us come back to the problem of multiplying distributions. We will follow Bony [6] who introduced *paraproducts* which provide a useful tool to decompose the multiplication into simpler problems. The usefulness of the paraproduct comes from the following simple observation:

Lemma 2.7. *There exists an annulus* $\mathscr{A}$ *such that for all* $j \geqslant 1$ *and all* $i \leqslant j - 2$
$$\operatorname{supp}(\mathscr{F}(\Delta_i u \Delta_j v)) \subset 2^j \mathscr{A}, \qquad u, v \in \mathscr{S}'.$$
Moreover, there exists a ball $\mathscr{B}$ *such that for all* $i, j \geqslant -1$ *with* $|i - j| \leqslant 1$
$$\operatorname{supp}(\mathscr{F}(\Delta_i u \Delta_j v)) \subset 2^j \mathscr{B}.$$

Proof. This is quite simple:
$$\begin{aligned}\operatorname{supp}(\mathscr{F}(\Delta_i u \Delta_j v)) &= \operatorname{supp}(\mathscr{F}\Delta_i u * \mathscr{F}\Delta_j u) \subset \overline{\operatorname{supp}(\mathscr{F}\Delta_i u) + \operatorname{supp}(\mathscr{F}\Delta_j u)} \\ &\subset 2^i \tilde{\mathscr{A}} + 2^j \tilde{\mathscr{A}} = 2^j (2^{i-j} \tilde{\mathscr{A}} + \tilde{\mathscr{A}})\end{aligned}$$
for an annulus $\tilde{\mathscr{A}}$. By our assumptions on the dyadic partition of unity we can choose $\tilde{\mathscr{A}}$ such that $2^{i-j} \tilde{\mathscr{A}} + \tilde{\mathscr{A}} \subset \mathscr{A}$ for a new annulus $\mathscr{A}$ and all $i \leqslant j - 2$.

If on the other hand $|i - j| \leqslant 1$, then all we can say is that $\operatorname{supp}(\mathscr{F}(\Delta_i u \Delta_j v)) \subset 2^j \mathscr{B}$ for a ball $\mathscr{B}$. □

Intuitively, this means that multiplying $\Delta_j v$, a function that lives on the spatial scale 2^{-j}, with $\Delta_i u$ for $i \leqslant j-2$, we obtain a new function $\Delta_i u \Delta_j v$ which still lives on the spatial scale 2^{-j}. The multiplication does not create any effects on larger scales. If on the other hand $|i-j| \leqslant 1$, then $\Delta_i u$ and Δ_j live on the spatial scale 2^{-j}, but multiplying the two together can create effects on the scale 1, i.e. small scale contributions work together to create an effect on large scales. We interpret this as a *resonance* phenomenon.

Example 2.8. Below we see a slowly oscillating function u (red curve) and a fast sine curve v (blue curve). The product uv is shown under the two curves. We see that the local fluctuations of uv are due to v, and that uv is essentially oscillating with the same speed as v.

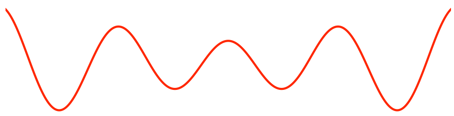


Figure 2.1. u oscillates slowly.

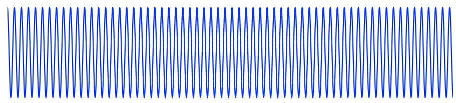


Figure 2.2. v is a fast sine curve.

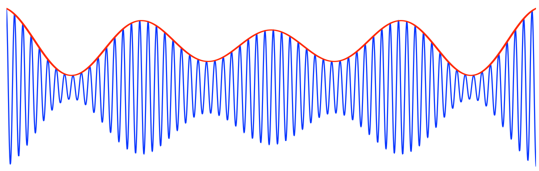


Figure 2.3. uv still lives on the same scale as v.

Formally we can decompose the product uv of two distributions as

$$uv = \sum_{i,j \geqslant -1} \Delta_i u \, \Delta_j v = u \varolessthan v + u \varogreaterthan v + u \odot v.$$

Here $u \varolessthan v$ is the part of the double sum with $i \leqslant j-2$, $u \varogreaterthan v$ is the part with $i \geqslant j+2$, and $u \odot v$ is the "diagonal" part, where $|i-j| \leqslant 1$. More precisely, we define

$$u \varolessthan v = v \varogreaterthan u = \sum_{j \geqslant -1} \Delta_{\leqslant j-2} u \, \Delta_j v \qquad \text{and} \qquad u \odot v = \sum_{i,j:|i-j| \leqslant 1} \Delta_i u \, \Delta_j v.$$

We call $u \varolessthan v$ and $u \varogreaterthan v$ *paraproducts*, and $u \odot v$ the *resonant term*.

Bony's crucial observation is that $u \varolessthan v$ (and thus $u \varogreaterthan v$) is always a well-defined distribution. Heuristically, $u \varolessthan v$ behaves at large frequencies (i.e. small spatial scales) like v and thus retains the same regularity, and u provides only a frequency modulation of v. This can also be seen in Example 2.8 above, where the product uv is actually equal to the paraproduct $u \varolessthan v$ because u has no rapidly oscillating components. The only difficulty in constructing uv for arbitrary distributions lies in handling the diagonal term $u \odot v$. The following key estimates provide the analytically precise formulation of the preceding heuristic discussion:

Theorem 2.9. (Paraproduct estimates) *For all $\beta \in \mathbb{R}$ and $u, v \in \mathscr{S}'$ we have*

$$\|u \varolessthan v\|_\beta \lesssim \|u\|_{L^\infty} \|v\|_\beta, \tag{2.1}$$

and for $\alpha < 0$ furthermore

$$\|u \varolessthan v\|_{\alpha+\beta} \lesssim \|u\|_\alpha \|v\|_\beta. \tag{2.2}$$

If $\alpha + \beta > 0$ we have

$$\|u \odot v\|_{\alpha+\beta} \lesssim \|u\|_\alpha \|v\|_\beta. \tag{2.3}$$

Proof. By Lemma 2.7 there exists an annulus $\mathscr{A}$ such that $\operatorname{supp}(\mathscr{F}(\Delta_{\leqslant j-2}u\Delta_j v))\subset 2^j\mathscr{A}$, and for $u\in L^\infty$ we have

$$\|\Delta_{\leqslant j-2}\,u\Delta_j\,v\|_{L^\infty}\leqslant\|\Delta_{\leqslant j-2}u\|_{L^\infty}\|\Delta_j v\|_{L^\infty}\lesssim\|u\|_{L^\infty}2^{-j\beta}\|v\|_\beta.$$

Inequality (2.1) now follows from Lemma 2.6. The proof of (2.2) and (2.3) works in the same way, except that for estimating $u\odot v$ we need $\alpha+\beta>0$ because now the terms of the series are supported in balls and not in annuli. □

The ill–posedness of $u\odot v$ for $\alpha+\beta\leqslant 0$ can be interpreted as a resonance effect since $u\odot v$ contains exactly those part of the double series where u and v are in the same frequency range. As discussed above, the paraproduct $u\otimes v$ can be interpreted as frequency modulation of v.

In combination with Exercise 2.2 above we deduce the following simple corollary:

Corollary 2.10. *Let $\alpha+\beta>0$. Then the product $(u,v)\mapsto uv$ of smooth functions can be extended to a bounded bilinear operator from $\mathscr{C}^\alpha\times\mathscr{C}^\beta$ to $\mathscr{C}^{\alpha\wedge\beta}$. While $u\otimes v$, $u\circledcirc v$, and $u\odot v$ depend on our specific dyadic partition of unity, the product uv does not.*

The condition $\alpha+\beta>0$ is essentially sharp:

Example 2.11. Let $\alpha,\beta\in\mathbb{R}$ and consider the functions $u_n(x)=n^{-\alpha}e^{inx}$ on $\mathbb{R}$, and $v_n(x)=n^{-\beta}e^{-inx}$. It is easy to see that $\|u_n\|_{\tilde{\alpha}}\to 0$ and $\|v_n\|_{\tilde{\beta}}\to 0$ for all $\tilde{\alpha}<\alpha$ and $\tilde{\beta}<\beta$. Nonetheless

$$u_n v_n\equiv n^{-(\alpha+\beta)}$$

diverges to ∞ whenever $\alpha+\beta<0$, and stays constant for $\alpha+\beta=0$.

Example 2.12. Let $(B_t)_{t\in[0,T]}$ be a Brownian motion and extend $B|_{(-\infty,0]}\equiv 0$ and $B|_{[T,\infty)}\equiv B_T$. Then $B\in\mathscr{C}^\alpha$ for all $\alpha<1/2$ almost surely, and as we saw above this implies $\partial_t B\in\mathscr{C}^{\alpha-1}$. Therefore, the sum of the regularities is $2\alpha-1<0$ and the product $B\partial_t B$ is ill-defined. This manifests itself in the probabilistic phenomenon that there are different reasonable interpretations for the integral $\int_0^t B_s\mathrm{d}B_s=\int_0^t(B_s\partial_s B_s)\mathrm{d}s$, for example Itô, Stratonovich, or backward Itô, roughly speaking because different approximations lead to different limits. In $d=1$ there is actually a certain stiffness, because if (B^n) is a sequence of smooth paths that converge to B in $\mathscr{C}^\alpha$, then we always have

$$B^n\partial_t B^n=\frac{1}{2}\partial_t((B^n)^2)\to\frac{1}{2}\partial_t(B^2),$$

i.e. there appears to be a canonical interpretation for the product $B\partial_t B$, which corresponds to the Stratonovich integral $\int_0^t B_s\circ\mathrm{d}B_s$. However, nobody is forcing us to approximate B with smooth functions, and if we take piecewise constant approximations and compute $\int_0^t B^n_s\mathrm{d}B^n_s$ as a discrete sum, then the limit is the Itô integral $\int_0^t B_s\mathrm{d}B_s$.

Remark 2.13. So far we only considered tempered distributions on $\mathbb{R}^d$, but the same works also on the torus $\mathbb{T}^d=(\mathbb{R}/\mathbb{Z})^d$. In that case we simply have $\mathscr{S}(\mathbb{T})=C^\infty(\mathbb{T})$, and the Fourier transform $\mathscr{F}u(k)=\mathscr{F}_{\mathbb{T}^d}u(k)=\hat{u}(k):=u(e^{-2\pi ik\cdot(\cdot)})$ is defined for $k\in\mathbb{Z}^d$. The Littlewood-Paley blocks are then defined in exactly the same way, and all the results of this section continue to hold. The proofs are mostly the same, only that we have to find a replacement for some scaling arguments: For example it is not true that $(\mathscr{F}^{-1}_{\mathbb{T}^d}\rho_0(2^{-j}\cdot))=2^{jd}(\mathscr{F}^{-1}_{\mathbb{T}^d}\rho_0)(2^j\cdot)$. But usually we can apply the Poisson summation formula to overcome this difficulty. See e.g. [19] for details.

3 The Φ_1^4 and Φ_2^4 equation

Here we work on the torus $\mathbb{T}^d$, and from now on we will slightly restrict the space $\mathscr{C}^\alpha$ and take it as the closure of $C^\infty(\mathbb{T}^d)$ with respect to $\|\cdot\|_\alpha$. This is a strict subspace of $B^\alpha_{\infty,\infty}$, and it is nicer in some respects, for example it is separable. An alternative characterization is

$$\mathscr{C}^\alpha = \left\{ u \in \mathscr{S}' \colon \lim_{j \to \infty} 2^{j\alpha} \|\Delta_j u\|_{L^\infty} = 0 \right\}.$$

So to show that $u \in \mathscr{S}'$ is in the new $\mathscr{C}^\alpha$ it suffices to show that $\|u\|_{\alpha'} < \infty$ for some $\alpha' > \alpha$.

3.1 Regularity analysis

Equipped with these tools, we now aim to solve the Φ_d^4 equation $\phi \colon \mathbb{R}_+ \times \mathbb{T}^d \to \mathbb{R}$,

$$\partial_t \phi = \Delta \phi - \phi^3 + \phi + \xi,$$

where ξ is a space-time white noise on $\mathbb{R}_+ \times \mathbb{T}^d$. Originally we were interested in $\mathbb{R}^d$ instead of $\mathbb{T}^d$, but as will see later, in infinite volume there are considerable technical problems additionally to the regularity problems that already appear in finite volume. So at least for now we work on $\mathbb{T}^d$ and we first learn how to deal with the regularity problems.

Let $(P_t)_{t \geqslant 0}$ be the semigroup generated by Δ, i.e.

$$P_t u = p(t, \cdot) * u$$

for the heat kernel

$$p(t,x) = \sum_{k \in \mathbb{Z}^d} e^{2\pi i k \cdot x} e^{-|2\pi k|^2 t} = \sum_{k \in \mathbb{Z}^d} (4\pi t)^{-d/2} \exp\left(-\frac{|x+k|^2}{4t} \right),$$

where the second step follows from the Poisson summation theorem, see Lemma 3.6 of [19]. By Duhamel's principle (= the variation of constants formula) the solution u to

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

is given by

$$u(t,x) = (p(t,\cdot) * u_0)(x) + \int_0^t (p(t-s,\cdot) * f(s,\cdot))(x) \mathrm{d}s,$$

or in slightly shorter notation

$$u(t) = P_t u_0 + \int_0^t P_{t-s} f(s) \mathrm{d}s.$$

This can be easily shown with the help of the Fourier transform and the finite-dimensional variation of constants formula.

Let's apply Duhamel's principle to our equation, and for simplicity we take $\phi(0) = \phi_0 = 0$. Then

$$\phi(t) = \int_0^t P_{t-s}(-\phi(s)^3 + \phi(s) + \xi(s)) \mathrm{d}s = \int_0^t P_{t-s}(-\phi(s)^3 + \phi(s)) \mathrm{d}s + \int_0^t P_{t-s} \xi(s) \mathrm{d}s.$$

We do not expect any substantial cancellations between the first and the second term on the right hand side, so ϕ should have the same regularity as

$$Z(t) = \int_0^t P_{t-s} \xi(s) \mathrm{d}s.$$

So to see in which function space we can hope to solve the equation for ϕ, let us compute the regularity of Z. For that purpose we use the following type of *Kolmogorov continuity criterion*, where we write for a Banach space X and $T>0$, $\gamma\in[0,1]$:

$$\|u\|_{C_T^\gamma X} := \sup_{t\in[0,T]} \|u(t)\|_X + \sup_{0\leqslant s<t\leqslant T} \frac{\|u(t)-u(s)\|_X}{|t-s|^\gamma}.$$

Note that for $\gamma=0$ this is equivalent to the supremum norm, and in that case we also write $C_T X := C_T^0 X$.

Lemma 3.1. *Let $(u(t))_{t\in[0,T]}$ be a stochastic process with values in $\mathscr{S}'(\mathbb{T}^d)$ and assume that for all $j\geqslant -1$, for all $0\leqslant s<t\leqslant T$, and for all $x\in\mathbb{T}^d$*

$$\mathbb{E}[|\Delta_j u(0,x)|^p]^{1/p} + \frac{\mathbb{E}[|\Delta_j u(t,x)-\Delta_j u(s,x)|^p]^{1/p}}{|t-s|^\gamma} \leqslant K 2^{-j\alpha}, \tag{3.1}$$

where $\gamma>1/p$. Then we have for all $\gamma'<\gamma-1/p$ and all $\alpha'<\alpha$

$$\mathbb{E}\Big[\|u\|^p_{C_T^{\gamma'}\mathscr{C}^{\alpha'-d/p}}\Big]^{1/p} \lesssim \mathbb{E}\Big[\|u\|^p_{C_T^{\gamma'}B^{\alpha'}_{p,p}}\Big]^{1/p} \lesssim K. \tag{3.2}$$

Proof. The first inequality is simply the Besov embedding theorem. To see the second inequality, note that

$$\begin{aligned}\mathbb{E}\Big[\|u(t)-u(s)\|^p_{B^{\alpha'}_{p,p}}\Big] &= \sum_{j\geqslant -1} 2^{j\alpha' p}\int_{\mathbb{T}^d} \mathbb{E}[|\Delta_j u(t,x)-\Delta_j u(s,x)|^p]\mathrm{d}x\\ &\leqslant \sum_{j\geqslant -1} 2^{j\alpha' p}\int_{\mathbb{T}^d} K^p|t-s|^{\gamma p}2^{-j\alpha p}\mathrm{d}x \lesssim K^p|t-s|^{\gamma p},\end{aligned}$$

where we used that $\alpha'<\alpha$ and that $\mathbb{T}^d$ has finite volume. Similarly $\mathbb{E}[\|u(0)\|^p_{B^{\alpha'}_{p,p}}] \lesssim K^p$, and since $\gamma-1/p>0$ we can apply Kolmogorov's continuity criterion (for Banach space valued processes) and obtain for $\gamma'<\gamma-1/p$

$$\mathbb{E}\Big[\|u\|^p_{C_T^{\gamma'}B^{\alpha'}_{p,p}}\Big]^{1/p} \lesssim K.$$

□

Remark 3.2. The proof crucially used that $\int_{\mathbb{T}^d} 1\mathrm{d}x<\infty$. On $\mathbb{R}^d$ a uniform bound as in (3.1) would only show that u has trajectories in a weighted Besov space, and this induces the technical difficulties alluded to above. We might get back to that point later.

As an application, we get the regularity of Z:

Lemma 3.3. *We have for all $\gamma<1/2$ and all $\alpha<1-d/2$ and all $p\in[1,\infty)$*

$$\mathbb{E}\big[\|Z\|^p_{C_T^0\mathscr{C}^\alpha} + \|Z\|^p_{C_T^\gamma\mathscr{C}^{\alpha-1}}\big] < \infty.$$

Proof. By Lemma 3.1 it suffices to show for all $\lambda\in[0,1]$, for all $p\in[1,\infty)$, and for all $x\in\mathbb{T}^d$ and $0\leqslant s<t\leqslant T$

$$\mathbb{E}[|\Delta_j Z(t,x)-\Delta_j Z(s,x)|^p]^{1/p} \lesssim 2^{j(d/2-1+\lambda)}|t-s|^{\lambda/2}. \tag{3.3}$$

Indeed, if we apply this inequality with $\lambda = 1$ we get $\mathbb{E}\big[\|Z\|^p_{C_T^\gamma \mathscr{C}^{\alpha-1}}\big] < \infty$, and if we apply it with $\lambda \simeq 0$ (but positive) and p large enough so that $p\lambda > 1$ we get $\mathbb{E}[\|Z\|^p_{C_T^\varepsilon \mathscr{C}^\alpha}] < \infty$ for some $\varepsilon > 0$ and thus $\mathbb{E}\big[\|Z\|^p_{C_T^0 \mathscr{C}^\alpha}\big] < \infty$.

So let us derive (3.3). Since $\Delta_j Z(t) - \Delta_j Z(s)$ is a Gaussian random variable (it is a linear functional of the Gaussian process ξ), we have with $c_p = \mathbb{E}[|X|^p]$ for a $\mathcal{N}(0,1)$ variable X, and using the orthogonality of the integrals $\int_0^s$ and $\int_s^t$,

$$
\begin{aligned}
&\mathbb{E}[|\Delta_j Z(t,x) - \Delta_j Z(s,x)|^p]^{2/p} = c_p^{2/p} \mathbb{E}[|\Delta_j Z(t,x) - \Delta_j Z(s,x)|^2]\\
&\simeq \mathbb{E}\bigg[\bigg|\int_s^t \int_{\mathbb{T}^d} (K_j * p(t-r))(x-z)\xi(r,z)\mathrm{d}z\mathrm{d}r\bigg|^2\bigg]\\
&\quad + \mathbb{E}\bigg[\bigg|\int_0^s \int_{\mathbb{T}^d} (K_j * (p(t-r) - p(s-r)))(x-z)\xi(r,z)\mathrm{d}z\mathrm{d}r\bigg|^2\bigg]\\
&= \int_s^t \int_{\mathbb{T}^d} |(K_j * p(t-r))(x-z)|^2 \mathrm{d}z\mathrm{d}r + \int_0^s \int_{\mathbb{T}^d} |(K_j * (p(t-r) - p(s-r)))(x-z)|^2 \mathrm{d}z\mathrm{d}r\\
&= \int_s^t \sum_{k \in \mathbb{Z}^d} |\rho_j(k)|^2 e^{-2|2\pi k|^2 (t-r)} \mathrm{d}r + \int_0^s \sum_{k \in \mathbb{Z}^d} |\rho_j(k)|^2 e^{-2|2\pi k|^2 (s-r)} (e^{-|2\pi k|^2 (t-s)} - 1)^2 \mathrm{d}r\\
&\lesssim \sum_{k \in \mathbb{Z}^d} |\rho_j(k)|^2 [\min\{|t-s|, |k|^{-2}\} + \min\{|k|^{-2}, |k|^{-2}|k|^2|t-s|\}]\\
&\lesssim 2^{jd} \min\{|t-s|, 2^{-2j}\},
\end{aligned}
$$

because $\rho_j(k) \neq 0$ for $O(2^{jd})$ values of k. Now (3.3) follows by interpolation. □

These computations are essentially sharp and a slight modification of the proof shows that $\mathbb{E}[\|Z(t)\|^p_{B^{1-d/2}_{p,p}}] = \infty$ for all $p \in [1, \infty)$ and all $t > 0$. Therefore, Z is function valued if and only if $d = 1$, and the regularity gets worse with increasing dimension. This is a first indication why also the solution theory gets more and more complicated with increasing dimension. Let us first see how to solve the Φ_d^4 equation in the easiest case $d = 1$.

3.2 The Φ_1^4 equation

We now assume that $d = 1$ and we fix $\alpha \in (0, 1/2)$ and consider a general initial condition $\phi_0 \in \mathscr{C}^\alpha$ (not necessarily $\phi_0 = 0$). We also replace the drift $-\phi^3 + \phi$ by $-\phi^3$ to simplify the notation, because the linear term does not introduce any additional difficulties. Then Duhamel's formula gives

$$
\phi(t) = P_t \phi_0 + \int_0^t P_{t-s}(-\phi(s)^3)\mathrm{d}s + Z(t) = \int_0^t P_{t-s}(-\phi(s)^3)\mathrm{d}s + \tilde{Z}(t)
$$

for $\tilde{Z}(t) = Z(t) + P_t \phi_0$. It is not hard to see that $\tilde{Z} \in C_T \mathscr{C}^\alpha$ for all $T > 0$ (here we need the new definition of $\mathscr{C}^\alpha$, otherwise $t \mapsto P_t \phi_0$ might have a discontinuity at 0). Also, we have by Young's inequality

$$
\|P_r u\|_\alpha = \sup_j 2^{j\alpha} \|p(r) * \Delta_j u\|_{L^\infty} \leqslant \sup_j 2^{j\alpha} \|p(r)\|_{L^1} \|\Delta_j u\|_{L^\infty} = \|u\|_\alpha
$$

uniformly in r, and therefore we can set up a Picard iteration by defining

$$
\begin{aligned}
\Psi \colon C_T \mathscr{C}^\alpha &\to C_T \mathscr{C}^\alpha\\
\Psi(\phi)(t) &= \tilde{Z}(t) + \int_0^t P_{t-s}(-\phi(s)^3)\mathrm{d}s.
\end{aligned}
$$

Indeed, we have by the paraproduct estimates (Corollary 2.10, here we need $\alpha > 0$) for some $K > 0$ which is independent of T

$$\sup_{t \leqslant T} \|\Psi(\phi)(t)\|_\alpha \leqslant \sup_{t \leqslant T} \|\tilde{Z}(t)\|_\alpha + \sup_{t \leqslant T} \int_0^t \|\phi(s)^3\|_\alpha \mathrm{d}s$$
$$\leqslant \sup_{t \leqslant T} \|\tilde{Z}(t)\|_\alpha + KT \sup_{t \leqslant T} \|\phi(t)\|_\alpha^3,$$

which shows that Ψ maps to $C_T \mathscr{C}^\alpha$. Set

$$M = \sup_{t \leqslant 1} \|\tilde{Z}(t)\|_\alpha.$$

If $T \in (0, 1]$ is small enough, depending on M, then Ψ leaves the ball $\mathscr{B}_{C_T \mathscr{C}^\alpha}(0, 2M)$ in $C_T \mathscr{C}^\alpha$ with center 0 and radius $2M$ invariant. Moreover, since

$$\|\phi_1^3 - \phi_2^3\| \lesssim (\|\phi_1\|_\alpha^2 + \|\phi_2\|_\alpha^2) \|\phi_1 - \phi_2\|_\alpha,$$

we get that for possibly smaller $T > 0$ the map Ψ is a contraction on $\mathscr{B}_{C_T \mathscr{C}^\alpha}(0, 2M)$. By the Banach fixed point theorem we thus find a unique solution to (3.2) on the interval $[0, T]$, where T depends on $\tilde{Z}$ through M and thus may be random. We can iterate the construction, but since the initial condition is part of $\tilde{Z}$ the time interval in the second iteration step might be strictly smaller. Ultimately we get the existence of a $T^* \in (0, \infty]$ and a unique solution $\phi \in C_T \mathscr{C}^\alpha$ for all $T < T^*$, such that in the case $T^* < \infty$

$$\lim_{t \to T^*} \|\phi(t)\|_\alpha = \infty.$$

In other words, $[0, T^*)$ is the maximal existence interval, and the solution blows up at T^*, or it exists for all times and $T^* = \infty$. Thus, we have established the following result:

Proposition 3.4. *Let $d = 1$ and $\alpha \in (0, 1/2)$ and $\phi_0 \in \mathscr{C}^\alpha$. There exists a random time $T^* \in (0, \infty]$ and a unique $\phi \in C_{T^*} \mathscr{C}^\alpha := \bigcup_{T < T^*} C_T \mathscr{C}^\alpha$ such that*

$$\phi(t) = P_t \phi_0 + \int_0^t P_{t-s}(-\phi(s)^3) \mathrm{d}s + Z(t), \qquad t \in [0, T^*).$$

By slightly refining the analysis we could show that the solution ϕ depends continuously on $Z \in C_T \mathscr{C}^\alpha$ and on $\phi_0 \in \mathscr{C}^\alpha$.

And actually we would not expect the solution to blow up, and indeed it does not. But to see this we would have to use the sign of the nonlinearity: The above analysis works also for the equation with $+\phi^3$ instead of $-\phi^3$, and in that case we expect it to blow up in finite time.

At least for now we do not worry about these problems, and instead we start increasing the dimension. The previous discussion breaks down if $d > 1$ because then we have to take $\alpha < 0$ and then the estimate $\|\phi(s)^3\|_\alpha \lesssim \|\phi(s)\|_\alpha^3$ is false (and $\phi(s)^3$ is not even well defined for $\phi \in C_T \mathscr{C}^\alpha$).

3.3 Schauder estimates and the Φ_2^4 equation

In $d = 2$ we have $Z \in C_T \mathscr{C}^{0-} := \bigcup_{\varepsilon > 0} C_T \mathscr{C}^{-\varepsilon}$, and therefore even $Z(s)^3$ is ill-defined, let alone $\phi(s)^3$. We ignore this problem for now and decompose $\phi = Z + v$, a strategy which is due to Da Prato and Debussche [13]. Then v should solve

$$(\partial_t - \Delta) v = -\phi^3 = -(Z^3 + 3Z^2 v + 3Z v^2 + v^3), \qquad v(0) = \phi_0 - Z(0) = \phi_0,$$

where we used that $Z(0)=0$. If we ignore that the products Z^3 and Z^2 are ill-defined and we simply apply the paraproduct estimates to them anyways, then we get $Z^2, Z^3 \in C_T\mathscr{C}^{0-}$ (convince yourself that if we ignore the constraint $\alpha+\beta>0$ for $f\odot g$, then the product $fg = f\varolessthan g + f\varogreaterthan g + f\odot g$ of $f\in\mathscr{C}^\alpha$ and $g\in\mathscr{C}^\beta$ should have regularity $fg\in\mathscr{C}^{\alpha\wedge\beta\wedge(\alpha+\beta)}$). So let us continue our computations under the assumption that $Z^2, Z^3\in C_T\mathscr{C}^{0-}$ are given – this may look like a bold assumption, but Z is a a Gaussian process, and in probability theory we deal with ill-defined nonlinear operations of Gaussian processes all the time: Recall the integral $\int B\mathrm{d}B$ for a Brownian motion; the reason why we can do more in the stochastic setting is that the analytic paraproduct theory gives only worst case estimates, while in the stochastic case there are cancellations coming from independence properties of the noise, and we are not in a worst case situation.

If $Z^2, Z^3\in C_T\mathscr{C}^{0-}$ are given, the right hand side of the equation for v is well defined as long as we can estimate $v\in C_T\mathscr{C}^{0+} = \bigcup_{\varepsilon>0} C_T\mathscr{C}^\varepsilon$, and in that case this right hand side is in $C_T\mathscr{C}^{0-}$. To close our estimate we need Schauder estimates, which tell us that inverting the operator $(\partial_t-\Delta)$ gains two space derivatives:

Lemma 3.5. (Schauder estimates) *Let $\alpha\in\mathbb{R}$ and let $(P_t)_{t\geqslant 0}$ be the semigroup generated by Δ. Given $f\in C_T\mathscr{C}^{\alpha-2}$ and $\varphi\in\mathscr{C}^\alpha$, let u be the unique weak solution to*

$$(\partial_t-\Delta)u = f, \qquad u(T)=\varphi.$$

Then

$$\|u\|_{C_T\mathscr{C}^\alpha} \lesssim (1+T)(\|f\|_{C_T\mathscr{C}^{\alpha-2}} + \|\varphi\|_\alpha).$$

Proof. Recall that by Duhamel's formula $u(t) = P_t\varphi + \int_0^t P_{t-s} f(s)\mathrm{d}s$. To derive the estimate for $\|u\|_{C_T\mathscr{C}^\alpha}$ note that for $j\geqslant 0$ the term $\Delta_j\psi$ only has frequencies of the order 2^j and since $\Delta_j P_t\psi = \mathcal{F}^{-1}\big(e^{-|2\pi\cdot|^2 t}\rho_j(\cdot)\hat{\psi}\big)$ we get at least formally for all $\gamma\geqslant 0$ and $\beta\in\mathbb{R}$ (and some $c>0$)

$$\begin{aligned}\|\Delta_j P_t\psi\|_{L^\infty} &\simeq e^{-ct2^{2j}}\|\Delta_j\psi\|_{L^\infty} = (t^{1/2}2^j)^{-\gamma}\sup_{r>0}(r^\gamma e^{-cr^2})\|\Delta_j\psi\|_{L^\infty}\\ &\lesssim t^{-\gamma/2}2^{-j(\gamma+\beta)}\|\psi\|_\beta.\end{aligned}$$

This can be made rigorous by using similar arguments as in the proof of Bernstein's inequality. For $\Delta_{-1}\psi$ we can only estimate

$$\|\Delta_{-1}P_t\psi\|_{L^\infty} \leqslant \|\Delta_{-1}\psi\|_{L^\infty} \leqslant 2^{-\gamma}2^{\gamma+\beta}\|\psi\|_\beta.$$

Thus, we have

$$\|P_t\psi\|_{\beta+\gamma} \lesssim (1+t^{-\gamma/2})\|\psi\|_\beta \tag{3.4}$$

for all $\gamma\geqslant 0$ and $\beta\in\mathbb{R}$. We apply this with $\gamma=0$, $\beta=\alpha$, and $\psi=\varphi$ to obtain

$$\|(t\mapsto P_t\varphi)\|_{C_T\mathscr{C}^\alpha} \lesssim \|\varphi\|_\alpha.$$

Also,

$$\begin{aligned}\left\|\int_0^t P_{t-s}f(s)\mathrm{d}s\right\|_{\alpha-\varepsilon} &\leqslant \int_0^t \|P_{t-s}f(s)\|_{\alpha-\varepsilon}\mathrm{d}s \lesssim \int_0^t (1+|t-s|^{-1+\varepsilon/2})\|f\|_{C_T\mathscr{C}^{\alpha-2}}\mathrm{d}s\\ &\lesssim (t+t^{\varepsilon/2})\|f\|_{C_T\mathscr{C}^{\alpha-2}}.\end{aligned}$$

For $\varepsilon=0$ there is a problem because $|t-s|^{-1}$ barely fails to be integrable. But here we already see that if we wanted to regularize f by $2-\varepsilon$ degrees of regularity, then we would get an estimate for u. To gain two full derivatives we have to be slightly more careful and use two different estimates, one for s close to 0, and one for s close to t, see Lemma A.9 in [16] for details. □

Proposition 3.6. *Let $d=2$ and $\alpha \in (1,2)$ and $\phi_0 \in \mathscr{C}^\alpha$. Assume that $Z, Z^2, Z^3 \in C(\mathbb{R}_+, \mathscr{C}^{\alpha-2})$ are given. Then there exists a time $T^* \in (0,\infty]$ depending on Z, Z^2, Z^3 and ϕ_0 and a unique $v \in C_{T^*}\mathscr{C}^\alpha := \bigcup_{T<T^*} C_T \mathscr{C}^\alpha$ such that*

$$v(t) = P_t \phi_0 + \int_0^t P_{t-s}(-(Z^3 + 3Z^2 v + 3Z v^2 + v^3)(s))\mathrm{d}s, \qquad t \in [0, T^*).$$

Moreover, v depends continuously on (ϕ_0, Z, Z^2, Z^3).

Proof. Let $\alpha' \in (1, \alpha)$ and $T \in (0,1]$. We define the map

$$\Psi : C_T \mathscr{C}^{\alpha'} \to C_T \mathscr{C}^{\alpha'}$$
$$\Psi(v)(t) = P_t \phi_0 + \int_0^t P_{t-s}(-(Z^3 + 3Z^2 v + 3Z v^2 + v^3)(s))\mathrm{d}s.$$

From the proof of Lemma 3.5 we get

$$\|\Psi(v)\|_{C_T \mathscr{C}^{\alpha'}} \lesssim \|\phi_0\|_\alpha + \|Z\|_{C_T \mathscr{C}^\alpha} + T^{(\alpha-\alpha')/2} \|Z^3 + 3Z^2 v + 3Z v^2 + v^3\|_{C_T \mathscr{C}^{\alpha-2}}.$$

Now we apply the product estimates (which is ok since $\alpha' + \alpha - 2 > 2 - 2 = 0$) to obtain

$$\begin{aligned} \|Z^3 + 3Z^2 v + 3Z v^2 + v^3\|_{C_T \mathscr{C}^{\alpha-2}} &\lesssim \|Z^3\|_{C_T \mathscr{C}^{\alpha-2}} + \|Z^2\|_{C_T \mathscr{C}^{\alpha-2}} \|v\|_{C_T \mathscr{C}^{\alpha'}} \\ &\quad + \|Z\|_{C_T \mathscr{C}^{\alpha-2}} \|v^2\|_{C_T \mathscr{C}^{\alpha'}} + \|v^3\|_{C_T \mathscr{C}^\alpha} \\ &\lesssim \|Z^3\|_{C_T \mathscr{C}^{\alpha-2}} + \|Z^2\|_{C_T \mathscr{C}^{\alpha-2}} \|v\|_{C_T \mathscr{C}^{\alpha'}} \\ &\quad + \|Z\|_{C_T \mathscr{C}^{\alpha-2}} \|v\|^2_{C_T \mathscr{C}^{\alpha'}} + \|v\|^3_{C_T \mathscr{C}^\alpha}, \end{aligned}$$

where in the second step we used that $\alpha' > 0$ and applied once more the product estimates. So overall

$$\|\Psi(v)\|_{C_T \mathscr{C}^{\alpha'}} \leqslant \|\phi_0\|_\alpha + \|Z\|_{C_T \mathscr{C}^\alpha} + T^{(\alpha-\alpha')/2}(1 + \|\mathbb{Z}\|)\big(1 + \|v\|^3_{C_T \mathscr{C}^{\alpha'}}\big),$$

where

$$\|\mathbb{Z}\| = \|Z\|_{C_T \mathscr{C}^{\alpha-2}} + \|Z^2\|_{C_T \mathscr{C}^{\alpha-2}} + \|Z^3\|_{C_T \mathscr{C}^{\alpha-2}}$$

and where we used that $|x| + |x|^2 \lesssim 1 + |x|^3$. Now we get a unique solution $v \in C_T \mathscr{C}^{\alpha'}$ by the same argument as for Propsition 3.4. Moreover, Lemma 3.5 shows that

$$\|v\|_{C_T \mathscr{C}^\alpha} \lesssim \|\phi_0\|_\alpha + \|Z^3 + 3Z^2 v + 3Z v^2 + v^3\|_{C_T \mathscr{C}^{\alpha-2}} < \infty,$$

i.e. a posteriori we even get $v \in C_T \mathscr{C}^\alpha$ instead of $C_T \mathscr{C}^{\alpha'}$.

The solution depends continuously on the data because all the operations in the equation are continuous (and we made the discontinuous parts Z^2, Z^3 part of the data).

To obtain the solution up to an explosion time $T^* > 0$ we need to iterate the construction. A subtlety is that at time T we no longer have $Z(T) = 0$ (which was needed before to get a nice initial condition for v). But in fact we should take now $v(T)$ as new initial condition, and we just saw that this is in $\mathscr{C}^\alpha$, so we can indeed iterate the construction up to a (possibly finite) explosion time. □

This analysis works as long as Z^2, Z^3 are given. To construct theses terms we use a Fourier truncation and set

$$Z_\varepsilon(t) := \mathscr{F}^{-1}(\varphi_\varepsilon(\cdot) \mathscr{F} Z(t, \cdot))$$

for $\varphi_\varepsilon(k) = \varphi(\varepsilon k)$ for a compactly supported bounded and even function φ which is continuous around 0 and satisfies $\varphi(0) = 1$ (you may think of $\varphi \in C_c^\infty$ with $\varphi(0) = 1$ or $\varphi = \mathbb{I}_{[-1,1]}$). If we can show that Z_ε^k converges in $C_T \mathscr{C}^{0-}$ to a limit as $\varepsilon \to 0$, then this limit would provide a natural interpretation of Z^k.

But it turns out that Z_ε^k diverges for all $k > 1$! However, the divergence is not so bad, and we can cure it by subtracting simple counterterms.

3.4 Stochastic computations

Unlike for $k=1$ and/or $d=1$, the construction of Z^k for $k>1$ and $d>1$ is nontrivial. Here we introduce the main tools for this (renormalization, hypercontractivity), and sketch the estimates. The presentation is strongly inspired by [28, 30].

Recall that $Z(t)=\int_0^t P_{t-s}\xi(s)\mathrm{d}s$, so

$$Z_\varepsilon(t)=\int_0^t P_{t-s}^\varepsilon \xi(s)\mathrm{d}s, \qquad P_t^\varepsilon f=\mathscr{F}^{-1}(\varphi_\varepsilon e^{-|2\pi\cdot|^2 t}\mathscr{F}f).$$

Lemma 3.7. *We have*

$$\mathbb{E}[\Delta_j(Z_\varepsilon^2)(t,x)]=\delta_{j=-1}C_\varepsilon(t),$$

where δ is the Kronecker delta and we have for $t>0$

$$C_\varepsilon(t)=t+\sum_{k\neq 0}\frac{|\varphi(\varepsilon k)|^2}{2|2\pi k|^2}(1-e^{-2|2\pi k|^2 t})=\begin{cases} O(1), & d=1,\\ O(|\log(\varepsilon)|), & d=2,\\ O(\varepsilon^{2-d}), & d\geqslant 3.\end{cases}$$

Proof. We have with $\mathscr{F}p^\varepsilon(t,k)=\varphi_\varepsilon(k)\mathscr{F}p(t,k)=\varphi_\varepsilon(k)e^{-|2\pi k|^2 t}$

$$\begin{aligned}\mathbb{E}[\Delta_j(Z_\varepsilon^2)(t,x)] &= \int_{\mathbb{T}^d}\mathrm{d}y K_j(x-y)\mathbb{E}[(Z_\varepsilon^2)(t,y)]\\ &= \int_{\mathbb{T}^d}\mathrm{d}y K_j(x-y)\int_0^t \mathrm{d}s\int_{\mathbb{T}^d}\mathrm{d}z p^\varepsilon(t-s,y-z)^2\\ &= \int_{\mathbb{T}^d}\mathrm{d}y K_j(x-y)\int_0^t \mathrm{d}s\sum_k |\varphi_\varepsilon(k)e^{-|2\pi k|^2 t}|^2\\ &= \left(\int_{\mathbb{T}^d}\mathrm{d}y K_j(y)\right)\times\left(t+\sum_{k\neq 0}\frac{|\varphi(\varepsilon k)|^2}{2|2\pi k|^2}(1-e^{-2|2\pi k|^2 t})\right),\end{aligned}$$

from where the claim follows. □

So for $d>1$ we cannot construct Z^2 in this way, since the expectation of $\Delta_{-1}Z_\varepsilon^2$ diverges. However, it turns out that we can cure this divergence by simply subtracting the expectation. To see this, we need the *Hermite polynomials*:

Definition 3.8. *For $x\in\mathbb{R}$ and $t\geqslant 0$ we set*

$$H_0(x,t)=1, \qquad H_n(x,t)=xH_{n-1}(x,t)-t\partial_x H_{n-1}(x,t).$$

The first few Hermite polynomials are

$$\begin{gathered} H_0(x,t)=1, \qquad H_1(x,t)=x, \qquad H_2(x,t)=x^2-t,\\ H_3(x,t)=x^3-3tx, \qquad H_4(x,t)=x^4-6tx^2+3t^2.\end{gathered}$$

Note that for $n=2$ we get

$$H_2(Z_\varepsilon(t,x),C_\varepsilon(t))=Z_\varepsilon(t,x)^2-C_\varepsilon(t)=Z_\varepsilon(t,x)^2-\mathbb{E}[Z_\varepsilon(t,x)^2],$$

and this suggests that $H_2(Z_\varepsilon(t,x),C_\varepsilon(t))$ might be better behaved than $Z_\varepsilon(t,x)^2$. It turns out that also the higher order Hermite polynomials $H_n(Z_\varepsilon(t,x),C_\varepsilon(t))$ are better behaved than $Z_\varepsilon(t,x)^n$. Intuitively this can be explained by the fact that the $H_n(Z_\varepsilon(t,x),C_\varepsilon(t))$ are (multiples of) orthogonal projections:

Exercise 3.1. Show that for $t>0$

$$H_n(x,t)=(-t)^n e^{x^2/(2t)}\partial_x^n e^{-x^2/(2t)}. \tag{3.5}$$

Conclude that the family $(H_n(\cdot,t))_{n\in\mathbb{N}_0}$ is orthogonal with respect to the centered Gaussian measure with variance t. Show also that

$$\partial_x H_n = nH_{n-1}, \qquad \partial_t H_n = -\frac{n(n-1)}{2}H_{n-2} = -\frac{1}{2}\partial_x^2 H_n, \tag{3.6}$$

i.e. that each Hermite polynomial solves the backward heat equation $\left(\partial_t+\frac{1}{2}\partial_x^2\right)H_n=0$.

Solution. We use induction. For $n=0$ the identity is obvious, so assume that it holds for $n-1$. Then

$$\begin{aligned}
(-t)^n e^{x^2/(2t)}\partial_x^n e^{-x^2/(2t)} &= (-t)^n e^{x^2/(2t)}\partial_x\big((-t)^{-(n-1)}e^{-x^2/(2t)}H_{n-1}(x,t)\big)\\
&= (-t)^n e^{x^2/(2t)} \times \big((-t)^{-n} x e^{-x^2/(2t)} H_{n-1}(x,t)\\
&\quad + (-t)^{-(n-1)}e^{-x^2/(2t)}\partial_x H_{n-1}(x,t)\big)\\
&= xH_{n-1}(x,t) - t\partial_x H_{n-1}(x,t)\\
&= H_n(x,t).
\end{aligned}$$

Consequently, we have for $m\leqslant n\in\mathbb{N}_0$ and $\mu_t(\mathrm{d}x)=\frac{1}{\sqrt{2\pi t}}e^{-x^2/(2t)}\mathrm{d}x$

$$\begin{aligned}
\int_{\mathbb{R}} H_n(x,t)H_m(x,t)\mu_t(\mathrm{d}x) &= \frac{(-t)^n}{\sqrt{2\pi t}}\int_{\mathbb{R}}\partial_x^n e^{-x^2/(2t)}H_m(x,t)\mathrm{d}x\\
&= \frac{t^n}{\sqrt{2\pi t}}\int_{\mathbb{R}} e^{-x^2/(2t)}\partial_x^n H_m(x,t)\mathrm{d}x.
\end{aligned}$$

Since H_m is an m-th degree polynomial the right hand side is equal to zero whenever $n>m$. For $n=m$ we use that the leading coefficient of H_n is 1, and thus $\partial_x^n H^n\equiv n!$, leading to

$$\int_{\mathbb{R}}|H_n(x,t)|^2\mu_t(\mathrm{d}x) = \frac{t^n}{\sqrt{2\pi t}}\int_{\mathbb{R}} e^{-x^2/(2t)} n!\mathrm{d}x = t^n n!.$$

Finally let us show (3.6): For $n=2$ we have $\partial_x H_1 = 2x = 2H_1$ and $\partial_t H_2 = -1 = -\frac{2\times 1}{2}H_0$, so assume that the identities hold for $n-1$:

$$\begin{aligned}
\partial_x H_n &= \partial_x(xH_{n-1} - t\partial_x H_{n-1})\\
&= H_{n-1} + x\partial_x H_{n-1} - t\partial_x(\partial_x H_{n-1})\\
&= H_{n-1} + x(n-1)H_{n-2} - t\partial_x((n-1)H_{n-2})\\
&= H_{n-1} + (n-1)(xH_{n-2} - t\partial_x H_{n-2})\\
&= nH_n,
\end{aligned}$$

and similarly

$$\begin{aligned}
\partial_t H_n &= \partial_t(xH_{n-1} - t\partial_x H_{n-1})\\
&= x\partial_t H_{n-1} - \partial_x H_{n-1} - t\partial_x\partial_t H_{n-1}\\
&= x\left(-\frac{1}{2}\partial_x^2 H_{n-1}\right) - \partial_x H_{n-1} - t\partial_x\left[-\frac{1}{2}\partial_x^2 H_{n-1}\right]\\
&= -\frac{1}{2}[\partial_x^2(xH_{n-1}) - t\partial_x^2\partial_x H_{n-1}]\\
&= -\frac{1}{2}\partial_x^2[xH_{n-1} - t\partial_x H_{n-1}]\\
&= -\frac{1}{2}\partial_x^2 H_n,
\end{aligned}$$

where we used that

$$\partial_x^2(x H_{n-1}) = \partial_x(H_{n-1} + x\partial_x H_{n-1}) = 2\partial_x H_{n-1} + x\partial_x^2 H_{n-1}.$$

Remark 3.9. One can also show that if (X, Y) are jointly centered Gaussian, then

$$\mathbb{E}[H_m(X, \mathbb{E}[X^2]) H_n(Y, \mathbb{E}[Y^2])] = \mathbb{I}_{m=n} n! \mathbb{E}[XY]^n,$$

see Lemma 1.1.1 in [32].

With the help of Hermite polynomials we can do very efficient computations. However, to control

$$\Delta_j(H_n(Z_\varepsilon(t,\cdot), C_\varepsilon(t)))(x) = \int K_j(x-y) H_n(Z_\varepsilon(t,y), C_\varepsilon(t)) \mathrm{d}y$$

we also need to understand sums of Hermite polynomials. This leads to the so called *Wiener-Ito chaos*. The n-th Wiener-Ito chaos will be described by a bounded map

$$W_n: L^2((\mathbb{R}_+ \times \mathbb{T}^d)^n) \to L^2(\Omega).$$

Let us write $E = \mathbb{R}_+ \times \mathbb{T}^d$ from now on. We start by defining for $\varphi \in L^2(E)$

$$W_n(\varphi^{\otimes n}) := H_n(\xi(\varphi), \|\varphi\|_{L^2}^2),$$

and then for $\varphi_1, \ldots, \varphi_m \in L^2(E)$

$$W_n\left(\sum_{k=1}^m \varphi_k^{\otimes n}\right) := \sum_{k=1}^m W_n(\varphi_k^{\otimes n}) = \sum_{k=1}^m H_n(\xi(\varphi_k), \|\varphi_k\|_{L^2}^2).$$

With this definition we get from Remark 3.9

$$\begin{aligned}\mathbb{E}\left[W_n\left(\sum_{k=1}^m \varphi_k^{\otimes n}\right)^2\right] &= \sum_{k,\ell=1}^m \mathbb{E}[H_n(\xi(\varphi_k), \|\varphi_k\|_{L^2}^2) H_n(\xi(\varphi_\ell), \|\varphi_\ell\|_{L^2}^2)] \\ &= \sum_{k,\ell=1}^m n! \langle \varphi_k, \varphi_\ell \rangle_{L^2(E)}^n \\ &= \sum_{k,\ell=1}^m n! \langle \varphi_k^{\otimes n}, \varphi_\ell^{\otimes n} \rangle_{L^2(E^n)} \\ &= n! \left\| \sum_{k=1}^m \varphi_k^{\otimes n} \right\|_{L^2(E^n)}^2,\end{aligned}$$

i.e. W_n is (the multiple of) an isometry from the functions of the type $\sum_{k=1}^m \varphi_k^{\otimes n}$ in $L^2(E^n)$ to $L^2(\Omega)$. Therefore, we can uniquely extend W_n to the closure of these functions in $L^2(E^n)$. By polarization we see that the closure contains all linear combinations of functions of the type $\frac{1}{n!}\sum_{\sigma \in \Sigma_n} \varphi_{\sigma(1)} \otimes \ldots \otimes \varphi_{\sigma(n)}$ for $\varphi_1, \ldots, \varphi_n \in L^2(E)$. Indeed, if Σ_n denotes the permutations of $\{1, \ldots, n\}$, then

$$\frac{1}{n!} \sum_{\sigma \in \Sigma_n} \varphi_{\sigma(1)} \otimes \ldots \otimes \varphi_{\sigma(n)} = \frac{1}{n! 2^n} \sum_{a_1, \ldots, a_n = \pm 1} a_1 \cdot \ldots \cdot a_n (a_1 \varphi_1 + \cdots + a_n \varphi_n)^{\otimes n}$$

which can be shown by expanding the bracket on the right hand side. Based on this we see that the closure contains $L^2_s(E^n)$, the subspace of functions in $L^2(E^n)$ that are symmetric in their n arguments. For a generic function $\varphi \in L^2(E^n)$ we now simply set $W_n(\varphi) = W_n(\tilde{\varphi})$, where

$$\tilde{\varphi}(x_1, \ldots, x_n) = \frac{1}{n!} \sum_{\sigma \in \Sigma_n} \varphi(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

is the symmetrization of φ. Since $\|\tilde{\varphi}\|_{L^2} \leqslant \|\varphi\|_{L^2}$ by the triangle inequality, we have established the following result:

Theorem 3.10. *For all $n \in \mathbb{N}_0$ the map W_n defines a bounded linear operator*

$$W_n \colon L^2(E^n) \to L^2(\Omega).$$

We write $\mathcal{H}_n \subset L^2(\Omega)$ for the image of W_n, which we call the n-th chaos. One can show that

$$L^2(\Omega, \sigma(\xi)) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n,$$

for $\sigma(\xi) := \sigma(\xi(\varphi) \colon \varphi \in L^2)$, which is called the chaos decomposition of $L^2(\Omega, \sigma(\xi))$.

To see that $L^2(\Omega, \sigma(\xi))$ has the chaos decomposition property, it suffices to note that x^n can be written as a linear combination of $H_k(x, t)$ with $k \leqslant n$, and therefore any random variable which is orthogonal to $\bigoplus_{n=0}^{\infty} \mathcal{H}_n$ is orthogonal to all polynomials $\xi(\varphi)^n$ for all $\varphi \in L^2$ and all $n \in \mathbb{N}_0$. See Theorem 1.1.1 of [32] for details.

To apply the Besov-Kolmogorov type result from Lemma 3.1 we need to control high moments. This can be done with the help of the following proposition:

Theorem 3.11. *For all $p \in (0, \infty)$ there exists a constant $C_p > 0$ such that*

$$\mathbb{E}[|W_n(\varphi)|^p] \leqslant C_p^n \mathbb{E}[|W_n(\varphi)|^2]^{p/2} \leqslant C_p^n (n!)^{p/2} \|\varphi\|_{L^2(E^n)}^p$$

for all $\varphi \in L^2(E^n)$. This is called the hypercontractivity of the n-th chaos.

We will use martingale arguments to prove this theorem. We start by noting that for $\varphi \in L^2(E)$ the process $M_t^{\varphi} = \xi(\varphi \mathbb{I}_{[0,t]})$ is a continuous martingale in the filtration

$$\mathcal{F}_t = \sigma\{\xi(\varphi \mathbb{I}_{[0,s]}) \colon \varphi \in L^2(E), s \leqslant t\}.$$

The martingale property follows from the independence of ξ over disjoint intervals, and if φ is bounded the continuity can be shown by Kolmogorov's continuity criterion and then it follows for general φ by approximation together with Doob's L^2 inequality. Moreover, we have for $0 \leqslant s \leqslant t$

$$\begin{aligned}\mathbb{E}[(M_t^{\varphi})^2 - (M_s^{\varphi})^2 | \mathcal{F}_s] &= \mathbb{E}[(M_t^{\varphi} - M_s^{\varphi})^2 | \mathcal{F}_s] = \mathbb{E}[\xi(\varphi \mathbb{I}_{[s,t]})^2 | \mathcal{F}_s] = \|\varphi \mathbb{I}_{[s,t]}\|_{L^2(E)}^2 \\ &= \int_s^t \int_{\mathbb{T}^d} \varphi(r, x)^2 \mathrm{d}x \mathrm{d}r,\end{aligned}$$

from where we deduce that

$$\langle M^{\varphi} \rangle_t = \int_0^t \int_{\mathbb{T}^d} \varphi(r, x)^2 \mathrm{d}x \mathrm{d}r, \qquad \langle M^{\varphi}, M^{\psi} \rangle_t = \int_0^t \int_{\mathbb{T}^d} \varphi(r, x) \psi(r, x) \mathrm{d}x \mathrm{d}r.$$

We want to write elements of the n-th chaos as martingales, and for that purpose we need the following simple lemma, which shows that Hermite polynomials are intimately connected to continuous martingales.

Lemma 3.12. *Let M be a continuous local martingale. Then*

$$H_n(M_t, \langle M\rangle_t) = n\int_0^t H_{n-1}(M_s, \langle M\rangle_s)\mathrm{d}M_s.$$

Proof. We apply Ito's formula to $H_n(M_t, \langle M\rangle_t)$: Since $H_n(0,0)=0$ and $\left(\partial_t + \frac{1}{2}\partial_x^2\right)H_n \equiv 0$ by (3.6), we get

$$\begin{aligned} H_n(M_t, \langle M\rangle_t) =& \int_0^t \partial_x H_n(M_s, \langle M\rangle_s)\mathrm{d}M_s + \int_0^t \left(\partial_t + \frac{1}{2}\partial_x^2\right)H_n(M_s, \langle M\rangle_s)\mathrm{d}\langle M\rangle_s \\ =& n\int_0^t H_{n-1}(M_s, \langle M\rangle_s)\mathrm{d}M_s, \end{aligned}$$

where the last step follows from the first identity in (3.6). □

Consequently, for all $\varphi \in L^2(E)$ the process

$$W_n(\varphi^{\otimes n}\mathbb{I}_{[0,t]}^{\otimes n}) = n\int_0^t W_{n-1}\big(\varphi^{\otimes(n-1)}\mathbb{I}_{[0,s]}^{\otimes(n-1)}\big)\mathrm{d}M_s^{\varphi}, \qquad t \geqslant 0,$$

is a continuous martingale, and the quadratic covariation of two such martingales is

$$\begin{aligned} &\langle W_n(\varphi^{\otimes n}\mathbb{I}_{[0,\cdot]}^{\otimes n}), W_n(\psi^{\otimes n}\mathbb{I}_{[0,\cdot]}^{\otimes n})\rangle_t \\ =& n^2\int_E W_{n-1}\big(\varphi^{\otimes n}(z,\cdot)\mathbb{I}_{[0,s]}^{\otimes(n-1)}\big)W_{n-1}\big(\psi^{\otimes n}(z,\cdot)\mathbb{I}_{[0,s]}^{\otimes(n-1)}\big)\mathbb{I}_{[0,t]}(s)\mathrm{d}z, \end{aligned}$$

where we write $z=(s,x)\in E$. By taking linear combinations of such $\varphi^{\otimes n}$ and by approximation we deduce that for all $\varphi \in L^2(E^n)$ the process $W_n(\varphi\mathbb{I}_{[0,t]}^{\otimes n})$, $t \geqslant 0$, is a martingale with quadratic variation

$$\langle W_n(\varphi\mathbb{I}_{[0,\cdot]}^{\otimes n})\rangle_t = n^2\int_E W_{n-1}\big(\varphi(z,\cdot)\mathbb{I}_{[0,s]}^{\otimes(n-1)}\big)^2\mathbb{I}_{[0,t]}(s)\mathrm{d}z.$$

Proof. (of Theorem 3.11) For $p<2$ there is nothing to show, so let $p \geqslant 2$. By the Burkholder-Davis-Gundy inequality, see Exercise 3.2 below, together with the Minkowski inequality $\|\int_E(\dots)\mathrm{d}z\|_{L^{p/2}(\Omega)} \leqslant \int_E\|\dots\|_{L^{p/2}(\Omega)}\mathrm{d}z$ we have

$$\begin{aligned} \mathbb{E}[|W_n(\varphi)|^p] \leqslant& \mathbb{E}\Big[\sup_{t>0}|W_n(\varphi\mathbb{I}_{[0,t]}^{\otimes n})|^p\Big] \leqslant C_p\mathbb{E}\left[\left(n^2\int_E W_{n-1}\big(\varphi(z,\cdot)\mathbb{I}_{[0,s]}^{\otimes(n-1)}\big)^2\mathrm{d}z\right)^{p/2}\right] \\ \leqslant& C_p\left(n^2\int_E \mathbb{E}\big[\big|W_{n-1}\big(\varphi(z_1,\cdot)\mathbb{I}_{[0,s_1]}^{\otimes(n-1)}\big)\big|^p\big]^{2/p}\mathrm{d}z_1\right)^{p/2} \\ \leqslant& C_p^2\left(n^2(n-1)^2\int_E\int_E \mathbb{E}\big[\big|W_{n-2}\big(\varphi(z_1,z_2,\cdot)\mathbb{I}_{[0,s_2]}^{\otimes(n-2)}\big)\big|^p\big]^{2/p}\mathbb{I}_{s_2\leqslant s_1}\mathrm{d}z_2\mathrm{d}z_1\right)^{p/2} \\ \leqslant& \dots \leqslant C_p^n\left((n!)^2\int_E\dots\int_E|\varphi(z_1,\dots,z_n)|^2\mathbb{I}_{s_n\leqslant\dots\leqslant s_1}\mathrm{d}z_n\dots\mathrm{d}z_1\right)^{p/2} \\ =& C_p^n\left(n!\int_E\dots\int_E|\varphi(z_1,\dots,z_n)|^2\mathrm{d}z_n\dots\mathrm{d}z_1\right)^{p/2}, \end{aligned}$$

where in the last step we assumed without loss of generality that φ is symmetric in its n arguments. In that case the right hand side equals $C_p^n (n!)^{p/2} \|\varphi\|_{L^2(E^n)}^p$, and this completes the proof. □

Remark 3.13. It seems to have been important that there was a designated "time variable" in the computations above. But in fact we can use the arguments whenever ξ is a white noise on $L^2(E)$ and $E = I \times E'$ for some interval $I \subset \mathbb{R}$, for example if $E = \mathbb{T}^d$ we can interpret this as $E \simeq [0,1] \times \mathbb{T}^{d-1}$ and this gives us an "artificial" time variable to work with.

Similar martingale arguments also allow to prove hypercontractivity results in discrete settings, see for example Lemma 5.1 in [26].

Exercise 3.2. Let M be a continuous local martingale with $M_0 = 0$ and let $p \geqslant 2$. Show that

$$\mathbb{E}\Big[\sup_{s \leqslant t} |M_s|^p\Big] \lesssim \mathbb{E}\big[\langle M \rangle_t^{p/2}\big].$$

Hint: Apply Ito's formula to $|M_t|^p$.

Solution. We have by Ito's formula

$$|M_t|^p = \int_0^t p |M_s|^{p-2} M_s \mathrm{d} M_s + \frac{1}{2} \int_0^t p(p-1) |M_s|^{p-2} \mathrm{d} \langle M \rangle_s$$

Assume now that M is bounded. Then the stochastic integral on the right hand side is a true martingale, and we get

$$\mathbb{E}[|M_t|^p] \leqslant \frac{p(p-1)}{2} \mathbb{E}\Big[\sup_{s \leqslant t} |M_s|^{p-2} \langle M \rangle_t\Big] \leqslant \frac{p(p-1)}{2} \mathbb{E}\Big[\sup_{s \leqslant t} |M_s|^p\Big]^{(p-2)/p} \mathbb{E}\big[\langle M \rangle_t^{p/2}\big]^{2/p},$$

where in the second step we applied Hölder's inequality with $r = p/(p-2)$ and $q = p/2$. Now Doob's L^p-inequality gives

$$\mathbb{E}\Big[\sup_{s \leqslant t} |M_t|^p\Big] \lesssim \mathbb{E}[|M_t|^p] \lesssim \mathbb{E}\Big[\sup_{s \leqslant t} |M_s|^p\Big]^{(p-2)/p} \mathbb{E}\big[\langle M \rangle_t^{p/2}\big]^{2/p}$$

and thus we can bring $\mathbb{E}[\sup_{s \leqslant t} |M_s|^p]^{(p-2)/p}$ to the left hand side to estimate

$$\mathbb{E}\Big[\sup_{s \leqslant t} |M_t|^p\Big] \lesssim \mathbb{E}\big[\langle M \rangle_t^{p/2}\big].$$

This is still under the assumption that M is bounded, but for the general case we can use a stopping argument.

Corollary 3.14. *Let $d = 2$ and $T > 0$ and define*

$$Z_\varepsilon^{:n:}(t,x) := H_n(Z_\varepsilon(t,x), C_\varepsilon(t)), \qquad t \in [0,T].$$

Let $\alpha < 2$ and $\gamma < 1/2$. There exists $Z^{:n:} \in C_T^0 \mathscr{C}^{\alpha-2} \cap C_T^\gamma \mathscr{C}^{\alpha-3}$ such that for all $p \in [1,\infty)$

$$\lim_{\varepsilon \to 0} \mathbb{E}\big[\|Z_\varepsilon^{:n:} - Z^{:n:}\|_{C_T^0 \mathscr{C}^{\alpha-2}}^p + \|Z_\varepsilon^{:n:} - Z^{:n:}\|_{C_T^\gamma \mathscr{C}^{\alpha-3}}^p\big] = 0. \tag{3.7}$$

Proof. 1. We have

$$Z_\varepsilon(t,x) = \int_0^t \int_{\mathbb{T}^d} p^\varepsilon(t-s, x-y) \xi(s,y) \mathrm{d}y \mathrm{d}s = \xi(\mathbb{I}_{[0,t]} p^\varepsilon(t-\cdot, x-\cdot))$$

and in Lemma 3.7 we saw that $C_\varepsilon(t) = \|\mathbb{I}_{[0,t]} p^\varepsilon(t-\cdot, x-\cdot)\|_{L^2}^2$, so that

$$Z_\varepsilon^{:n:}(t,x) = W_n(\mathbb{I}_{[0,t]}^{\otimes n} p^\varepsilon(t-\cdot, x-\cdot)^{\otimes n}),$$

and by construction W_n is a continuous linear map, which means that we can exchange it with integrals and we obtain

$$\begin{aligned}\Delta_j Z_\varepsilon^{:n:}(t,x) &= \int_{\mathbb{T}^d} K_j(x-y) W_n(\mathbb{I}_{[0,t]}^{\otimes n} p^\varepsilon(t-\cdot, y-\cdot)^{\otimes n}) \mathrm{d}y \\ &= W_n\left(\int_{\mathbb{T}^d} K_j(x-y) \mathbb{I}_{[0,t]}^{\otimes n} p^\varepsilon(t-\cdot, y-\cdot)^{\otimes n} \mathrm{d}y\right).\end{aligned}$$

So by Theorem 3.11 we have

$$\begin{aligned}\mathbb{E}[|\Delta_j Z_\varepsilon^{:n:}(t,x)|^p]^{2/p} &\lesssim \int_{E^n} \left| \int_{\mathbb{T}^d} K_j(x-y) \prod_{j=1}^n [\mathbb{I}_{s_j \leqslant t} p^\varepsilon(t-s_j, y-x_j)] \mathrm{d}y \right|^2 \mathrm{d}z_n \ldots \mathrm{d}z_1 \\ &= \int_{E^n} \int_{\mathbb{T}^d} K_j(x-y_1) K_j(x-y_2) \\ &\quad \times \prod_{j=1}^n [\mathbb{I}_{s_j \leqslant t} p^\varepsilon(t-s_j, y_1-x_j) p^\varepsilon(t-s_j, y_2-x_j)] \mathrm{d}y_1 \mathrm{d}y_2 \mathrm{d}z_n \ldots \mathrm{d}z_1 \\ &= \int_{[0,t]^n} \int_{\mathbb{T}^d} K_j(x-y_1) K_j(x-y_2) \prod_{j=1}^n \tilde{p}^\varepsilon(2(t-s_j), y_2-y_1) \mathrm{d}y_1 \mathrm{d}y_2 \mathrm{d}s_n \ldots \mathrm{d}s_1 \\ &= \int_{\mathbb{T}^d} K_j(x-y_1) K_j(x-y_2) \left(\int_0^t \tilde{p}^\varepsilon(2s, y_2-y_1) \mathrm{d}s\right)^n \mathrm{d}y_1 \mathrm{d}y_2,\end{aligned}$$

where $\mathscr{F}\tilde{p}^\varepsilon(t,k) = \varphi_\varepsilon(k)^2 \mathscr{F}p(t,k)$ and we used the semigroup property $p(r,\cdot) * p(s,\cdot) = p(r+s,\cdot)$. Now we apply Parseval's identity and obtain

$$\begin{aligned}&\mathbb{E}[|\Delta_j Z_\varepsilon^{:n:}(t,x)|^p]^{2/p} \\ &\lesssim \sum_{k_1,\ldots,k_n} \rho_j(k_1+\cdots+k_n)^2 \prod_{j=1}^n \left(\int_0^t |\mathscr{F}\tilde{p}^\varepsilon(2s,k_j)|^2 \mathrm{d}s\right) \\ &= n! \sum_{k_1,\ldots,k_n} \rho_j(k_1+\cdots+k_n)^2 \varphi_\varepsilon(k_1)^2 \ldots \varphi_\varepsilon(k_n)^2 \prod_{j=1}^n \left(\mathbb{I}_{k_j=0}\, t + \mathbb{I}_{k_j \neq 0} \frac{1-e^{-4|2\pi k_j|^2 t}}{4|2\pi k_j|^2}\right) \\ &\lesssim_{n,t} \sum_{k_1,\ldots,k_n} \rho_j(k_1+\cdots+k_n)^2 \prod_{j=1}^n \frac{1}{1+|k_j|^2}.\end{aligned}$$

2. Next, we use Lemma 3.15 below and get for all $\delta > 0$ with $n\delta < 2$ and with $\langle k \rangle = \sqrt{1+|k|^2}$

$$\begin{aligned}&\sum_{k_1,\ldots,k_n} \rho_j(k_1+\cdots+k_n)^2 \prod_{j=1}^n \langle k_j \rangle^{-2} \\ &\lesssim \sum_{\ell_{n-1}} \sum_{k_1,\ldots,k_{n-2}} \rho_j(k_1+\cdots+k_{n-2}+\ell_{n-1})^2 \prod_{j=1}^{n-2} \langle k_j \rangle^{-2} \sum_{k_{n-1}} \langle \ell_{n-1} - k_{n-1} \rangle^{-2+\delta} \langle k_{n-1} \rangle^{-2+\delta} \\ &\lesssim \sum_{\ell_{n-1}} \sum_{k_1,\ldots,k_{n-2}} \rho_j(k_1+\cdots+k_{n-2}+\ell_{n-1})^2 \prod_{j=1}^{n-2} \langle k_j \rangle^{-2} \langle \ell_{n-1} \rangle^{-2+2\delta} \\ &\lesssim \ldots \lesssim \sum_{\ell_1} \rho_j(\ell_1)^2 \langle \ell_1 \rangle^{-2+n\delta} \lesssim 2^{2j} 2^{j(-2+n\delta)} = 2^{jn\delta}.\end{aligned}$$

3. Thus, we have shown that $\mathbb{E}[|\Delta_j Z_\varepsilon^{:n:}(t,x)|^p]^{1/p} \lesssim 2^{jn\delta/2}$ for all $\delta > 0$ and $p > 0$. Combining the above analysis with similar arguments as in the proof of Lemma 3.3, we also get

$$\mathbb{E}[|\Delta_j Z_\varepsilon^{:n:}(t,x) - \Delta_j Z_\varepsilon^{:n:}(s,x)|^p]^{1/p} \lesssim 2^{j(n\delta/2+\lambda)}|t-s|^{\lambda/2}$$

for all $\lambda \in [0,1]$, and therefore it follows from Lemma 3.1 that we can bound

$$\mathbb{E}\big[\|Z_\varepsilon^{:n:}\|^p_{C_T^0 \mathscr{C}^{\alpha-2}} + \|Z_\varepsilon^{:n:}\|^p_{C_T^\gamma \mathscr{C}^{\alpha-3}}\big]$$

uniformly in ε. Since moreover $\varphi_{\varepsilon'}(k)^2 - \varphi_\varepsilon(k)^2 \to 0$ as $\varepsilon, \varepsilon' \to 0$ for all k, and the difference is uniformly bounded, we see from the dominated convergence theorem that $(Z_\varepsilon^{:n:})_\varepsilon$ is a Cauchy sequence in $L^p(C_T^0 \mathscr{C}^{\alpha-2} \cap C_T^\gamma \mathscr{C}^{\alpha-3})$. Therefore it converges to the limit

$$Z^{:n:}(t,x) = W_n(\mathbb{I}_{[0,t]}^{\otimes n} p(t-\cdot, x-\cdot)^{\otimes n}),$$

which is of course not defined pointwise in x but only as a distribution. □

Lemma 3.15. *Let $\alpha, \beta < d$ be such that $\alpha + \beta > d$ and write $\langle k \rangle = \sqrt{1+|k|^2}$. Then*

$$\sum_{k' \in \mathbb{Z}^d} \langle k - k' \rangle^{-\alpha} \langle k' \rangle^{-\beta} \lesssim \langle k \rangle^{d-\alpha-\beta}.$$

Proof. See [30], Lemma 4.1. □

Remark 3.16. Most of the proof is independent of the dimension. However, in step 2. we crucially used that $d = 2$. Convince yourself that in $d = 3$ the same construction works for $Z^{:2:}$ but not for $Z^{:3:}$, and in $d = 4$ it does not even work for $Z^{:2:}$.

For much more details on the Wiener-Ito chaos and related concepts see [25, 32].

3.5 Back to the Φ_2^4 equation

In Section 3.3 we saw that if we were able to construct Z^2, Z^3 in $d = 2$, then we would be able to solve the Φ_2^4 equation by setting $\phi = Z + v$, where

$$(\partial_t - \Delta)v = -(Z^3 + 3Z^2 v + 3Zv^2 + v^3), \qquad v(0) = \phi_0.$$

However, then we saw in Section 3.4 that we can only construct the renormalized products

$$Z^{:n:}(t,x) = \lim_{\varepsilon \to 0} Z_\varepsilon^{:n:}(t,x),$$

where the convergence is in $C_T \mathscr{C}^{\alpha-2} \cap C_T^\gamma \mathscr{C}^{\alpha-3}$ whenever $\alpha < 2$, $\gamma < 1/2$, and $T > 0$. So it is natural to replace the equation for v with the "renormalized" equation

$$(\partial_t - \Delta)v = -(Z^{:3:} + 3Z^{:2:} v + 3Zv^2 + v^3), \qquad v(0) = \phi_0,$$

which can be solved by the same arguments as in Section 3.3. Of course, this changes the equation and the question is whether we still have a useful interpretation for the new equation. For that purpose we go to the approximations $Z_\varepsilon^{:n:}$. Since we saw that the solution v depends continuously on the data, we have $v = \lim_{\varepsilon \to 0} v_\varepsilon$, where

$$\begin{aligned}(\partial_t - \Delta)v_\varepsilon &= -(Z_\varepsilon^{:3:} + 3Z_\varepsilon^{:2:} v_\varepsilon + 3Z_\varepsilon v_\varepsilon^2 + v_\varepsilon^3) \\ &= -[(Z_\varepsilon^3 - 3C_\varepsilon Z_\varepsilon) + 3(Z_\varepsilon^2 - C_\varepsilon)v_\varepsilon + 3Z_\varepsilon v_\varepsilon^2 + v_\varepsilon^3] \\ &= -[Z_\varepsilon^3 + 3Z_\varepsilon^2 v_\varepsilon + 3Z_\varepsilon v_\varepsilon^2 + v_\varepsilon^3 - 3C_\varepsilon(Z_\varepsilon + v_\varepsilon)] \\ &= -(\phi_\varepsilon^3 - 3C_\varepsilon \phi_\varepsilon),\end{aligned}$$

where $\phi_\varepsilon = Z_\varepsilon + v_\varepsilon$. In other words, $\phi = Z + v$ is given as the limit of ϕ_ε solving the renormalized equation

$$(\partial_t - \Delta)\phi_\varepsilon = -(\phi_\varepsilon^3 - 3C_\varepsilon\phi_\varepsilon) + \xi_\varepsilon,$$

where $\xi_\varepsilon = \varphi_\varepsilon * \xi$.

Exercise 3.3. It may seem like a small miracle that replacing Z_ε^n by $Z_\varepsilon^{:n:}$ leads to a closed equation for ϕ_ε. In fact, this follows from a general result on Hermite polynomials: Show that for $t > 0$ and $x, y \in \mathbb{R}$ and $n \in \mathbb{N}_0$ we have

$$H_n(x + y, t) = \sum_{k=0}^{n} \binom{n}{k} H_k(x, t) y^{n-k}.$$

Solution. For $n = 0$ this is obvious. Then

$$\begin{aligned} H_n(x+y,t) =& (x+y)H_{n-1}(x+y,t) - t\partial_x H_{n-1}(x+y,t) \\ =& ((x+y) - t\partial_x)\sum_{k=0}^{n-1} \binom{n-1}{k} H_k(x,t) y^{n-1-k} \\ =& \sum_{k=0}^{n-1} \binom{n-1}{k} [H_k(x,t)y^{n-k} - H_{k+1}(x,t)y^{n-1-k}] \\ =& y^n + \sum_{k=1}^{n-1} \left[\binom{n-1}{k} + \binom{n-1}{k-1}\right] H_k(x,t)y^{n-k} + H_n(x,t) \\ =& \sum_{k=0}^{n} \binom{n}{k} H_k(x,t) y^{n-k}. \end{aligned}$$

Combining this exercise together with a similar analysis as in Section 3.3 we can construct for all n the limit ϕ of ϕ_ε solving

$$(\partial_t - \Delta)\phi_\varepsilon = -\phi_\varepsilon^{:n:} + \xi_\varepsilon := -H_n(\phi_\varepsilon, C_\varepsilon) + \xi_\varepsilon,$$

i.e. the Φ_2^{n+1} model.

Remark 3.17. We have some freedom how to choose the renormalization, because of course also the equation with renormalization $C_\varepsilon + \lambda$ for $\lambda \in \mathbb{R}$ converges to a limit. Our renormalization depends on time:

$$(\partial_t - \Delta)\phi_\varepsilon(t,x) = -(\phi_\varepsilon^3(t,x) - 3C_\varepsilon(t)\phi_\varepsilon(t,x)) + \xi_\varepsilon(t,x),$$

where

$$C_\varepsilon(t) = t + \sum_{k \neq 0} \frac{|\varphi(\varepsilon k)|^2}{2|2\pi k|^2}(1 - e^{-2|2\pi k|^2 t}).$$

To obtain a more natural time-independent renormalization, note that with $c_\varepsilon = \sum_{k\neq 0} \frac{|\varphi(\varepsilon k)|^2}{2|2\pi k|^2}$ we have for all $t > 0$

$$\lim_{\varepsilon \to 0} (C_\varepsilon(t) - c_\varepsilon) = t,$$

i.e. the difference of $C_\varepsilon(t)$ and the constant c_ε converges to a finite limit. This suggests that also the solution $\tilde{\phi}_\varepsilon$ to

$$(\partial_t - \Delta)\tilde{\phi}_\varepsilon = -(\tilde{\phi}_\varepsilon^3 - 3c_\varepsilon\tilde{\phi}_\varepsilon) + \xi_\varepsilon$$

should converge, and this provides a more natural renormalization because the original equation was time-homogeneous in law and thus also the renormalized equation should be time-homogeneous in law.

The only problem is that while for all $t > 0$ the convergence $C_\varepsilon(s) - c_\varepsilon \to s$ is uniform on (t, ∞), there is a divergence at $t = 0$ (since $C_\varepsilon(0) = 0$). This can be cured by allowing the solution ϕ_ε to have a singularity at $t = 0$, i.e. by considering the norm

$$\sup_{t \in [0,T]} t^\gamma \|\phi_\varepsilon(t)\|_{\mathscr{C}^\alpha} = \|(t \mapsto t^\gamma \phi_\varepsilon(t))\|_{C_T \mathscr{C}^\alpha}.$$

A closely related problem is that in Section 3.3 we had to take an initial condition $\phi_0 \in \mathscr{C}^\alpha$ while the solution $\phi = Z + v$ only has regularity $\phi \in C_T \mathscr{C}^{\alpha - 2}$. It is more natural to take ϕ_0 of the same regularity as Z, and this also can be done by allowing for a singularity at $t = 0$.

In fact, once we are able to solve the equation with more singular initial conditions, we can also take a stationary version of Z, so that the expectation $C_\varepsilon(t)$ becomes independent of t. Of course, then we no longer have $Z = 0$ and thus the new initial condition for v becomes $\phi_0 - Z(0)$, which is typically in $\mathscr{C}^{\alpha - 2}$ since $Z(0) \in \mathscr{C}^{\alpha - 2}$.

4 The Φ_3^4 equation and paracontrolled distributions

Let us now consider the case $d = 3$. Then $Z \in C_T \mathscr{C}^{-1/2-}$ by Lemma 3.3, so if we assume as for Φ_2^4 that we can construct the renormalized powers $Z^{:n:}$ with their canonical regularities, then we would get $Z^{:2:} \in C_T \mathscr{C}^{-1-}$ and $Z^{:3:} \in C_T \mathscr{C}^{-3/2-}$. Therefore, we expect that by Schauder estimates the solution v to

$$(\partial_t - \Delta) v = -(Z^{:3:} + 3 Z^{:2:} v + 3 Z v^2 + v^3)$$

has 2 degrees of regularity more than $Z^{:3:}$, i.e. $v \in C_T \mathscr{C}^{1/2-}$. But now we have a problem, because the product $Z^{:2:} v$ is ill-defined if $Z^{:2:} \in C_T \mathscr{C}^{-1-}$ and $v \in C_T \mathscr{C}^{1/2-}$ since the sum of the regularities is well below 0. We try to overcome this problem by cancelling again the most irregular term on the right hand side, let us write

$$Y = \int_0^t P_{t-s}(-Z^{:3:}(s)) \mathrm{d}s,$$

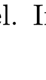
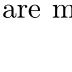
so that $(\partial_t - \Delta) Y = -Z^{:3:}$, and $v^{(2)} = \phi - Z - Y$. Then $v = Y + v^{(2)}$, and therefore

$$(\partial_t - \Delta) v^{(2)} = -(3 Z^{:2:}(Y + v^{(2)}) + 3 Z (Y + v^{(2)})^2 + (Y + v^{(2)})^3).$$

The right hand side contains the product $Z^{:2:}(Y + v^{(2)})$, and we saw that a product has at best the lower of the two regularities of its factors, i.e. $Z^{:2:}(Y + v^{(2)}) \in C_T \mathscr{C}^{-1-}$ should have the same regularity as $Z^{:2:}$. So by Schauder estimates we expect $v^{(2)} \in C_T \mathscr{C}^{1-}$, and this means that the product $Z^{:2:}(Y + v^{(2)})$ is ill-defined.

Unfortunately, this problem is more substantial than before, and we cannot solve it by expanding $v^{(2)}$ further, say by setting $v^{(2)} = U + v^{(3)}$ for an arbitrary U: In that case we would still get the term $3 Z^{:2:} v^{(3)}$ in the equation for $v^{(3)}$, and by the same arguments as above we expect $v^{(3)} \in C_T \mathscr{C}^{1-}$ so that the product $Z^{:2:} v^{(3)}$ is ill-defined.

The solution is to use a more complicated expansion of $v^{(2)}$, based on paraproducts. But before that, let us introduce more efficient notation: Already in the considerations above we needed $Z, Z^{:n:}, Y$, and we would need even more letters in the expansions below. To make the variable names more transparent and compact, we use trees to denote the polynomials of ξ that appear in the expansion of ϕ. We use a dot $\bullet$ to represent an instance of ξ, and a line $\backslash$ to denote the convolution with the heat kernel. If two trees are multiplied with each other, we simply join them. So for example

$Z =$ , $Z^2 =$ , $Z^3 =$ , $Y = -$, $YZ = -$.

In some cases we will also write Υ^2 etc. if this seems simpler than writing out the full tree. For now we do not worry about the fact that these trees are not well defined analytically and that we may have to renormalize them, we will address this point after developing the paracontrolled solution theory under the assumption that each tree in our expansion is given with its canonical regularity. At this point the attentive reader may recall the observation from Remark 3.16 that $Z^{:3:}$ cannot be constructed in $d=3$. This will not be a problem because we only encounter Υ in the expansions below, and while Ψ is not a continuous function of time with values in a space of distributions, it is a space-time distribution and we can construct its convolution against the heat kernel. We expand on this after developing the analytic theory under the assumption that all trees are well defined and have the right regularity.

4.1 The main commutator estimate

With our new notation, the expansion for $v := v^{(2)} = \phi - \mathsf{I} + \Upsilon$ can be rewritten with $\mathscr{L} = \partial_t - \Delta$ as

$$\begin{aligned}\mathscr{L}v &= -\Big(3\vee\Big(-\Upsilon+v\Big)+3\mathsf{I}\,\Big(-\Upsilon+v\Big)^2+\Big(-\Upsilon+v\Big)^3\Big)\\ &= -3\Big(v-\Upsilon\Big)\vee+\tau_0+\tau_1 v+\tau_2 v^2-v^3,\end{aligned}$$

where we wrote

$$\tau_0 := -3\mathsf{I}\times\Upsilon^2+\Upsilon^3, \qquad \tau_1 := 6\Upsilon\!\vee-3\Upsilon^2, \qquad \tau_2 := -3\mathsf{I}+3\Upsilon.$$

Let us use the underbrace $\underbrace{X}_{\alpha}$ to express that we expect $X \in C_T\mathscr{C}^\alpha$. Then

$$\mathscr{L}v = -3(\underbrace{v}_{1-}-\underbrace{\Upsilon}_{\frac{1}{2}-})\underbrace{\vee}_{-1-}+\underbrace{\tau_0}_{-\frac{1}{2}-}+\underbrace{\tau_1}_{-\frac{1}{2}-}\underbrace{v}_{1-}+\underbrace{\tau_2}_{-\frac{1}{2}-}\underbrace{v^2}_{1-}-\underbrace{v^3}_{1-},$$

which as we discussed before means that $v\vee$ is ill-posed. Note however that if τ_2, τ_1, τ_0 are given with their canonical regularities, then the other products involving v are all well-defined. The idea for dealing with the ill-defined product is to decompose it further with the help of paraproducts: Ignoring the fact that the resonant product is ill-defined, we would expect to have the regularities

$$\Big(v-\Upsilon\Big)\vee = \underbrace{\Big(v-\Upsilon\Big)\prec\vee}_{-1-}+\underbrace{\Big(v-\Upsilon\Big)\succ\vee}_{-\frac{1}{2}-}+\underbrace{\Big(v-\Upsilon\Big)\odot\vee}_{-\frac{1}{2}-},$$

and therefore

$$\mathscr{L}v = -3\Big(v-\Upsilon\Big)\prec\vee+C_T\mathscr{C}^{-1/2-},$$

by which we mean that $\mathscr{L}v+3\Big(v-\Upsilon\Big)\prec\vee \in C_T\mathscr{C}^{-1/2-}$. In other words, we expect $\mathscr{L}v$ to be given by a paraproduct plus a more regular remainder. We will see below that the convolution with the heat kernel in a certain sense commutes with the paraproduct (modulo a smoother remainder term), and this leads to the *paracontrolled ansatz*

$$\phi = \mathsf{I}-\Upsilon+v, \qquad v = v'\prec\Upsilon+v^\sharp, \qquad v' \in C_T\mathscr{C}^{1/2-}, \qquad v \in C_T\mathscr{C}^{1-}, \qquad v^\sharp \in C_T\mathscr{C}^{3/2-}.$$

Why should this be useful? First note that we can now expand the ill-defined resonant product further as

$$v \odot \vee = \underbrace{(v' \varolessthan \mathsf{Y})}_{1-} \odot \underbrace{\vee}_{-1-} + \underbrace{v^\sharp}_{\frac{3}{2}-} \odot \underbrace{\vee}_{-1-},$$

and the second term on the right hand side is well defined. Therefore, it remains to understand the resonant product $(v' \varolessthan \mathsf{Y}) \odot \vee$. Recall that we saw in Example 2.8 that the paraproduct $v' \varolessthan \mathsf{Y}$ is a "frequency modulation" of Y and it looks like Y on small scales. But the difficulty we have with defining $(v' \varolessthan \mathsf{Y}) \odot \vee$ comes from interactions of small scale contributions of $v' \varolessthan \mathsf{Y}$ and $\vee$ which in the product might create diverging resonances. So if we understand how the small scale contributions of Y interact with those of $\vee$ and that no diverging resonances arise in the product, then we might also hope that $v' \varolessthan \mathsf{Y}$ has no diverging resonances with $\vee$. This can be made precise with the help of the following commutator estimate, the main technical result in paracontrolled distributions:

Lemma 4.1. *Assume that $\alpha \in (0,1)$ and $\beta, \gamma \in \mathbb{R}$ are such that $\alpha + \beta + \gamma > 0$ and $\beta + \gamma < 0$. Then the trilinear operator on $\mathscr{S}^3$, defined by*

$$C(f, g, h) = ((f \varolessthan g) \odot h) - f(g \odot h),$$

satisfies

$$\|C(f, g, h)\|_{\alpha+\beta+\gamma} \lesssim \|f\|_\alpha \|g\|_\beta \|h\|_\gamma, \qquad (4.1)$$

and can thus be canonically extended to a bounded trilinear operator from $\mathscr{C}^\alpha \times \mathscr{C}^\beta \times \mathscr{C}^\gamma$ to $\mathscr{C}^{\alpha+\beta+\gamma}$.

Remark 4.2. It would be aesthetically more pleasing to have the same estimate for

$$(f \varolessthan g) \odot h - f \varolessthan (g \odot h),$$

but unfortunately this is not true.

To prove Lemma 4.1 we need the following auxiliary result:

Lemma 4.3. *Let $\alpha \in (0,1)$ and $\beta \in \mathbb{R}$ and let $f \in \mathscr{C}^\alpha$ and $g \in \mathscr{C}^\beta$. Then for all $j \geqslant -1$*

$$\|\Delta_j(f \varolessthan g) - f \Delta_j g\|_{L^\infty} \lesssim 2^{-j(\alpha+\beta)} \|f\|_\alpha \|g\|_\beta.$$

Proof. We give the proof on $\mathbb{R}^d$, on $\mathbb{T}^d$ it is slightly more complicated because there is no exact scaling relation $K_j(x) = 2^{jd} K_0(2^j x)$, but we can use Poisson summation to overcome this problem. We have

$$\begin{aligned}\Delta_j(f \varolessthan g) - f \Delta_j g &= \sum_{i \sim j} (\Delta_j(\Delta_{<i-1} f \Delta_i g) - f \Delta_j \Delta_i g) \\ &= \sum_{i \sim j} (\Delta_j(\Delta_{<i-1} f \Delta_i g) - \Delta_{<i-1} f \Delta_j \Delta_i g) - \sum_{i \sim j} \Delta_{\geqslant i-1} f \Delta_j \Delta_i g.\end{aligned}$$

Since $\alpha > 0$, the second term on the right hand side is easily estimated by

$$\left\| \sum_{i \sim j} \Delta_{\geqslant i-1} f \Delta_j \Delta_i g \right\|_{L^\infty} \lesssim \sum_{i \sim j} 2^{-i\alpha} \|f\|_\alpha 2^{-i\beta} \|g\|_\beta \simeq 2^{-j(\alpha+\beta)} \|f\|_\alpha \|g\|_\beta.$$

For the remaining term we write the action of Δ_j as a convolution:

$$
\begin{aligned}
&|(\Delta_j(\Delta_{<i-1}f\Delta_i g)-\Delta_{<i-1}f\Delta_j\Delta_i g)(x)|\\
=&\left|\int K_j(x-y)(\Delta_{<i-1}f(y)-\Delta_{<i-1}f(x))\Delta_i g(y)\mathrm{d}y\right|\\
\lesssim&\int |K_j(x-y)|\max_{|\mu|=1}\|\partial^\mu\Delta_{<i-1}f\|_{L^\infty}|x-y|\|\Delta_i g\|_{L^\infty}\mathrm{d}y\\
\lesssim&2^{i(1-\alpha-\beta)}\max_{|\mu|=1}\|\partial^\mu f\|_{\alpha-1}\|g\|_\beta\int |K_j(x-y)||x-y|\mathrm{d}y\\
\lesssim&2^{i(1-\alpha-\beta)}\|f\|_\alpha\|g\|_\beta\int |K_j(y)||y|\mathrm{d}y,
\end{aligned}
$$

where we used that $\alpha-1<0$. If $j=-1$ this estimate is sufficient. For $j\geqslant 0$ we have

$$\int |K_j(y)||y|\mathrm{d}y=\int |2^{jd}K_0(2^j y)|2^{-j}|2^j y|\mathrm{d}y=2^{-j}\int |K_0(y)|\times|y|\mathrm{d}y\simeq 2^{-j},$$

from where our claim follows. □

Proof. (of Lemma 4.1) We have

$$
\begin{aligned}
C(f,g,h)=&\sum_{|i-j|\leqslant 1}[\Delta_i(f\circledcirc g)-f\Delta_i g]\Delta_j h\\
=&\sum_{|i-j|\leqslant 1}[\Delta_i(\Delta_{\lesssim i}f\circledcirc g)-\Delta_{\lesssim i}f\Delta_i g]\Delta_j h+\sum_{|i-j|\leqslant 1}\Delta_{\gtrsim i}f\Delta_i g\Delta_j h\\
=:&C_1(f,g,h)+C_2(f,g,h),
\end{aligned}
$$

where C_1 and C_2 are defined by the equality. Note that for fixed i the term

$$\sum_j \mathbb{I}_{|i-j|\leqslant 1}[\Delta_i(\Delta_{\lesssim i}f\circledcirc g)-\Delta_{\lesssim i}f\Delta_i g]\Delta_j h$$

has spectral support in a ball $2^i\mathscr{B}$. Moreover, since $\alpha\in(0,1)$ we get from Lemma 4.3

$$
\begin{aligned}
&\left\|\sum_j \mathbb{I}_{|i-j|\leqslant 1}[\Delta_i(\Delta_{\lesssim i}f\circledcirc g)-\Delta_{\lesssim i}f\Delta_i g]\Delta_j h\right\|_{L^\infty}\\
\lesssim&\sum_j \mathbb{I}_{|i-j|\leqslant 1}2^{-i(\alpha+\beta)}\|\Delta_{\lesssim i}f\|_\alpha\|g\|_\beta 2^{-j\gamma}\|h\|_\gamma\\
\lesssim&2^{-j(\alpha+\beta+\gamma)}\|f\|_\alpha\|g\|_\beta\|h\|_\gamma,
\end{aligned}
$$

so since $\alpha+\beta+\gamma>0$ the claimed regularity for $C_1(f,g,h)$ follows from Lemma 2.6. Let us get to $C_2(f,g,h)$:

$$C_2(f,g,h)=\sum_{|i-j|\leqslant 1}\Delta_{\gtrsim i}f\Delta_i g\Delta_j h=\sum_k\left(\sum_{|i-j|\leqslant 1}\mathbb{I}_{i\lesssim k}\Delta_k f\Delta_i g\Delta_j h\right),$$

and the term inside the brackets has spectral supprt in a ball $2^k\mathscr{B}$. Moreover, since $\beta+\gamma<0$,

$$
\begin{aligned}
\left\|\sum_{|i-j|\leqslant 1}\mathbb{I}_{i\lesssim k}\Delta_k f\Delta_i g\Delta_j h\right\|_{L^\infty}\lesssim&2^{-k\alpha}\|f\|_\alpha\sum_{i\lesssim k}2^{-i(\beta+\gamma)}\|g\|_\beta\|h\|_\gamma\\
\lesssim&2^{-k(\alpha+\beta+\gamma)}\|f\|_\alpha\|g\|_\beta\|h\|_\gamma,
\end{aligned}
$$

and now we use once more that $\alpha+\beta+\gamma>0$ and apply Lemma 2.6 to conclude the proof. □

So if ϕ satisfies the paracontrolled ansatz, our commutator estimate suggests to define

$$\left(v'\varolessthan\Upsilon\right)\odot\vee=C(\underbrace{v'}_{\frac{1}{2}-},\underbrace{\Upsilon}_{1-},\underbrace{\vee}_{-1-})+\underbrace{v'}_{\frac{1}{2}-}\underbrace{\left(\Upsilon\odot\vee\right)}_{0-}.$$

The sum of the regularities of the arguments of C is strictly positive, and therefore $C(v', \Upsilon,\vee)$ is well defined and in $C_T\mathscr{C}^{1/2-}$ by Lemma 4.1. For the remaining term, the resonant product $\Upsilon\odot\vee$ is still not defined, but if we make it part of the data, call it $\Upsilon_{\vee}$, and assume it has its canonical regularity $\Upsilon_{\vee}\in C_T\mathscr{C}^{0-}$, then $v'\Upsilon_{\vee}$ is well defined and in $C_T\mathscr{C}^{0-}$.

In the following we make this rigorous by defining a suitable Banach space of paracontrolled distributions in which we can set up a fixed point iteration to solve the Φ^4_3 equation. To not obscure the presentation with technicalities we first consider a simplified linear equation which, from the regularity analysis point of view, contains all the difficulties of the Φ^4_3 equation.

4.2 A linearized Φ^4_3 equation

Consider now the equation

$$(\partial_t-\Delta)v=\left(v+\Psi\right)\vee.$$

In the previous discussion we considered function spaces of the type $C_T\mathscr{C}^\alpha$, thus we only quantified the space regularity. It turns out however that we also need to make more precise assumptions on the time regularity. For that purpose we define the following function spaces:

Definition 4.4. *For $\alpha\in(0,2]$ and $T>0$ we set*

$$\mathscr{L}^\alpha_T:=C^{\alpha/2}_T L^\infty\cap C_T\mathscr{C}^\alpha,$$

equipped with the norm $\|u\|_{\mathscr{L}^\alpha_T}=\|u\|_{C^{\alpha/2}_T}+\|u\|_{C_T\mathscr{C}^\alpha}$. We also write

$$\mathscr{L}^\alpha=C^{\alpha/2}_{\mathrm{loc}}(\mathbb{R},L^\infty)\cap C(\mathbb{R},\mathscr{C}^\alpha),$$

with the obvious definition of the space of locally Hölder continuous functions $C^{\alpha/2}_{\mathrm{loc}}(\mathbb{R},L^\infty)$.

We assume that $\vee\in C(\mathbb{R},\mathscr{C}^{-1-})$, $\Psi\in\mathscr{L}^{1/2-}$, and also that the resonant product $\Psi_{\vee}=\Psi\odot\vee\in C(\mathbb{R},\mathscr{C}^{-1/2-})$ is given. We define

$$\Upsilon(t)=\int_0^t P_{t-s}\vee(s)\mathrm{d}s,\qquad t\geqslant 0,$$

which as we will see below is in $\mathscr{L}^{1-}_T$ for all $T>0$, and we also assume that $\Upsilon_{\vee}=\Upsilon\odot\vee\in C(\mathbb{R},\mathscr{C}^{0-})$ is given. We parametrize the regularities as $1/2-=\alpha$, $-1-=2\alpha-2$, $-1/2-=\alpha-1$, $0-=2\alpha-1$ for $\alpha\in(2/5,1/2)$ which will be fixed in what follows. We write

$$\mathcal{T}^\alpha=\mathscr{L}^\alpha\times C\mathscr{C}^{2\alpha-2}\times C\mathscr{C}^{\alpha-1}\times C\mathscr{C}^{2\alpha-1}$$

and

$$\mathbb{Z}=(\Psi,\vee,\Psi_{\vee},\Upsilon_{\vee})$$

for a generic element of $\mathcal{T}^\alpha$, as well as for $T>0$

$$\|\mathbb{Z}\|_{\mathcal{T}^\alpha_T}:=\|\Psi\|_{\mathscr{L}^\alpha_T}+\|\vee\|_{C_T\mathscr{C}^{2\alpha-2}}+\|\Psi_{\vee}\|_{C_T\mathscr{C}^{\alpha-1}}+\|\Upsilon_{\vee}\|_{C_T\mathscr{C}^{2\alpha-1}},$$

where we implicitly restrict all the functions inside the norms to the time interval $[0, T]$.

We also need a modified version of the paraproduct:

Definition 4.5. *We define the modified paraproduct as*

$$u \prec\!\!\!\prec v(t) = \int_0^t P_{t-s}(u(s) \varolessthan \mathscr{L} v(s)) \mathrm{d}s + P_t(u(0) \varolessthan v(0)),$$

whenever this is well defined.

This definition is due to Bailleul, Bernicot and Frey [5], although they did not include the term $P_t(u(0) \varolessthan v(0))$, which we will need to compare $u \prec\!\!\!\prec v$ and $u \varolessthan v$. They beauty of the definition is that by construction $\mathscr{L}(u \prec\!\!\!\prec v) = u \varolessthan \mathscr{L} v$, and this will be very useful for deriving paracontrolled Schauder estimates below.

Definition 4.6. *Let $\beta \in (2/5, \alpha]$ and $T > 0$. We say that*

$$(v, v', v^\sharp) \in \mathscr{L}_T^{\alpha+\beta} \times \mathscr{L}_T^\alpha \times \mathscr{L}_T^{2\alpha+\beta}$$

is paracontrolled by $\mathbb{Z}$ and write $(v, v', v^\sharp) \in \mathscr{D}_T^\beta \mathbb{Z}$ if

$$v = v' \prec\!\!\!\prec \text{Y} + v^\sharp.$$

If there is no ambiguity about v' and $v^\sharp$, we also write $v \in \mathscr{D}_T^\beta \mathbb{Z}$, or $v \in \mathscr{D}_T^\beta$ if there is no ambiguity about $\mathbb{Z}$. We set

$$\|v\|_{\mathscr{D}_T^\beta} = \|v\|_{\mathscr{L}_T^{\alpha+\beta}} + \|v'\|_{\mathscr{L}_T^\beta} + \|v^\sharp\|_{\mathscr{L}_T^{2\alpha+\beta}},$$

and if $\tilde{v} \in \mathscr{D}_T \tilde{\mathbb{Z}}$, then we also define

$$\|v - \tilde{v}\|_{\mathscr{D}_T^\beta} = \|v - \tilde{v}\|_{\mathscr{L}_T^{\alpha+\beta}} + \|v' - \tilde{v}'\|_{\mathscr{L}_T^\beta} + \|v^\sharp - \tilde{v}^\sharp\|_{\mathscr{L}_T^{2\alpha+\beta}}.$$

Note that $\|v - \tilde{v}\|_{\mathscr{D}_T^\beta}$ is only formal notation, this is not a norm since v and $\tilde{v}$ do not even live in the same space.

Remark 4.7. We sometimes call v' the derivative and $v^\sharp$ the remainder. Note that in general v' and $v^\sharp$ are not uniquely determined by v and $\mathbb{Z}$: Consider e.g. arbitrary v, Y, $v' \in \mathscr{L}_T^{2\alpha+\beta}$, then we always have $v^\sharp = v - v' \prec\!\!\!\prec \text{Y} \in \mathscr{L}_T^{2\alpha+\beta}$. So we really have to keep track of the tuple $(v, v', v^\sharp)$.

Our aim is now to make sense of the product $(v + \Psi) \vee$ for $v \in \mathscr{D}_T^\beta$. We use the paracontrolled structure and get

$$(v + \Psi) \vee := (v + \Psi) \varolessthan \vee + (v + \Psi) \varogreaterthan \vee + v^\sharp \odot \vee + (v' \prec\!\!\!\prec \text{Y}) \odot \vee + \Psi\!\!\vee.$$

If instead of $v' \prec\!\!\!\prec \text{Y}$ we had $v' \varolessthan \text{Y}$, then we could set

$$(v' \varolessthan \text{Y}) \odot \vee = C(v', \text{Y}, \vee) + v' \,\text{Y}\!\!\vee.$$

This leads us to compare the modified paraproduct with the "usual" paraproduct, and for that purpose we need the following auxiliary estimate. The result is from [34], Lemma 5.3.20, but the formulation here is stronger and the technical details of the proof are slightly simpler, although the idea is the same as in [34].

Lemma 4.8. *Let $\alpha<1$, $\beta\in\mathbb{R}$, and $\delta>-1$. Then we have uniformly in $\varepsilon>0$*

$$\begin{aligned}&\|\mathscr{F}^{-1}(\varphi(\varepsilon\cdot))*(u\circledcirc v)-u\circledcirc\mathscr{F}^{-1}(\varphi(\varepsilon\cdot))*v\|_{\alpha+\beta+\delta}\\&\lesssim\varepsilon^{-\delta}\Big(\|x\mathscr{F}^{-1}\varphi\|_{B^{\delta+1}_{1,\infty}}+\|\mathscr{F}^{-1}\varphi\|_{B^{\delta}_{1,\infty}}\Big)\|u\|_\alpha\|v\|_\beta.\end{aligned}$$

Proof. Let us write $\hat{\varphi}_\varepsilon=\mathcal{F}^{-1}(\varphi(\varepsilon\cdot))$. The proof is similar to that of Lemma 4.3, and we argue again on $\mathbb{R}^d$ because here we can use scaling.

1. By Lemma 2.6 it suffices to control for all j the L^∞ norm of

$$\begin{aligned}&\hat{\varphi}_\varepsilon*(\Delta_{<j-1}u\Delta_jv)(x)-\Delta_{<j-1}u\circledcirc\hat{\varphi}_\varepsilon*v(x)\\&=\int\hat{\varphi}_\varepsilon(x-y)(\Delta_{<j-1}u(y)-\Delta_{<j-1}u(x))\Delta_jv(y)\mathrm{d}y.\end{aligned}$$

2. For $\varepsilon 2^j\geqslant 1$ we use that $\Delta_{<j-1}u\Delta_jv$ and Δ_j are both spectrally supported in an annulus $2^j\mathcal{A}$, and thus we can find $\psi\in C^\infty_c$ with $\psi|_{\mathcal{A}=1}$ and replace $\hat{\varphi}_\varepsilon$ by $\hat{\varphi}_\varepsilon*\hat{\psi}_{2^{-j}}$, where

$$\hat{\psi}_{2^{-j}}=\mathscr{F}^{-1}(\psi(2^{-j}\cdot)).$$

As in Lemma 4.3 we get, using that $\alpha<1$,

$$\begin{aligned}&\left|\int\hat{\varphi}_\varepsilon*\hat{\psi}_{2^{-j}}(x-y)(\Delta_{<j-1}u(y)-\Delta_{<j-1}u(x))\Delta_jv(y)\mathrm{d}y\right|\\&\lesssim\int|\hat{\varphi}_\varepsilon*\hat{\psi}_{2^{-j}}(y)|\times|y|\mathrm{d}y2^{j(1-\alpha-\beta)}\|u\|_\alpha\|v\|_\beta,\end{aligned}$$

so it suffices to bound the integral on the right hand side. For that purpose note first that

$$\begin{aligned}y\hat{\varphi}_\varepsilon*\hat{\psi}_{2^{-j}}(y)&=y\mathscr{F}^{-1}(\varphi(\varepsilon\cdot)\psi(2^{-j}\cdot))(y)\\&=\varepsilon\mathscr{F}^{-1}((\nabla\varphi)(\varepsilon\cdot)\psi(2^{-j}\cdot))(y)+2^{-j}\mathscr{F}^{-1}(\varphi(\varepsilon\cdot)(\nabla\psi)(2^{-j}\cdot))(y)\\&=\varepsilon^{1-d}\mathscr{F}^{-1}((\nabla\varphi)\psi(\varepsilon^{-1}2^{-j}\cdot))(\varepsilon^{-1}y)\\&\quad+2^{-j}\varepsilon^{-d}\mathscr{F}^{-1}(\varphi(\nabla\psi)(\varepsilon^{-1}2^{-j}\cdot))(\varepsilon^{-1}y),\end{aligned}$$

and therefore

$$\int|y\hat{\varphi}_\varepsilon*\hat{\psi}_{2^{-j}}(y)|\mathrm{d}y=\varepsilon\|\mathscr{F}^{-1}(\nabla\varphi\psi(\varepsilon^{-1}2^{-j}\cdot))\|_{L^1}+2^{-j}\|\mathscr{F}^{-1}(\varphi(\nabla\psi)(\varepsilon^{-1}2^{-j}\cdot))\|_{L^1}.$$

The first term on the right hand side can be estimated by

$$\begin{aligned}\|\mathscr{F}^{-1}(\nabla\varphi\psi(\varepsilon^{-1}2^{-j}\cdot))\|_{L^1}&\leqslant\sum_{i:2^i\sim\varepsilon 2^j}\|\Delta_i\mathscr{F}^{-1}(\nabla\varphi)*\mathscr{F}^{-1}\psi(\varepsilon^{-1}2^{-j}\cdot)\|_{L^1}\\&\leqslant\sum_{i:2^i\sim\varepsilon 2^j}\|\Delta_i\mathscr{F}^{-1}(\nabla\varphi)\|_{L^1}\|\mathscr{F}^{-1}\psi\|_{L^1}\\&\lesssim\sum_{i:2^i\sim\varepsilon 2^j}\|\Delta_i(x\mathscr{F}^{-1}\varphi)\|_{L^1}\\&\simeq(\varepsilon 2^j)^{-\delta-1}\|x\mathscr{F}^{-1}\varphi\|_{B^{\delta+1}_{1,\infty}},\end{aligned}$$

and by the same argument

$$\|\mathscr{F}^{-1}(\varphi(\nabla\psi)(\varepsilon^{-1}2^{-j}\cdot))\|_{L^1}\lesssim(\varepsilon 2^j)^{-\delta}\|\mathscr{F}^{-1}\varphi\|_{B^{\delta}_{1,\infty}},$$

so overall

$$\begin{aligned}&\|\hat{\varphi}_\varepsilon*(\Delta_{<j-1}u\Delta_jv)-\Delta_{<j-1}u\circledcirc\varphi_\varepsilon*v\|_{L^\infty}\\&\lesssim\Big(\varepsilon(\varepsilon 2^j)^{-\delta-1}\|x\mathscr{F}^{-1}\varphi\|_{B^{\delta+1}_{1,\infty}}+2^{-j}(\varepsilon 2^j)^{-\delta}\|\mathscr{F}^{-1}\varphi\|_{B^{\delta}_{1,\infty}}\Big)2^{j(1-\alpha-\beta)}\|u\|_\alpha\|v\|_\beta\\&\lesssim\varepsilon^{-\delta}2^{-j(\delta+\alpha+\beta)}\Big(\|x\mathscr{F}^{-1}\varphi\|_{B^{\delta+1}_{1,\infty}}+\|\mathscr{F}^{-1}\varphi\|_{B^{\delta}_{1,\infty}}\Big)\|u\|_\alpha\|v\|_\beta,\end{aligned}$$

which is the desired bound if $\varepsilon 2^j \geqslant 1$.

3. For $\varepsilon 2^j < 1$ we do not make use of the function ψ. Then we obtain

$$\begin{aligned}&\left|\int \hat{\varphi}_\varepsilon(\Delta_{<j-1}u(y) - \Delta_{<j-1}u(x))\Delta_j v(y)\mathrm{d}y\right|\\ \lesssim& \int |\hat{\varphi}_\varepsilon(y)| \times |y|\mathrm{d}y 2^{j(1-\alpha-\beta)}\|u\|_\alpha \|v\|_\beta\\ =&\varepsilon \int |\mathscr{F}^{-1}(y)y|\mathrm{d}y 2^{j(1-\alpha-\beta)}\|u\|_\alpha\|v\|_\beta\\ =&\varepsilon 2^j 2^{-j(\alpha+\beta)}\|x\mathscr{F}^{-1}\varphi\|_{L^1}\|u\|_\alpha\|v\|_\beta\\ \lesssim&(\varepsilon 2^j)^{-\delta}2^{-j(\alpha+\beta)}\|x\mathscr{F}^{-1}\varphi\|_{B^{\delta+1}_{1,\infty}}\|u\|_\alpha\|v\|_\beta,\end{aligned}$$

where in the last step we used that $\delta > -1$ and thus $\varepsilon 2^j < (\varepsilon 2^j)^{-\delta}$ and $\|x\mathscr{F}^{-1}\varphi\|_{L^1} \lesssim \|x\mathscr{F}^{-1}\varphi\|_{B^{\delta+1}_{1,\infty}}$. □

We will apply this with $\varphi(x) = e^{-|2\pi x|^2}$ and $\varepsilon = t^{1/2}$, so that $\mathscr{F}^{-1}(\varphi(\varepsilon\cdot)) * u = P_t u$.

Corollary 4.9. *Let $\alpha \in (0,1)$ and $\beta \in \mathbb{R}$. We have*

$$\|u \prec\!\!\prec w - u \olessthan w\|_{C_T\mathscr{C}^{\alpha+\beta}} \lesssim (1+T)\|u\|_{\mathscr{L}^\alpha_T}(\|w(0)\|_\beta + \|\mathscr{L}w\|_{C_T\mathscr{C}^{\beta-2}}).$$

Proof. We decompose

$$\begin{aligned}&u \prec\!\!\prec w(t) - u \olessthan w(t)\\ =&\int_0^t P_{t-s}(u(s)\olessthan \mathscr{L}w(s))\mathrm{d}s + P_t(u(0)\olessthan w(0))\\ &-u(t)\olessthan\left(\int_0^t P_{t-s}\mathscr{L}w(s)\mathrm{d}s + P_t w(0)\right)\\ =&\int_0^t P_{t-s}((u(s)-u(t))\olessthan \mathscr{L}w(s))\mathrm{d}s + \int_0^t [P_{t-s}(u(t)\olessthan\mathscr{L}w(s)) - u(t)\olessthan P_{t-s}\mathscr{L}w(s)]\mathrm{d}s\\ &+P_t((u(0)-u(t))\olessthan w(0)) + [P_t(u(t)\olessthan w(0)) - u(t)\olessthan P_t w(0)]\\ =:&A_1 + A_2 + A_3 + A_4.\end{aligned}$$

Now we simply estimate each A_i separately and apply the commutation estimate between paraproduct and P_t (Lemma 4.8), the estimate $\|P_t f\|_{\gamma+\delta} \lesssim t^{-\delta/2}\|f\|_\gamma$, and the time regularity of u along the way: For example we get with $\alpha' < \alpha$

$$\begin{aligned}\|A_1\|_{\alpha'+\beta} \lesssim& \int_0^t \|P_{t-s}((u(s)-u(t))\olessthan\mathscr{L}w(s))\|_{\alpha'+\beta}\mathrm{d}s\\ \lesssim& \int_0^t (1+(t-s)^{-(2+\alpha')/2})\|(u(s)-u(t))\olessthan\mathscr{L}w(s)\|_{\beta-2}\mathrm{d}s\\ \lesssim& \int_0^t (1+(t-s)^{-1-\alpha'/2})\|u(s)-u(t)\|_{L^\infty}\|\mathscr{L}w\|_{C_T\mathscr{C}^{\beta-2}}\mathrm{d}s\\ \lesssim& \int_0^t (|t-s|^{\alpha/2} + (t-s)^{-1+(\alpha-\alpha')/2})\|u\|_{\mathscr{L}^\alpha_T}\|\mathscr{L}w\|_{C_T\mathscr{C}^{\beta-2}}\mathrm{d}s\\ \lesssim& T^{(\alpha-\alpha')/2}(1+T)\|u\|_{\mathscr{L}^\alpha_T}\|\mathscr{L}w\|_{C_T\mathscr{C}^{\beta-2}}.\end{aligned}$$

As indicated in the proof of Lemma 3.5, we can estimate the individual blocks $\Delta_j A_1$ and refine the analysis to avoid the loss of regularity and get the same result for $\alpha' = \alpha$. □

Definition 4.10. *For $v \in \mathscr{D}_T^{\beta}\mathbb{Z}$ we define*

$$(v+\Psi)\vee := (v+\Psi)\varolessthan\vee + (v+\Psi)\varogreaterthan\vee + v^{\sharp}\odot\vee + (v'\prec\!\!\!\prec Y - v'\varolessthan Y)\odot\vee$$
$$+ C(v', Y, \vee) + v' Y^{\vee} + \Psi^{\vee}.$$

Lemma 4.11. *For $v \in \mathscr{D}_T^{\beta}\mathbb{Z}$ we have*

$$\|(v+\Psi)\vee\|_{C_T\mathscr{C}^{2\alpha-2}} + \|(v+\Psi)\vee - (v+\Psi)\varolessthan\vee\|_{C_T\mathscr{C}^{3\alpha-2}} \leqslant M_T\big(1+\|v\|_{\mathscr{D}_T^{\beta}}\big),$$

where $M_T > 0$ depends polynomially on $\|\mathbb{Z}\|_{\mathcal{T}_T^{\alpha}}$. Moreover, if $\tilde{v} \in \mathscr{D}_T^{\beta}\tilde{\mathbb{Z}}$, then

$$\|(v+\Psi)\vee - (\tilde{v}+\tilde{\Psi})\tilde{\vee}\|_{C_T\mathscr{C}^{2\alpha-2}} \leqslant M_T\big(\|\mathbb{Z}-\tilde{\mathbb{Z}}\|_{\mathcal{T}_T^{\alpha}} + \|v-\tilde{v}\|_{\mathscr{D}_T^{\beta}}\big),$$

where now $M_T > 0$ depends polynomially on $\|\mathbb{Z}\|_{\mathcal{T}_T^{\alpha}}$, $\|\tilde{\mathbb{Z}}\|_{\mathcal{T}_T^{\alpha}}$, $\|v\|_{\mathscr{D}_T^{\beta}\mathbb{Z}}$, $\|\tilde{v}\|_{\mathscr{D}_T^{\beta}\tilde{\mathbb{Z}}}$, and only on $\|\mathbb{Z}\|_{\mathcal{T}_T^{\alpha}}$ if $\mathbb{Z} = \tilde{\mathbb{Z}}$.

Proof. We simply have to estimate each term in the definition of $(v+\Psi)\vee$:

$$\|(v+\Psi)\varolessthan\vee\|_{C_T\mathscr{C}^{2\alpha-2}} \lesssim \Big(\|v\|_{C_T\mathscr{C}^{\alpha}} + \|\Psi\|_{C_T\mathscr{C}^{\alpha}}\Big)\|\vee\|_{C_T\mathscr{C}^{2\alpha-2}},$$
$$\|(v+\Psi)\varogreaterthan\vee\|_{C_T\mathscr{C}^{3\alpha-2}} \lesssim \Big(\|v\|_{C_T\mathscr{C}^{\alpha}} + \|\Psi\|_{C_T\mathscr{C}^{\alpha}}\Big)\|\vee\|_{C_T\mathscr{C}^{2\alpha-2}},$$
$$\|v^{\sharp}\odot\vee\|_{C_T\mathscr{C}^{4\alpha+\beta-2}} \lesssim \|v^{\sharp}\|_{C_T\mathscr{C}^{2\alpha+\beta}}\|\vee\|_{C_T\mathscr{C}^{2\alpha-2}},$$
$$\|(v'\prec\!\!\!\prec Y - v'\varolessthan Y)\odot\vee\|_{C_T\mathscr{C}^{4\alpha+\beta-2}} \lesssim \|v'\|_{\mathscr{L}_T^{\beta}}\Big(\|Y(0)\|_{2\alpha} + \|\mathscr{L}Y\|_{C_T\mathscr{C}^{2\alpha-2}}\Big)\|\vee\|_{C_T\mathscr{C}^{2\alpha-2}}$$
$$= \|v'\|_{\mathscr{L}_T^{\beta}}\|\vee\|^2_{C_T\mathscr{C}^{2\alpha-2}},$$
$$\|C(v', Y, \vee)\|_{C_T\mathscr{C}^{4\alpha+\beta-2}} \lesssim \|v'\|_{C_T\mathscr{C}^{\beta}}\|Y\|_{C_T\mathscr{C}^{2\alpha}}\|\vee\|_{C_T\mathscr{C}^{2\alpha-2}},$$
$$\|v' Y^{\vee}\|_{C_T\mathscr{C}^{2\alpha-1}} \lesssim \|v'\|_{C_T\mathscr{C}^{\beta}}\|Y^{\vee}\|_{C_T\mathscr{C}^{2\alpha-1}},$$
$$\|\Psi^{\vee}\|_{C_T\mathscr{C}^{\alpha-1}} \lesssim \|\Psi^{\vee}\|_{C_T\mathscr{C}^{\alpha-1}},$$

where we used that $3\alpha > 1$ and $4\alpha+\beta > 2$, which both follow from the fact that $\alpha > \beta > 2/5 > 1/3$. The term with the lowest regularity is $(v+\Psi)\varolessthan\vee \in C_T\mathscr{C}^{2\alpha-2}$, and if we subtract it from the product the remaining terms all have at least regularity $3\alpha-2$. Therefore, the first claimed estimate follows.

The estimate for the difference of the two products follows from the same arguments by using the bi- or tri-linearity of all the operators involved. □

To recap, by now we defined a Banach space of paracontrolled functions $\mathscr{D}_T^{\beta}$ that can be decomposed into a paraproduct plus a more regular remainder, and we showed that for such functions the right hand side of our equation is well defined and again given as a paraproduct plus a more regular remainder. The last ingredient we need to set up a paracontrolled Picard iteration is a paracontrolled Schauder estimate, but by the definition of the modified paraproduct this is a triviality:

Proposition 4.12. *Let $\gamma \in (0,2)$, then for all $\gamma' \in (0,\gamma]$*

$$\|u - u'\prec\!\!\!\prec w\|_{\mathscr{L}_T^{\gamma'}} \lesssim \|u(0) - u'(0)\varolessthan w(0)\|_{\gamma} + T^{(\gamma-\gamma')/2}(1+T)\|\mathscr{L}u - u'\varolessthan w\|_{C_T\mathscr{C}^{\gamma-2}}.$$

Proof. We have by definition of the modified paraproduct

$$\mathscr{L}\,(u - u'\prec\!\!\!\prec w) = \mathscr{L}u - u'\varolessthan w, \qquad u(0) - u'(0)\varolessthan w(0),$$

so the claim follows from the usual Schauder estimates in parabolic spaces, see Lemma 4.14 below. ☐

By the very definition of the modified paraproduct, paracontrolled Schauder estimates are a triviality, and therefore we can now set up a Picard iteration in a space of paracontrolled distributions:

Corollary 4.13. *Let $T>0$, $\mathbb{Z}\in\mathcal{T}^\alpha$, and $v_0\in\mathscr{C}^{2\alpha+\beta}$ for $\beta\in(2/5,\alpha)$. Then the map*

$$\Pi\colon \mathscr{D}^\beta_T\to\mathscr{D}^\beta_T,\qquad \Pi(v,v',v^\sharp):=(\Pi v,(\Pi v)',(\Pi v)^\sharp):=(w,w',w^\sharp),$$

where

$$\begin{aligned} w(t) &:= P_t v_0+\int_0^t P_{t-s}((v+\Psi)\vee(s))\mathrm{d}s,\qquad t\in[0,T]\\ w' &:= v+\Psi,\\ w^\sharp &:= w-w'\prec\!\!\prec \Upsilon, \end{aligned}$$

is well defined. If $T>0$ is sufficiently small (depending only on $\|\mathbb{Z}\|_{\mathcal{T}^\alpha_T}$ but not on v_0), then Π^2 is a contraction and therefore Π has a unique fixed point.

Proof. We have to estimate $w,w',w^\sharp$. Throughout the proof M_T denotes a changing constant that only depends on $\|\mathbb{Z}\|_{\mathcal{T}^\alpha_T}$ and T and that is increasing in T.

1. We have by Lemma 4.11

$$\|\mathscr{L}w-(v+\Psi)\circledcirc\vee\|_{C_T\mathscr{C}^{3\alpha-2}}\leqslant M_T\big(1+\|v\|_{\mathscr{D}^\beta_T}\big),$$

so since $\Upsilon(0)=0$ we can apply Proposition 4.12 and obtain

$$\|w^\sharp\|_{\mathscr{L}^{2\alpha+\beta}_T}\lesssim\|v_0\|_{2\alpha+\beta}+T^{(\alpha-\beta)/2}(1+T)M_T\big(1+\|v\|_{\mathscr{D}^\beta_T}\big).$$

Note that $\|v_0\|_{2\alpha+\beta}$ does not come with a small factor, but this term drops out if we compare $\Pi v-\Pi\tilde{v}$.

2. Again by Lemma 4.11

$$\|\mathscr{L}w\|_{C_T\mathscr{C}^{2\alpha-2}}\leqslant M_T\big(1+\|v\|_{\mathscr{D}^\beta_T}\big),$$

so a direct application of Lemma 4.14 gives

$$\|w\|_{\mathscr{L}^{\alpha+\beta}_T}\lesssim\|v_0\|_{2\alpha+\beta}+T^{(\alpha-\beta)/2}(1+T)M_T\big(1+\|v\|_{\mathscr{D}^\beta_T}\big).$$

3. So w and $w^\sharp$ have the right regularity, and the factor $T^{(\alpha-\beta)/2}$ gives us a contraction for small T depending on M_T, i.e.

$$\|\Pi v-\Pi\tilde{v}\|_{\mathscr{L}^{\alpha+\beta}_T}+\|(\Pi v)^\sharp-(\Pi\tilde{v})^\sharp\|_{\mathscr{L}^{2\alpha+\beta}_T}\leqslant\frac{1}{2}\|v-\tilde{v}\|_{\mathscr{D}^\beta_T}.$$

Since the equation is linear we can choose M_T to only depend on $\mathbb{Z}$ but not on v_0 or $v,\tilde{v}$.

4. Let us get to $w'=v+\Psi$: We have

$$\|w'\|_{\mathscr{L}^\beta_T}\leqslant\|v\|_{\mathscr{L}^{\alpha+\beta}_T}+\|\Psi\|_{\mathscr{L}^\alpha_T}\lesssim(1+\|\mathbb{Z}\|_{\mathcal{T}^\alpha_T})(1+\|v\|_{\mathscr{D}^\beta_T}),$$

so w' also has the right regularity. But unlike for $w^\sharp$, we do not gain a small factor here and thus Π is not a contraction.

5. Therefore, we consider Π^2 instead. Clearly Π^2 is a contraction for v and $v^\sharp$ (possibly after reducing the value of T some more), because

$$\begin{aligned}\|(\Pi^2 v)^\sharp - (\Pi^2 \tilde{v})^\sharp\|_{\mathscr{L}_T^{2\alpha+\beta}} &\leqslant T^{(\alpha-\beta)/2} M_T \|(\Pi v, \Pi v', \Pi v^\sharp) - (\Pi\tilde{v}, \Pi\tilde{v}', \Pi\tilde{v}^\sharp)\|_{\mathscr{D}_T^\beta} \\ &\lesssim T^{(\alpha-\beta)/2} M_T \|v - \tilde{v}\|_{\mathscr{D}_T^\beta},\end{aligned}$$

and similarly

$$\|(\Pi^2 v) - (\Pi^2 \tilde{v})\|_{\mathscr{L}_T^{\alpha+\beta}} \lesssim T^{(\alpha-\beta)/2} M_T \|v - \tilde{v}\|_{\mathscr{D}_T^\beta}.$$

So let us look at the derivative:

$$\begin{aligned}\|(\Pi^2 v)' - (\Pi^2 \tilde{v})'\|_{\mathscr{L}_T^\alpha} &= \|\Pi v + \Upsilon - (\Pi\tilde{v} + \Upsilon)\|_{\mathscr{L}_T^\alpha} \\ &\leqslant \|\Pi v - \Pi \tilde{v}\|_{\mathscr{L}_T^{\alpha+\beta}} \\ &\leqslant T^{(\alpha-\beta)/2} M_T \|v - \tilde{v}\|_{\mathscr{D}_T^\beta},\end{aligned}$$

which proves the contraction property of Π^2 (for small $T > 0$).

6. By Banach's fixed point theorem Π^2 has a unique fixed point, and it only remains to show that also Π has a unique fixed point. Uniqueness is clear, because any fixed point of Π is one of Π^2 as well. So it suffices to show that the fixed point v of Π^2 is also a fixed point of Π. Let us write $w := \Pi v$. Then

$$\Pi w = \Pi^2 v = v,$$

and thus

$$\Pi^2 w = \Pi(\Pi w) = \Pi v = w.$$

Hence w is a fixed point of Π^2, and by uniqueness we have $w = v$, i.e. $\Pi v = v$.

□

We needed the following parabolic Schauder estimate, which is a refinement of Lemma 3.5:

Lemma 4.14. *Let $\gamma \in (0, 2)$, then we have for all $\gamma' \in (0, \gamma]$*

$$\begin{aligned}\|(t \mapsto P_t \varphi)\|_{\mathscr{L}_T^\gamma} &\lesssim \|\varphi\|_\gamma, \\ \left\|\left(t \mapsto \int_0^t P_{t-s} f(s) \mathrm{d}s\right)\right\|_{\mathscr{L}_T^{\gamma'}} &\lesssim T^{(\gamma-\gamma')/2}(1+T)\|f\|_{C_T \mathscr{C}^{\gamma-2}}.\end{aligned}$$

Proof. The space regularity is controlled in Lemma 3.5 (we saw in its proof that we can gain a factor $T^{(\gamma-\gamma')/2}$ by giving up some regularity). Therefore, we only have to control the time regularity.

1. For that purpose note first that for $\beta \in (0, 2)$

$$\|(P_t - \mathrm{id})\Delta_j \psi\|_{L^\infty} = \left\|\int_0^t \partial_s P_s \Delta_j \psi \mathrm{d}s\right\|_{L^\infty} \leqslant \int_0^t \|\Delta P_s \Delta_j \psi\|_{L^\infty} \mathrm{d}s \leqslant t 2^{-j(\beta-2)} \|\psi\|_\beta$$

and also

$$\|(P_t - \mathrm{id})\Delta_j \psi\|_{L^\infty} \lesssim 2^{-j\beta} \|\psi\|_\beta,$$

so with j_0 such that $2^{-j_0} \simeq t^{1/2}$ (if $t \leqslant 1$, otherwise we take $j_0 = -2$)

$$\begin{aligned}\|(P_t - \mathrm{id})\psi\|_{L^\infty} &\leqslant \sum_j \|(P_t - \mathrm{id})\Delta_j \psi\|_{L^\infty} \\ &\lesssim \left(\sum_{j \leqslant j_0} t 2^{-j(\beta-2)} + \sum_{j > j_0} 2^{-j\beta}\right) \|\psi\|_\beta \\ &\lesssim (t 2^{-j_0(\beta-2)} + 2^{-j_0\beta}) \|\psi\|_\beta \\ &\lesssim t^{\beta/2} \|\psi\|_\beta.\end{aligned}$$

2. This gives for example

$$\|P_t\varphi - P_s\varphi\|_{L^\infty} = \|(P_{t-s} - \mathrm{id})P_s\varphi\|_{L^\infty} \lesssim |t-s|^{\gamma/2} \|\varphi\|_\gamma.$$

3. For the convolution, note that it suffices to show

$$\left\| \left(t \mapsto \int_0^t P_{t-s} f(s) \mathrm{d}s \right) \right\|_{C_T^{\gamma/2} L^\infty} \lesssim (1+T) \|f\|_{C_T \mathscr{C}^\alpha},$$

the estimate for the $C_T^{\gamma'/2} L^\infty$ norm then follows from $|t-s|^{(\gamma-\gamma')/2} \leqslant T^{(\gamma-\gamma')/2}$ for $s, t \in [0, T]$. We have

$$\begin{aligned}\left\| \int_0^t P_{t-r} f(r) \mathrm{d}r - \int_0^s P_{s-r} f(r) \mathrm{d}r \right\|_{L^\infty} \leqslant & \left\| \int_s^t P_{t-r} f(r) \mathrm{d}r \right\|_{L^\infty} \\ & + \left\| (P_{t-s} - \mathrm{id}) \int_0^s P_{s-r} f(r) \mathrm{d}r \right\|_{L^\infty},\end{aligned}$$

and the second term on the right hand side is controlled with our estimate from above together with the Schauder estimates from Lemma 3.5:

$$\begin{aligned}\left\| (P_{t-s} - \mathrm{id}) \int_0^s P_{s-r} f(r) \mathrm{d}r \right\|_{L^\infty} &\lesssim |t-s|^{\gamma/2} \left\| \int_0^s P_{s-r} f(r) \mathrm{d}r \right\|_\gamma \\ &\lesssim |t-s|^{\gamma/2} (1+T) \|f\|_{C_T \mathscr{C}^{\gamma-2}}.\end{aligned}$$

For the remaining term, we decompose again into high and low frequencies and obtain with $2^{-j_0} \simeq |t-s|^{1/2}$

$$\begin{aligned}&\left\| \int_s^t P_{t-r} f(r) \mathrm{d}r \right\|_{L^\infty} \\ &\leqslant \sum_{j \leqslant j_0} \int_s^t \|P_{t-r} \Delta_j f(r)\|_{L^\infty} \mathrm{d}r + \left\| \Delta_{>j_0} \int_s^t P_{t-r} f(r) \mathrm{d}r \right\|_{L^\infty} \\ &\lesssim \sum_{j \leqslant j_0} |t-s| 2^{-j(\gamma-2)} \|f\|_{C_T \mathscr{C}^{\gamma-2}} + 2^{-j_0\gamma} \left\| \int_s^t P_{t-r} f(r) \mathrm{d}r \right\|_\gamma \\ &\lesssim (|t-s| 2^{-j_0(\gamma-2)} + 2^{-j_0\gamma}) \|f\|_{C_T \mathscr{C}^\alpha} \\ &\simeq |t-s|^{\gamma/2} \|f\|_{C_T \mathscr{C}^\alpha}.\end{aligned}$$

□

Remark 4.15. The following estimate from the proof is often useful in its own right:

$$\|(P_t - \mathrm{id})\psi\|_{L^\infty} \lesssim t^{\beta/2} \|\psi\|_\beta, \qquad \beta \in (0, 2).$$

4.3 The full Φ_3^4 equation

To study the full Φ_3^4 equation we first need to define a space of extended data where all analytically ill-defined trees live. To simplify the renormalization, we replace the operator Δ in the equation by $A = \Delta - 1$. This has the advantage that now the semigroup $(e^{tA} = e^{-t} P_t)_{t \geqslant 0}$ is integrable over all of $\mathbb{R}_+$. Of course, we can recover the original equation by considering $(\partial_t - A)\phi = -\phi^3 + \phi + \xi$ instead, but as discussed before the linear term $+\phi$ on the right hand side poses no additional difficulty for our small scale solution theory and therefore we simply omit it. On the other hand this term might have a strong effect on the long time behavior of the solution, but at least for now we do not care about that.

In the following we write for a Banach space X

$$C_{\mathbb{R}} X = C(\mathbb{R}, X) \qquad \text{and} \qquad C_{\text{pg}}(X) = \{u \in C_{\mathbb{R}} X : \exists k \in \mathbb{N}\, s.t.\ \|u(t)\|_X \lesssim 1 + |t|^k\},$$

equipped with the distance

$$d(u, v)_{C_{\mathbb{R}} X} = \sum_{n=1}^{\infty} 2^{-n} (\|u|_{[-n,n]} - v|_{[-n,n]}\|_{C([-n,n],X)} \wedge 1)$$

under which $C_{\mathbb{R}} X$ is complete (of course the subspace $C_{\text{pg}} X$ of polynomially growing functions is not closed). By adapting the proof of Lemma 3.5 is not hard to see that for $u \in C_{\text{pg}} \mathscr{C}^{\alpha}$ we have

$$t \mapsto \int_{-\infty}^{t} e^{(t-s)A} u(s) \mathrm{d}s \in C_{\text{pg}} \mathscr{C}^{\alpha+2}.$$

Definition 4.16. *Let $\alpha \in (1/3, 1/2)$ and let*

$$\mathcal{T}^{\alpha} \subset C_{\mathbb{R}} \mathscr{C}^{\alpha-1} \times C_{\mathbb{R}} \mathscr{C}^{2\alpha-2} \times C_{\mathbb{R}} \mathscr{C}^{2\alpha} \times C_{\mathbb{R}} \mathscr{C}^{\alpha} \times C_{\mathbb{R}} \mathscr{C}^{2\alpha-1} \times C_{\mathbb{R}} \mathscr{C}^{\alpha-1} \times C_{\mathbb{R}} \mathscr{C}^{2\alpha-1},$$

be the closure of the image of the map

$$\begin{gathered} \Theta : C_{\text{pg}} L^{\infty} \times \mathbb{R} \times \mathbb{R} \\ \to C_{\mathbb{R}} \mathscr{C}^{\alpha-1} \times C_{\mathbb{R}} \mathscr{C}^{2\alpha-2} \times C_{\mathbb{R}} \mathscr{C}^{2\alpha} \times C_{\mathbb{R}} \mathscr{C}^{\alpha} \times C_{\mathbb{R}} \mathscr{C}^{2\alpha-1} \times C_{\mathbb{R}} \mathscr{C}^{\alpha-1} \times C_{\mathbb{R}} \mathscr{C}^{2\alpha-1}, \\ \Theta(Z, c_1, c_2) = \left(\mathrm{I}, \vee, \mathsf{Y}, \Psi, \Psi\mathrm{I}, \Psi\vee, \mathsf{Y}\vee \right), \end{gathered}$$

where

$$\begin{aligned} \mathrm{I}(t) &= Z(t), \\ \vee(t) &= (\mathrm{I}(t)^2 - c_1), \\ \mathsf{Y}(t) &= \int_{-\infty}^{t} e^{(t-s)A} \vee(s) \mathrm{d}s, \\ \Psi(t) &= \int_{-\infty}^{t} e^{(t-s)A} (\mathrm{I}(s)^3 - 3c_1 \mathrm{I}(s)) \mathrm{d}s, \\ \Psi\mathrm{I}(t) &= \Psi(t) \odot \mathrm{I}(t), \\ \Psi\vee(t) &= \Psi(t) \odot \vee(t) - 6c_2 \mathrm{I}(t), \\ \mathsf{Y}\vee(t) &= \left(\int_{-\infty}^{t} e^{(t-s)A} \vee(s) \mathrm{d}s \right) \odot \vee(t) - 2c_2. \end{aligned}$$

We write $\mathcal{T} = \bigcup_{\alpha \in (1/3,1/2)} \mathcal{T}^{\alpha}$. We also write $\|\mathbb{Z}\|_{\mathcal{T}_T^{\alpha}}$ $\|\mathbb{Z} - \tilde{\mathbb{Z}}\|_{\mathcal{T}_T^{\alpha}}$ for the canonical norm on the product space $\mathcal{T}^{\alpha}|_{[0,T]}$ of functions in $\mathcal{T}^{\alpha}$ restricted to the time interval $[0, T]$.

Remark 4.17. We did not use the canonical regularities that we would guess from paraproduct and Schauder estimates. It would be more natural to take

$$\Psi \in C_{\mathbb{R}}\mathscr{C}^{3\alpha-1}, \qquad \Psi_{\vee} \in C_{\mathbb{R}}\mathscr{C}^{4\alpha-2}, \qquad \Psi_{\Psi} \in C_{\mathbb{R}}\mathscr{C}^{5\alpha-3}, \qquad Y_{\vee} \in C_{\mathbb{R}}\mathscr{C}^{4\alpha-2}.$$

But since we are interested in the case $\alpha = 1/2 - \varepsilon$ and to simplify the presentation we formally identify $2\alpha + k\alpha = 1 + k\alpha$ whenever $k > 0$.

Our aim is now to solve the Φ_3^4 equation for given data $\mathbb{Z} \in \mathcal{T}$. For that purpose we define a Banach space of distributions, depending on τ, in which we can use paracontrolled arguments to make sense of the equation and in which we can set up a Picard iteration:

Definition 4.18. *Let $\alpha \in (1/3, 1/2)$, $\beta \in (1/3, \alpha]$, $T > 0$, and $\mathbb{Z} \in \mathcal{T}^{\alpha}$. We say that*

$$(\phi, v', v^{\sharp}) \in C_T\mathscr{C}^{\alpha-1} \times C_T\mathscr{C}^{\beta} \times C_T\mathscr{C}^{\beta+1}$$

is paracontrolled by $\mathbb{Z}$ and write $(\phi, v', v^{\sharp}) \in \mathcal{D}_T^{\beta}\mathbb{Z}$ if

$$\phi = \uparrow - \Psi + v, \qquad v = v' \prec\!\!\prec Y + v^{\sharp}.$$

If there is no ambiguity about v' and $v^{\sharp}$, we also write $v \in \mathcal{D}_T^{\beta}\mathbb{Z}$ or $\phi \in \mathcal{D}_T^{\beta}\mathbb{Z}$.

Definition 4.19. *For $(\phi, v', v^{\sharp}) \in \mathcal{D}_T^{\beta}\mathbb{Z}$ we define the space-time distribution*

$$-\phi^{:3:} := -(\partial_t - A)\Psi - 3\Big(v - \Psi\Big)\vee + \tau_0 + \tau_1 v + \tau_2 v^2 - v^3$$

where the coefficients $\tau_0, \tau_1, \tau_2 \in C_{\mathbb{R}}\mathscr{C}^{\alpha-1}$ are defined as

$$\tau_0 := -3\uparrow \times \Psi^2 + \Psi^3 := -3\Big(\uparrow \otimes \Psi^2 + \uparrow \oslash \Psi^2 + \uparrow \odot \Big(\Psi \odot \Psi\Big) + 2C\Big(\Psi, \Psi, \uparrow\Big) + \Psi\Psi_{\vee}\Big) + \Psi^3,$$

$$\tau_1 := 6\Psi_{\vee} - 3\Psi^2 := 6\Psi_{\vee} + 6\Psi \otimes \uparrow + 6\Psi \oslash \uparrow - 3\Psi^2,$$

$$\tau_2 := -3\uparrow + 3\Psi,$$

and where

$$(v + \Psi)\vee := (v + \Psi) \otimes \vee + (v + \Psi) \oslash \vee + v^{\sharp} \odot \vee + (v' \prec\!\!\prec Y - v' \otimes Y) \odot \vee$$
$$+ C(v', Y, \vee) + v' Y_{\vee} + \Psi_{\vee}.$$

The following lemma gives a more intuitive representation of $\phi^{:3:}$ in the case of regular data $\mathbb{Z}$:

Lemma 4.20. *If $\mathbb{Z} = \Theta(Z, c_1, c_2)$ for $Z \in C_{\mathrm{pg}}L^{\infty}$, then*

$$-\phi^{:3:} = -\phi^3 + 3(c_1 + c_2)\phi.$$

Proof. If $c_1 = c_2 = 0$, then this immediately follows from the considerations at the beginning of this section, so we only have to keep track where c_1 and c_2 appear. This can be done in a lengthy but straightforward computation, noting that now all products are well defined and we can combine all the paraproducts and commutators etc. to form usual products. □

While $(\partial_t - A)\Psi$ is not necessarily a function of time with values in a space of distributions, the renormalized cube $\phi^{:3:}$ is indeed a function of time once we subtract its most singular contribution:

Lemma 4.21. *Let $\alpha \in (1/3, 1/2)$, $\beta \in (1/3, \alpha]$, $T > 0$, and $\mathbb{Z}, \tilde{\mathbb{Z}} \in \mathcal{T}^\alpha$ as well as $(\phi, v', v^\sharp) \in \mathcal{D}_T^\beta \mathbb{Z}$ and $(\tilde{\phi}, \tilde{v}', \tilde{v}^\sharp) \in \mathcal{D}_T^\beta \tilde{\mathbb{Z}}$. Then*

$$\left\| \phi^{:3:} - (\partial_t - A)\Upsilon \right\|_{C_T \mathscr{C}^{2\alpha-2}} + \left\| \phi^{:3:} - (\partial_t - A)\Upsilon - 3\Big(v - \Upsilon\Big) \odot \vee \right\|_{C_T \mathscr{C}^{\alpha-1}}$$
$$\leqslant P(\|\mathbb{Z}\|_{\mathcal{T}_T^\alpha})(1 + \|v'\|^3_{C_T \mathscr{C}^\beta} + \|v^\sharp\|^3_{C_T \mathscr{C}^{\beta+1}})$$

for a polynomial P, and also

$$\left\| \phi^{:3:} - (\partial_t - A)\Upsilon - \Big(\tilde{\phi}^{:3:} - (\partial_t - A)\tilde{\Upsilon}\Big) \right\|_{C_T \mathscr{C}^{2\alpha-2}}$$
$$+ \left\| \phi^{:3:} - (\partial_t - A)\Upsilon - 3\Big(v - \Upsilon\Big) \odot \vee - \Big(\tilde{\phi}^{:3:} - (\partial_t - A)\tilde{\Upsilon} - 3\Big(\tilde{v} - \tilde{\Upsilon}\Big) \odot \tilde{\vee}\Big) \right\|_{C_T \mathscr{C}^{\alpha-1}}$$
$$\leqslant M_T(\|\mathbb{Z} - \tilde{\mathbb{Z}}\|_{\mathcal{T}_T^\alpha} + \|v' - \tilde{v}'\|_{C_T \mathscr{C}^\beta} + \|v^\sharp - \tilde{v}^\sharp\|_{C_T \mathscr{C}^{\beta+1}}),$$

where for another polynomial P

$$M_T = P(\|\mathbb{Z}\|_{\mathcal{T}_T^\alpha}, \|\tilde{\mathbb{Z}}\|_{\mathcal{T}_T^\alpha}, \|v'\|_{C_T \mathscr{C}^{\beta+1}}, \|v^\sharp\|_{C_T \mathscr{C}^\beta}, \|\tilde{v}'\|_{C_T \mathscr{C}^\beta}, \|\tilde{v}^\sharp\|_{C_T \mathscr{C}^{\beta+1}}).$$

Proof. This easily follows by combining the Definition 4.19 of $\phi^{:3:}$ with the paraproduct estimates from Theorem 2.9 and the commutator estimate from Lemma 4.1. □

Now consider the equation

$$(\partial_t - A)\phi = -\phi^{:3:} + \xi,$$

where we wrote

$$\xi := (\partial_t - A)\mathsf{I},$$

which is a space-time distribution. We are looking for paracontrolled solutions, so we should decompose $\phi = \mathsf{I} - \Upsilon + v$, and the equation for v is

$$(\partial_t - A)v = -\phi^{:3:} + (\partial_t - A)\Upsilon, \qquad v(0) = \phi_0 - \mathsf{I}(0) + \Upsilon(0),$$

which we can control with Lemma 4.21. From here we can set up a Picard iteration and solve the equation uniquely on a small time interval, just as in the case of the linear equation of Section 4.2. Unlike in the linear case, now the length of the time interval on which we obtain a contraction depends on the initial condition, and therefore we only obtain local existence up to an explosion time. To show that the explosion time is infinite, we have to use the sign of the nonlinearity $-\phi^{:3:}$ and apply more refined estimates, see [29].

5 Parabolic Anderson model & Anderson Hamiltonian

5.1 The parabolic Anderson model

Let now ξ be a space white noise on $\mathbb{T}^2$, i.e. a centered Gaussian process with values in $\mathscr{S}'$ such that for all $\varphi, \psi \in C^\infty(\mathbb{T}^2)$ we have $\mathbb{E}[\xi(\varphi)\xi(\psi)] = \langle \varphi, \psi \rangle_{L^2(\mathbb{T}^2)}$. We want to study the parabolic Anderson model (PAM)

$$\mathscr{L}u = (\partial_t - \Delta)u = u\xi,$$

which is a continuum model for a branching population in a random potential: We consider independent diffusing particles on $\mathbb{T}^2$ that in the point x branch with rate $\xi(x)^+ = \max\{\xi(x), 0\}$ and get killed with rate $\xi(x)^- = \max\{-\xi(x), 0\}$. Of course $\xi(x)$ does not make any sense because ξ is only a distribution, and also the solution u is not integer valued; but we can derive u as a continuum limit of a discrete model behaving as described above [26].

Exercise 5.1. Show that if η is a space white noise on $\mathbb{T}^d$, then

$$\mathbb{E}[\|\eta\|_{-d/2-\kappa}^p] < \infty$$

for all $\kappa, p > 0$.

Hint: Compare with Lemma 3.3.

Thus we have $\xi \in \mathscr{C}^{-1-\kappa}$ for all $\kappa > 0$, which is the same regularity that we had for the tree $\vee$ in the linearized Φ_3^4 equation. In other words, the parabolic Anderson model is a simplified linearized Φ_3^4 equation and we can solve it by using the same arguments as in Section 4.2. Since ξ does not depend on time we can now work with "extended data" that does not depend on time either: Let $\alpha \in (2/3, 1)$ so that $\xi \in \mathscr{C}^{\alpha-2}$, and consider

$$X = (1-\Delta)^{-1}\xi \in \mathscr{C}^\alpha.$$

Then

$$\mathscr{L}X = -\Delta X = (1-\Delta)X - X = \xi - X = \xi + \mathscr{C}^\alpha.$$

Exercise 5.2. Let $\rho\colon \mathbb{R}^2 \to \mathbb{R}$ be an even, compactly supported function which is continuous in 0 and which satisfies $\rho(0) = 1$. Define

$$\xi^\varepsilon := \mathscr{F}^{-1}(\rho(\varepsilon\cdot)\mathscr{F}\xi), \qquad X^\varepsilon := (1-\Delta)^{-1}\xi^\varepsilon = \mathscr{F}^{-1}(\rho(\varepsilon\cdot)\mathscr{F}X),$$

as well as

$$c^\varepsilon := \mathbb{E}[X^\varepsilon(0)\xi^\varepsilon(0)].$$

Show that for all $x \in \mathbb{T}^2$

$$c^\varepsilon = \mathbb{E}[X^\varepsilon(x)\xi^\varepsilon(x)] = \mathbb{E}[X^\varepsilon \odot \xi^\varepsilon(x)] = O(|\log\varepsilon|)$$

and that there exists $X \diamond \xi$ such that for all $p \in [1, \infty)$

$$\lim_{\varepsilon\to 0}\mathbb{E}[\|\xi^\varepsilon - \xi\|_{\alpha-2}^p + \|X^\varepsilon - X\|_\alpha^p + \|X\diamond\xi - (X^\varepsilon \odot \xi^\varepsilon - c^\varepsilon)\|_{2\alpha-2}^p] = 0.$$

Hint: Use the tools from Section 3.4.

From here it is not difficult to slightly adapt the arguments from Section 4.2 to show that for all $\beta \in (2/3, \alpha]$ and for all u_0 with $u_0 - u_0 \varolessthan X \in \mathscr{C}^{\alpha+\beta}$ there exists a unique paracontrolled solution $u = u \prec\!\!\prec X + u^\sharp$ with $u \in \mathscr{L}_T^\beta$ and $u^\sharp \in \mathscr{L}^{\alpha+\beta}$ to

$$\mathscr{L}u = u\xi := u \varolessthan \xi + u \varogreaterthan \xi + u^\sharp \odot \xi + (u \prec\!\!\prec X - u \varolessthan X) \odot \xi + C(u, X, \xi) + u(X \diamond \xi)$$

with initial condition $u(0) = u_0$. Note that unlike before we do not have $\mathscr{L}(u \prec\!\!\prec X) = u \varolessthan \xi$, but instead

$$\mathscr{L}(u \prec\!\!\prec X) = u \varolessthan \mathscr{L}X = u \varolessthan \xi - u \varolessthan X,$$

but the term $u \varolessthan X$ has positive regularity and therefore $\int_0^t P_{t-s}(u \varolessthan X)\mathrm{d}s \in \mathscr{L}_T^{\alpha+\beta}$.

Moreover, since $X \diamond \xi = \lim_\varepsilon (X^\varepsilon \odot \xi^\varepsilon - c^\varepsilon)$, we have $u = \lim_\varepsilon u^\varepsilon$, where

$$\mathscr{L}u^\varepsilon = u^\varepsilon(\xi^\varepsilon - c^\varepsilon).$$

(It is a good exercise to convince yourself of this! Where does the c^ε enter the equation?)

Our aim is now to analyze this equation in a bit more detail. We will first extend the solution theory to much more general initial conditions, then we will present a strong maximum principle, then we will study the Anderson Hamiltonian, i.e. the infinitesimal generator of the solution semigroup, and finally we will combine all these tools to obtain a quite precise understanding of the long time behavior of the (periodic) parabolic Anderson model.

5.2 General Besov spaces and more general initial conditions

As before, the condition on the initial condition is quite unnatural. A canonical initial condition for our population would be the Dirac delta, which would model the start from a unit mass at 0, and then we could explore how this mass diffuses through the system. By Exercise 2.1 we have $\delta \in B^{-2(1-1/p)}_{p,\infty}$ for all $p \in [1, \infty)$ (since we redefined the space $\mathscr{C}^\gamma$ as the closure of the smooth functions we only have $\delta \in \mathscr{C}^{-2-\kappa}$ for $\kappa > 0$).

The estimate $\|P_t \varphi\|_{\beta+\gamma} \lesssim (1 + t^{-\gamma/2}) \|\varphi\|_\beta$ for $\gamma > 0$ from (3.4) can be easily generalized to

$$\|P_t \varphi\|_{B^{\beta+\gamma}_{p,q}} \lesssim (1 + t^{-\gamma/2}) \|\varphi\|_{B^\beta_{p,q}}.$$

To obtain an integrable singularity at $t = 0$ we need $\gamma < 2$, so since the initial condition for the paracontrolled remainder would be $u^\sharp(0) = \delta - \delta \circledcirc X \in \mathscr{C}^{-2-\kappa}$ we could at best obtain

$$u^\sharp(t) = \underbrace{P_t u^\sharp(0)}_{\mathscr{C}^{-2\kappa}} + \underbrace{\int_0^t P_{t-s} \mathscr{L} u^\sharp(s) \mathrm{d}s}_{??} \in \mathscr{C}^{-2\kappa}.$$

Of course, this is way too irregular to make sense of $u^\sharp(t) \odot \xi$. If on the other hand we could work in the space $B_{1,\infty}$, then we would have $u^\sharp(0) \in B^0_{1,\infty}$ and then $P_t u^\sharp(0) \in B^{2-\kappa}_{1,\infty}$, which gives us some hope to make sense of $u^\sharp(t) \odot \xi$, because $\alpha + 2 - \kappa > 0$.

This is the purpose of the current subsection: We develop the solution theory for paracontrolled distributions in general Besov spaces, and we finally give details on how to include singularities at $t = 0$ which allow us to take less regular initial conditions. We start by defining the following function spaces:

Definition 5.1. *Let $p \in [1, \infty]$ and $\alpha \in \mathbb{R}$, $\beta \in (0, 2)$, $\gamma \in [0, 1)$. Then we define*

$$\mathscr{C}^\alpha_p := B^\alpha_{p,\infty}, \qquad \mathscr{L}^\beta_{p,T} := C^{\beta/2}_T L^p \cap C_T \mathscr{C}^\beta_p,$$

both equipped with their canonical norm, as well as

$$\mathscr{M}^\gamma_T \mathscr{C}^\alpha_p := \{f \in C([0, T], \mathscr{S}') : t \mapsto t^\gamma \varphi(t) \in C_T \mathscr{C}^\alpha_p\},$$
$$\mathscr{L}^{\gamma,\beta}_{p,T} := \{\varphi \in C([0, T], \mathscr{S}') : t \mapsto t^\gamma \varphi(t) \in \mathscr{L}^\beta_{p,T}\},$$

with canonical norms

$$\|\varphi\|_{\mathscr{M}^\gamma_T \mathscr{C}^\alpha_p} := \|(t \mapsto t^\gamma \varphi(t))\|_{C_T \mathscr{C}^\alpha_p}, \qquad \|\varphi\|_{\mathscr{L}^{\gamma,\beta}_{p,T}} := \|(t \mapsto t^\gamma \varphi(t))\|_{\mathscr{L}^\beta_{p,T}}.$$

Now we need to translate the ingredients for paracontrolled distributions to this new functional setting. Most of this was worked out by Prömel and Trabs in [35] in more generality than we need it here, and the following lemma is a collection of weaker versions of their results:

Lemma 5.2. ([35], Lemma 2.1)

Let $p \in [1, \infty]$ and let $\beta \in \mathbb{R}$ and $u, v \in \mathscr{S}'$. Then we have for all $\alpha > 0$

$$\|u \circledcirc v\|_{\mathscr{C}^\beta_p} \lesssim \min\left\{\|u\|_{L^p} \|v\|_\beta, \|u\|_{L^\infty} \|v\|_{\mathscr{C}^\beta_p}\right\},$$

and for $\alpha < 0$ furthermore

$$\|u \circledcirc v\|_{\mathscr{C}^{\alpha+\beta}_p} \lesssim \min\left\{\|u\|_{\mathscr{C}^\alpha_p} \|v\|_\beta, \|u\|_\alpha \|v\|_{\mathscr{C}^\beta_p}\right\}.$$

If $\alpha+\beta>0$ *we have*

$$\|u\odot v\|_{\mathscr{C}_p^{\alpha+\beta}}\lesssim \min\left\{\|u\|_{\mathscr{C}_p^{\alpha}}\|v\|_{\beta},\|u\|_{\alpha}\|v\|_{\mathscr{C}_p^{\beta}}\right\}.$$

If $\alpha\in(0,1)$ *and* $\gamma\in\mathbb{R}$ *is such that* $\beta+\gamma<0$ *but* $\alpha+\beta+\gamma>0$*, then*

$$\|C(u,v,w)\|_{\mathscr{C}_p^{\alpha+\beta+\gamma}}\lesssim\|u\|_{\mathscr{C}_p^{\alpha}}\|v\|_{\beta}\|w\|_{\gamma}.$$

We also need Schauder estimates in $\mathscr{L}_{p,T}^{\gamma,\alpha}$ spaces:

Lemma 5.3. ([20], Lemma 6.6)
Let $\alpha\in(0,2)$ *and* $\gamma\in[0,1)$*, as well as* $p\in[1,\infty]$ *and* $T>0$*. Then we have*

$$\|(t\mapsto P_tu)\|_{\mathscr{L}_{p,T}^{\gamma,\alpha}}\lesssim\|u\|_{\mathscr{C}_p^{\alpha-2\gamma}},$$
$$\left\|\left(t\mapsto\int_0^t P_{t-s}u(s)\mathrm{d}s\right)\right\|_{\mathscr{L}_{p,T}^{\gamma,\alpha}}\lesssim(1+T)\|u\|_{\mathscr{M}_T^{\gamma}\mathscr{C}_p^{\alpha-2}}.$$

Adapting the proof of this lemma to deal with the modified paraproduct leads to the following generalization of Corollary 4.9

Lemma 5.4. *Let* $\alpha\in(0,1)$*,* $\beta\in\mathbb{R}$*, and* $\gamma\in[0,1)$*, as well as* $p\in[1,\infty]$ *and* $T>0$*. Then*

$$\|u\prec\!\!\!\prec w-u\circledcirc w\|_{\mathscr{M}_T^{\gamma}\mathscr{C}_p^{\alpha+\beta}}\lesssim(1+T)\|u\|_{\mathscr{L}_{p,T}^{\gamma,\alpha}}(\|w(0)\|_{\beta}+\|\mathscr{L}w\|_{C_T\mathscr{C}^{\beta-2}}).$$

And finally we need an interpolation estimate:

Lemma 5.5. ([20], Lemma 6.8)
Let $\alpha\in(0,2)$*,* $\gamma\in(0,1)$*, and* $\kappa\in[0,\alpha\wedge2\gamma)$*, as well as* $p\in[1,\infty]$ *and* $T>0$*. Then*

$$\|u\|_{\mathscr{L}_{p,T}^{\gamma-\kappa/2,\alpha-\kappa}}\lesssim\|u\|_{\mathscr{L}_{p,T}^{\gamma,\alpha}}.$$

We now fix $p\in[1,\infty]$ consider an initial condition $u_0\in\mathscr{C}_p^0$. As before we let $\alpha\in(2/3,1)$ and $\beta\in(2/3,\alpha]$ We define paracontrolled distributions in our new setting as follows:

Definition 5.6. *We say* $(u,u',u^\sharp)$ *is paracontrolled,* $u\in\mathscr{D}_{p,T}^{\beta}$*, if*

$$u=u'\prec\!\!\!\prec X+u^\sharp$$

with

$$u\in\mathscr{L}_{p,T}^{\beta/2,\beta},\qquad u'\in\mathscr{L}_{p,T}^{\beta/2,\beta},\qquad u^\sharp\in\mathscr{L}_{p,T}^{(\alpha+\beta)/2,\alpha+\beta}.$$

Now consider $u\in\mathscr{D}_{p,T}^{\alpha}$, and note that by the interpolation estimate we have $\mathscr{D}_{p,T}^{\beta}\subset\mathscr{D}_{p,T}^{\alpha}$. Thus we get from the estimates by Prömel and Trabs, using that $2\alpha+\beta>2$,

$$u\xi=\underbrace{u\circledcirc\xi}_{\mathscr{M}_T^{\alpha/2}\mathscr{C}_p^{2\alpha-2}}+\underbrace{u\circledcirc\xi}_{\mathscr{M}_T^{\beta/2}\mathscr{C}_p^{\alpha-2}}+\underbrace{u^\sharp\odot\xi}_{\mathscr{M}_T^{(\alpha+\beta)/2}\mathscr{C}_p^{2\alpha+\beta-2}}+\underbrace{(u'\prec\!\!\!\prec X-u'\circledcirc X)\odot\xi}_{\mathscr{M}_T^{\beta/2}\mathscr{C}_p^{2\alpha+\beta-2}}$$
$$+\underbrace{C(u',X,\xi)}_{\mathscr{M}_T^{\beta/2}\mathscr{C}_p^{2\alpha+\beta-2}}+\underbrace{u'(X\diamond\xi)}_{\mathscr{M}_T^{\beta/2}\mathscr{C}_p^{2\alpha-2}}.$$

Therefore,

$$u\xi-u\circledcirc\xi\in\mathscr{M}_T^{(\alpha+\beta)/2}\mathscr{C}_p^{2\alpha-2},$$

and then

$$\|\mathscr{L}u^\sharp\|_{\mathscr{M}_T^\alpha \mathscr{C}_p^{2\alpha-2}} \leqslant \|u \circleddash X\|_{\mathscr{M}_T^\alpha \mathscr{C}_p^{2\alpha-2}} + \|u\xi - u \circleddash \xi\|_{\mathscr{M}_T^\alpha \mathscr{C}_p^{2\alpha-2}} \lesssim T^{(\alpha-\beta)/2},$$

which gives us the small factor for the contraction. We also have

$$t \mapsto P_t u^\sharp(0) = P_t(u_0 - u_0 \prec\!\!\prec X) \in \mathscr{M}_T^\alpha \mathscr{C}_p^{2\alpha},$$

and thus we can control $u^\sharp \in \mathscr{L}_{p,T}^{\alpha,2\alpha}$. Moreover,

$$\|u\|_{\mathscr{L}_{p,T}^{\alpha/2,\alpha}} \lesssim \|u^\sharp\|_{\mathscr{L}_{p,T}^{\alpha,2\alpha}} + T^{(\alpha-\beta)/2} \|\mathscr{L}(u \prec\!\!\prec X)\|_{\mathscr{M}_T^{\beta/2} \mathscr{C}_p^{\alpha-2}} + \|u_0 \circleddash X\|_{\mathscr{C}_p^0},$$

and

$$\mathscr{L}(u \prec\!\!\prec X) = \underbrace{u \circleddash \xi}_{\mathscr{M}_T^{\beta/2} \mathscr{C}_p^{\alpha-2}} - \underbrace{u \circleddash X}_{\mathscr{M}_T^{\beta/2} \mathscr{C}_p^{\alpha}}.$$

We can now apply the same strategy as in Section 4.2 to obtain the following result:

Theorem 5.7. *Let $p \in [1,\infty]$, let $(\xi, X, X \diamond \xi) \in \mathscr{C}^{\alpha-2} \times \mathscr{C}^\alpha \times \mathscr{C}^{2\alpha-2}$, and let $u_0 \in \mathscr{C}_p^0$. Then for all $T > 0$ there exists a unique $u \in \mathscr{D}_T^\alpha$ such that*

$$\mathscr{L}u = u\xi, \qquad u(0) = u_0.$$

Moreover, u depends continuously on $(u_0, \xi, X, X \diamond \xi)$, and for all $t > 0$ we have $u(t) \in \mathscr{C}^\alpha$.

Proof. Everything is clear by now, only the last point $u(t) \in \mathscr{C}^\alpha$ merits some discussion. By construction, have $u(s) \in \mathscr{C}_p^\alpha$ for all $s > 0$. So by Besov embedding (Lemma 2.5) we have for $p_1 > p$ with $\alpha - 2\left(\frac{1}{p} - \frac{1}{p_1}\right) = 0$:

$$u(t/2) \in \mathscr{C}_{p_1}^{\alpha - 2\left(\frac{1}{p} - \frac{1}{p_1}\right)} = \mathscr{C}_{p_1}^0.$$

From here we can bootstrap to increase the integrability to the L^∞ scale: We now use $u(t/2)$ as new initial condition for the equation on $[t/2, 2t/3]$, so at $u(2t/3)$ we get an even better integrability $\mathscr{C}_{p_2}^0$ with $p_2 > p_1$ such that $\alpha - 2\left(\frac{1}{p_1} - \frac{1}{p_2}\right) = 0$. This bootstrapping ends after finally many steps, when we arrive at the L^∞ scale. □

5.3 A strong maximum principle

We saw in the previous section that $u(t) \in \mathscr{C}^\alpha$ is a continuous function for all $t > 0$, even if the initial condition is only a Dirac delta $\delta \in \mathscr{C}_1^0$. In particular, it makes sense to speak of the sign of $u(t,x)$. By approximation it is not hard to see that whenever u_0 is a positive measure in $\mathscr{C}_p^0$ for some $p \in [1,\infty]$, then $u(t,x) \geqslant 0$ for all $t > 0$ and $x \in \mathbb{T}^2$. Here we present a nice argument due to Cannizzaro, Friz, Gassiat [9], inspired by Mueller [31], which shows that in fact the solution becomes instantly strictly positive, as long as u_0 can be approximated by positive functions.

We start with a simple observation about the heat kernel:

Lemma 5.8. ([9], (5.3))

Let p_t the Gaussian density on $\mathbb{R}^d$ with variance $2t$. For all $\rho > 0$ there exists $t_\rho > 0$ such that whenever $u \geqslant 0$ satisfies $u \geqslant 1$ on the ball $B(x,\kappa)$, then for all $t \in [0, t_\rho]$

$$p_t * u(y) \geqslant \frac{1}{4}, \qquad y \in B(x, \kappa + t\rho).$$

Proof. Without loss of generality we take $x=0$. For any $y\in B(x,\kappa)$ there exists a unique $|z|\leqslant 1$ with $y=z\times(\kappa+t\rho)$. So if Z is a standard Gaussian variable, then

$$\begin{aligned}p_t*u(y)\geqslant P_t\mathbb{I}_{B(0,\kappa)}(y)=&\mathbb{P}(|y+\sqrt{2t}Z|\leqslant\kappa)\\ =&\mathbb{P}\bigg(Z\in B\bigg(\frac{y}{\sqrt{2t}},\frac{\kappa}{\sqrt{2t}}\bigg)\bigg)\\ =&\mathbb{P}\bigg(Z\in B\bigg(z\bigg(\frac{\kappa}{\sqrt{2t}}+\sqrt{\frac{t}{2}}\rho\bigg),\frac{\kappa}{\sqrt{2t}}\bigg)\bigg),\end{aligned}$$

and for $t\to 0$ we have

$$\begin{aligned}&\lim_{t\to 0}\inf_{|z|\leqslant 1}\mathbb{P}\bigg(Z\in B\bigg(z\bigg(\frac{\kappa}{\sqrt{2t}}+\sqrt{\frac{t}{2}}\rho\bigg),\frac{\kappa}{\sqrt{2t}}\bigg)\bigg)\\ =&\lim_{t\to 0}\inf_{|z|\leqslant 1}\mathbb{P}\bigg(Z\in B\bigg(|z|e_1\bigg(\frac{\kappa}{\sqrt{2t}}+\sqrt{\frac{t}{2}}\rho\bigg),\frac{\kappa}{\sqrt{2t}}\bigg)\bigg)\\ =&\lim_{t\to 0}\mathbb{P}\bigg(Z\in B\bigg(e_1\bigg(\frac{\kappa}{\sqrt{2t}}+\sqrt{\frac{t}{2}}\rho\bigg),\frac{\kappa}{\sqrt{2t}}\bigg)\bigg)=\mathbb{P}(Z_1>0)=\frac{1}{2}.\end{aligned}$$

This proves the claim. □

Theorem 5.9. (Strong maximum principle, [9], Theorem 5.1)
Let $u_0\in\mathscr{C}_p^0$ be a positive measure. Assume furthermore that $u_0\neq 0$. Then

$$u(t,x)>0,\qquad t>0,x\in\mathbb{T}^2.$$

Proof. By the previous discussion we know that $u(s)\in\mathscr{C}^\alpha$ for all strictly positive times $s>0$. Moreover, from a convolution argument the positive measure u_0 can be approximated by a sequence of positive functions u_0^n, and therefore $u(s)\geqslant 0$ by approximation, and for sufficiently small $s>0$ we have $u(s)\neq 0$ by continuity. By considering $u(s)$ as a new initial condition, we may assume without loss of generality that u_0 is a positive continuous function with $u_0\neq 0$. Moreover, it turns out to be easier to interpret u as a periodic function on $\mathbb{R}^2$.

Then there exists a ball $B(x,\kappa)\subset\mathbb{R}^2$ on which $u_0\geqslant\varepsilon$ for some $\varepsilon>0$. Since the equation for u is linear, $\varepsilon^{-1}u$ solves the same equation but with initial condition $\varepsilon^{-1}u_0$. Since $\varepsilon^{-1}u>0$ if and only if $u>0$, we may forget about the multiplication with ε^{-1}, and thus we assume without loss of generality that

$$u_0\geqslant 1\qquad\text{on}\qquad B(x,\kappa)\subset\mathbb{R}^2.$$

We can decompose

$$u(s)=P_su_0+\int_0^s P_{s-r}(u(r)\xi)\mathrm{d}r=:P_su_0+w(s).$$

By adapting the proof of Theorem 5.7 to initial conditions in $\mathscr{C}^\alpha$ (for which we get a smaller blow-up factor γ) we see that $w\in C_T^{\alpha/2}L^\infty$. Let now $t>0$. Since $w(0)=0$, there exists $C>0$ such that for all $s\in[0,t]$

$$\|w(s)\|_{L^\infty}\leqslant Cs^{\alpha/2}.$$

Moreover,

$$P_su_0=\sum_{k\in\mathbb{Z}^2}p_s(\cdot+k)*u_0\geqslant p_s*\mathbb{I}_{B(x,\kappa)},$$

so that by Lemma 5.8 we get for all $\rho > 0$ and $s < t_\rho$

$$P_s u_0 \geqslant \frac{1}{4}$$

on $B(x, \kappa + s\rho)$. So if $s \in (0, t_\rho)$ is small enough such that

$$C s^{\alpha/2} < \frac{1}{8},$$

we get

$$u(s, y) \geqslant \frac{1}{8}, \qquad y \in B(x, \kappa + s\rho).$$

Using the linearity of the equation, we can repeat the argument on $[s, 2s]$ and obtain $u(2s, y) \geqslant 1/64$ for $y \in B(x, \kappa + 2s\rho)$, and so on, until we arrive at

$$u(t, y) > 0, \qquad y \in B(x, \kappa + t\rho).$$

Since $\rho > 0$ was arbitrary, this completes the proof. □

This argument is very flexible, essentially we only used that the equation is linear. In particular, it extends to all linear equations that can be solved with paracontrolled distributions or regularity structures.

5.4 The Anderson Hamiltonian

The parabolic Anderson model can be written as

$$\partial_t u = \Delta u + u\xi = (\Delta + \xi) u =: \mathscr{H} u,$$

where $\mathscr{H}$ is the Anderson Hamiltonian,

$$\mathscr{H} = \Delta + \xi$$

(or, taking renormalization into account, $\mathscr{H} = \Delta + \xi - \infty$). So we formally have $u(t) = e^{t\mathscr{H}} u_0$, and we hope to obtain information about the behavior of u from $\mathscr{H}$.

To do this, we first have to construct $\mathscr{H}$. In principle, we can define $\mathscr{H} u$ for all *paracontrolled* (with slightly different definition than before) u of the form

$$u = u' \otimes X + u^\sharp,$$

where $X = (1 - \Delta)^{-1} \xi \in \mathscr{C}^\alpha$, $u, u' \in \mathscr{C}_p^\alpha$, $u^\sharp \in \mathscr{C}_p^{2\alpha}$ for $\alpha \in (2/3, 1)$, because then we have

$$\mathscr{H} u = \Delta(u' \otimes X) + \Delta u^\sharp + u \otimes \xi + u \otimes \xi + u^\sharp \odot \xi + C(u', X, \xi) + u'(X \diamond \xi).$$

The problem is that while the right hand side is well defined, it is still only in $\mathscr{C}_p^{2\alpha - 2}$ and thus only a distribution and not a function. Also, we would like to construct $\mathscr{H}$ as a self-adjoint operator on a Hilbert space, so that we can use spectral theory.

The first problem could be overcome very easily by restricting our attention to a subspace of the paracontrolled distributions: If we write T_t for the map that sends u_0 to $u(t)$, where u is the solution of the PAM, then for all $u_0 \in \mathscr{C}_p^0$ both $T_s u_0$ and the integral $\int_0^t T_s u_0 \mathrm{d}s$ are paracontrolled, and

$$\mathscr{H} \int_0^t T_s u_0 \mathrm{d}s = \int_0^t \mathscr{H} T_s u_0 \mathrm{d}s = \int_0^t \partial_s T_s u_0 \mathrm{d}s = T_t u_0 - u_0,$$

so if u_0 is “nice enough” (depending on the space on which we want to define $\mathscr{H}$), then $\int_0^t T_s u_0 \mathrm{d}s$ is in the domain of $\mathscr{H}$. Since also $t^{-1} \int_0^t T_s u_0 \mathrm{d}s$ converges to u_0, we would obtain that the domain is dense. The problem with this approach is that it seems not so easy to obtain information about the spectrum of $\mathscr{H}$ from that construction.

Therefore, we take a different approach which goes back to Allez and Chouk [2]: For $\lambda > 0$ we consider the *resolvent equation*

$$(\lambda - \mathscr{H})u = v$$

for $v \in L^2 \subset \mathscr{C}_2^0$. We can rewrite this equation as a paracontrolled PDE as follows:

$$(\lambda - \Delta)u = u\xi + v \qquad \Leftrightarrow \qquad u = (\lambda - \Delta)^{-1}(u\xi + v)$$

As there is no time variable, we do not need the modified paraproduct and for large λ the equation is actually easier to solve than the parabolic Anderson model. The space of paracontrolled distributions for this problem is with $\beta \in (2/3, \alpha)$

$$(u, u', u^\sharp) \in \mathscr{C}_2^\alpha \times \mathscr{C}_2^\beta \times \mathscr{C}_2^{\alpha+\beta}: \qquad u^\sharp = u - u' \varolessthan X.$$

We obtain a small factor for the contraction property by choosing λ large (which we fix from now on). The solution u is then in $\mathscr{C}_2^\beta \subset B_{2,2}^{\beta'}$ for all $\beta' \in (0, \beta)$.

Exercise 5.3. For $\gamma \in \mathbb{R}$ we define the L^2 Sobolev space

$$H^\gamma = \left\{ u \in \mathscr{S}'(\mathbb{T}^d): \|u\|_{H^\gamma}^2 := \sum_{k \in \mathbb{Z}^d} |\hat{u}(k)|^2 (1 + |k|^2)^\gamma < \infty \right\}.$$

i. Show that $H^\gamma = B_{2,2}^\gamma$ with equivalent norms.

ii. Show that $H^0 = L^2$.

iii. Show that bounded sets in H^γ are relatively compact in $H^{\gamma'}$ whenever $\gamma > \gamma'$.

By the exercise we see that the operator

$$R_\lambda: L^2 \ni v \mapsto (\lambda - \mathscr{H})^{-1} v = u \in L^2$$

is compact (i.e. it maps bounded sets to relatively compact sets). Moreover, it is self-adjoint: Assume for the moment that ξ is a bounded function, then $\lambda - \mathscr{H} = \lambda - \Delta - \xi$ is self-adjoint because multiplication operators are trivially self-adjoint and Δ is self-adjoint as well, and therefore

$$\langle R_\lambda v, w \rangle_{L^2} = \langle R_\lambda v, (\lambda - \mathscr{H}) R_\lambda w \rangle_{L^2} = \langle (\lambda - \mathscr{H}) R_\lambda v, R_\lambda w \rangle_{L^2} = \langle v, R_\lambda w \rangle_{L^2}.$$

By approximation, this carries over to our situation.

So R_Λ is a compact self-adjoint operator on the Hilbert space $L^2(\mathbb{T}^2)$, and by the spectral theorem for compact operators there exists an orthonormal basis of eigenfunctions $(e_n)_{n \in \mathbb{N}}$ and real valued eigenvalues $(\kappa_n)_{n \in \mathbb{N}}$ such that $|\kappa_1| \geqslant |\kappa_2| \geqslant \cdots$ with $|\kappa_n| \to 0$ and

$$R_\lambda e_n = \kappa_n e_n, \qquad n \in \mathbb{N}.$$

We just need one more information about the κ_n:

Lemma 5.10. *We have $\kappa_n > 0$ for all $n \in \mathbb{N}$.*

Proof. By definition of $R_\lambda = (\lambda - \mathscr{H})^{-1}$ we have $R_\lambda v \neq 0$ for all $v \neq 0$, and therefore $|\kappa_n| > 0$ for all n. Thus it suffices to show that $\kappa_n \geqslant 0$, which follows immediately once we show that R_λ is a positive operator, i.e. that

$$\langle R_\lambda v, v \rangle_{L^2} \geqslant 0$$

for all $v \in L^2$. Indeed, then

$$\kappa_n = \kappa_n \langle e_n, e_n \rangle_{L^2} = \langle R_\lambda e_n, e_n \rangle_{L^2} \geqslant 0.$$

To see that R_λ is positive, we use the representation

$$R_\lambda v = \int_0^\infty e^{-\lambda t} T_t v \mathrm{d}t.$$

Indeed, we have for $S > 0$

$$(\lambda - \mathscr{H}) \int_0^S e^{-\lambda t} T_t v \mathrm{d}t = \int_0^S -\partial_t (e^{-\lambda t} T_t v) \mathrm{d}t = v - e^{-\lambda S} T_S v,$$

and since the PAM is linear we have $\|T_S v\|_{L^2} \leqslant K e^{KS} \|v\|_{L^2}$ for some $K > 0$. Without loss of generality we assume that $\lambda > K$ (in fact we already had to take $\lambda > K$ in the construction of R_λ), and then we can send $S \to \infty$ to get

$$(\lambda - \mathscr{H}) \int_0^\infty e^{-\lambda t} T_t v \mathrm{d}t = v,$$

which proves the representation for R_λ. So now it suffices to show that T_t is a positive operator for all t. But

$$\langle T_t v, v\rangle_{L^2} = \langle T_{t/2} T_{t/2} v, v\rangle_{L^2} = \langle T_{t/2} v, T_{t/2} v\rangle_{L^2} = \|T_{t/2} v\|_{L^2}^2 \geqslant 0,$$

where we used the semigroup property of (T_t), and that T_t is self-adjoint for all t: If ξ is a bounded function, this follows for example from the Feynman-Kac formula, because we may interpret v as a periodic function on $\mathbb{R}^2$ and then obtain with a two-dimensional Brownian motion

$$\begin{aligned}
\int_{\mathbb{T}^2} T_t v(x) w(x) \mathrm{d}x &= \int_{\mathbb{T}^2} \mathbb{E}\left[v(x + B_t) \exp\left(\int_0^t \xi(x + B_s) \mathrm{d}s\right)\right] w(x) \mathrm{d}x \\
&= \mathbb{E}\left[\int_{\mathbb{T}^2} v(x + B_t) \exp\left(\int_0^t \xi(x + B_s) \mathrm{d}s\right) w(x) \mathrm{d}x\right] \\
&= \mathbb{E}\left[\int_{\mathbb{T}^2} v(x) \exp\left(\int_0^t \xi(x + B_s - B_t) \mathrm{d}s\right) w(x - B_t) \mathrm{d}x\right] \\
&= \mathbb{E}\left[\int_{\mathbb{T}^2} v(x) \exp\left(\int_0^t \xi(x - (B_t - B_{t-s})) \mathrm{d}s\right) w(x - (B_t - B_{t-t})) \mathrm{d}x\right] \\
&= \mathbb{E}\left[\int_{\mathbb{T}^2} v(x) \exp\left(\int_0^t \xi(x + B_s) \mathrm{d}s\right) w(x + B_t) \mathrm{d}x\right] \\
&= \int_{\mathbb{T}^2} v(x) T_t w(x) \mathrm{d}x,
\end{aligned}$$

where we used that $s \mapsto -(B_t - B_{t-s})$, $s \in [0, t]$, is a Brownian motion. By approximation, T_t is also self-adjoint in the white noise case. □

Consequently, the eigenvalues κ_n are not only decreasing in absolute value, they are actually decreasing. We claim that the e_n are also eigenfunctions for $\mathscr{H}$. Indeed,

$$(\lambda - \mathscr{H}) e_n = \kappa_n^{-1} (\lambda - \mathscr{H}) \kappa_n e_n = \kappa_n^{-1} (\lambda - \mathscr{H}) R_\lambda e_n = \kappa_n^{-1} e_n,$$

and thus

$$\mathscr{H} e_n = (\lambda - \kappa_n^{-1}) e_n =: \lambda_n e_n$$

Since the κ_n are decreasing, also the λ_n are decreasing, and since $\kappa_n \to 0$ we have $\lambda_n \to -\infty$ for $n \to \infty$. We thus obtained a spectral decomposition

$$\mathscr{H} v = \sum_{n=1}^\infty \lambda_n \langle e_n, v\rangle_{L^2} e_n,$$

which is valid whenever v is in the domain of $\mathscr{H}$.

5.5 Long time behavior of the periodic parabolic Anderson model

We formally have $T_t = e^{t\mathscr{H}}$, and with spectral calculus we can make this rigorous:

$$T_t u = \sum_{n=1}^{\infty} e^{t\lambda_n} \langle e_n, u \rangle_{L^2} e_n.$$

For u such that $\langle u, e_n \rangle_{L^2} \neq 0$ only for finitely many n this representation follows immediately from the fact that both $T_t u$ and $\sum_{n=1}^{\infty} e^{t\lambda_n} \langle e_n, u \rangle_{L^2} e_n$ solve the equation $\partial_t v = \mathscr{H} v$, and then it extends to general u by approximation.

Lemma 5.11. *The operator $\mathscr{H}$ has a spectral gap, i.e. $\lambda_1 > \lambda_2$. Moreover, $e_1(x) > 0$ for all $x \in \mathbb{T}^2$.*

Proof. Consider the cone

$$K = \{v \in L^2 : v \geqslant 0\} \subset L^2.$$

We say $v \geqslant w$ (resp. $v \gg w$) if $v - w$ is in K (resp. in the interior K° of K), or in other words if $v - w \geqslant 0$ (resp. $v - w > 0$) almost everywhere. Then T_t is a compact linear operator that is strongly positive, i.e. such that

$$T_t v \gg 0$$

whenever $v \geqslant 0$; indeed, this follows from the strong maximum principle Theorem 5.9.

By (a consequence of) the Krein-Rutman theorem, see Theorem 19.3 in [14], we have $\lambda_1 > \lambda_2$ and $e_1 \gg 0$. Since $e_1 = \kappa_1^{-1} R_\lambda e_1$ we know that e_1 is paracontrolled, i.e. there exist $e_1 = e_1' \otimes X + e_1^\sharp$ with $e_1' \in \mathscr{C}_2^\beta$ and $e_1^\sharp \in \mathscr{C}_2^{\alpha+\beta}$. The paraproduct estimates together with the Besov embedding theorem thus show that $e_1 \in \mathscr{C}^\varepsilon$ for some $\varepsilon > 0$ and thus e_1 is a continuous function and $e_1(x) > 0$ for all x. □

We now collected all ingredients needed to describe the long time behavior of the PAM:

Theorem 5.12. *There exists $\kappa > 0$ such that for all $u \in L^2$ with $u \geqslant 0$ and $u \neq 0$*

$$\lim_{t \to \infty} \left\| \frac{T_t u}{e^{t\lambda_1} \langle u, e_1 \rangle_{L^2} e_1} - 1 \right\|_{L^\infty} = 0$$

Consequently, we have for $u' \geqslant 0$ with $u' \neq 0$

$$\lim_{t \to \infty} \left\| \frac{T_t u}{T_t u'} - \frac{\langle u, e_1 \rangle_{L^2}}{\langle u', e_1 \rangle_{L^2}} \right\|_{L^\infty} = 0,$$

i.e. the ratio of two solutions for different initial conditions becomes constant for large times.

Proof. First note that $\langle u, e_1 \rangle_{L^2} > 0$ since $e_1(x) > 0$ for all x, and therefore the division by $e^{t\lambda_1} \langle u, e_1 \rangle_{L^2} e_1$ is allowed. We also have $e_1(x) \geqslant \varepsilon > 0$ for all x, because e_1 is continuous and $\mathbb{T}^2$ is compact and thus e_1 attains its minimum. Thus

$$\frac{T_t u}{e^{t\lambda_1} \langle u, e_1 \rangle_{L^2} e_1} = 1 + \frac{1}{\langle u, e_1 \rangle_{L^2} e_1} \sum_{n=2}^{\infty} e^{t(\lambda_n - \lambda_1)} \langle u, e_n \rangle_{L^2} e_n.$$

From here it is trivial to show that $\frac{T_t u}{e^{t\lambda_1}\langle u, e_1\rangle_{L^2} e_1} - 1$ converges to zero in L^2, but we have to work a bit more the get the convergence in L^∞. We have for $\tau > 0$ and $t > \tau$

$$\begin{aligned}
&\left| \frac{T_t u(x)}{e^{t\lambda_1}\langle u, e_1\rangle_{L^2} e_1(x)} - 1 \right| \\
&\leqslant \left\| \frac{1}{\langle u, e_1\rangle_{L^2} e_1} \right\|_{L^\infty} \sum_{n=2}^{\infty} e^{t(\lambda_n - \lambda_1)} |\langle u, e_n\rangle_{L^2}| |e_n(x)| \\
&\leqslant |\langle u, e_1\rangle_{L^2} \varepsilon|^{-1} \left(\sum_{n=2}^{\infty} e^{2\tau(\lambda_n - \lambda_1)} |e_n(x)|^2 \right)^{1/2} \left(\sum_{n=2}^{\infty} e^{2(t-\tau)(\lambda_n - \lambda_1)} |\langle u, e_n\rangle_{L^2}|^2 \right)^{1/2} \\
&\lesssim \left(\sum_{n=1}^{\infty} e^{2\tau\lambda_n} |e_n(x)|^2 \right)^{1/2} e^{2(t-\tau)(\lambda_2 - \lambda_1)} \sum_{n=2}^{\infty} |\langle u, e_n\rangle_{L^2}|^2 \\
&\lesssim \left(\sum_{n=1}^{\infty} e^{2\tau\lambda_n} |e_n(x)|^2 \right)^{1/2} e^{2(t-\tau)(\lambda_2 - \lambda_1)} \|u\|_{L^2}^2 .
\end{aligned}$$

Now we use that $\lambda_1 > \lambda_2$ by Lemma 5.11, so the claim follows once we show that $\sum_{n=1}^{\infty} e^{2\tau\lambda_n} |e_n(x)|^2 < \infty$. But if ρ^ε is a smooth approximation of the Dirac delta, then

$$\begin{aligned}
\sum_{n=1}^{\infty} e^{2\tau\lambda_n} |e_n(x)|^2 &\leqslant \liminf_{\varepsilon \to 0} \sum_{n=1}^{\infty} e^{2\tau\lambda_n} \langle e_n, \rho^\varepsilon(x - \cdot)\rangle_{L^2} \langle e_n, \rho^\varepsilon(x - \cdot)\rangle_{L^2} \\
&= \int u^{\rho^\varepsilon(x-\cdot)}(2\tau, y) \rho^\varepsilon(y) \mathrm{d}y,
\end{aligned}$$

where $u^{\rho^\varepsilon(x-\cdot)}$ is the solution to PAM with initial condition ρ^ε. By our previous results $u^{\rho^\varepsilon(x-\cdot)}(2\tau, \cdot)$ converges uniformly to $u^{\delta(x-\cdot)}$ and thus we get

$$\sum_{n=1}^{\infty} e^{2\tau\lambda_n} |e_n(x)|^2 \leqslant u^{\delta(x-\cdot)}(2\tau, x) < \infty.$$

To derive the limiting behavior of $T_t u / T_t u'$, note that

$$\frac{T_t u}{T_t u'} = \frac{T_t u}{e^{t\lambda_1}\langle u, e_1\rangle_{L^2} e_1} \times \frac{e^{t\lambda_1}\langle u', e_1\rangle_{L^2} e_1}{T_t u'} \times \frac{\langle u, e_1\rangle_{L^2}}{\langle u', e_1\rangle_{L^2}},$$

and now note that if $\|f_n - 1\|_{L^\infty} \to 0$, then also $\|1/f_n - 1\|_{L^\infty} \to 0$, from where we deduce the second claimed convergence. □

Consequently, the initial condition does not influence the limiting shape of the solution to the PAM at all and only contributes through the scalar factor $\langle u, e_1\rangle$.

We can also solve the PAM on $\mathbb{R}^2$, and there the situation is much more complicated. Then the operator $\mathscr{H}$ does not have discrete spectrum, its spectrum is unbounded from above, and it does not generate a continuous contraction semigroup. While we can still solve the PAM on $\mathbb{R}^2$, the solution at time t lives in a larger space than at time 0, with more permissive weights capturing the growth/decay at infinity; see [24] or [26] for details.

5.6 Some related linear equations

The spectral point of view provides us with easy solution theories for some other linear equations, for example the stochastic Schroedinger equation

$$i\partial_t u = \Delta u + u\xi = \mathscr{H}\xi$$

with $i = \sqrt{-1}$, or the stochastic wave equation

$$\partial_{tt} u = (\mathscr{H} - \lambda)\xi$$

for $\lambda \geqslant \lambda_1$. In the first case we can set

$$u(t) = \sum_{n=1}^{\infty} e^{it\lambda_n} \langle u, e_n \rangle_{L^2} e_n,$$

and in the second case we consider the equation as a system,

$$\begin{aligned} \partial_t u &= v \\ \partial_t v &= (\mathscr{H} - \lambda) u, \end{aligned}$$

so that we can set

$$K_t(u_0, v_0),$$

where

$$K_t = \begin{pmatrix} \cos(t(\lambda - \mathscr{H})^{1/2}) & (\lambda - \mathscr{H})^{-1/2} \sin(t(\lambda - \mathscr{H})^{1/2}) \\ -(\lambda - \mathscr{H})^{1/2} \sin(t(\lambda - \mathscr{H})^{1/2}) & \cos(t(\lambda - \mathscr{H})^{1/2}) \end{pmatrix}.$$

Based on this point of view we can then introduce nonlinear perturbations to the equations, for example by considering the mild formulation based on $(e^{it\mathscr{H}})_t$ resp. $(K_t)_t$, see [21] for details.

6 Relation with regularity structures

Hairer's regularity structures [22] provide another approach towards dealing with singular SPDEs, and they are based on closely related ideas, although they use very different technical tools. They are based on generalizations of the Taylor expansion and of increment characterizations of regularity. Here we discuss some links between paracontrolled distributions and regularity structures, essentially how the different descriptions of regularity are compatible.

My aim is not to give an introduction to regularity structures, this section addresses mainly readers with some previous exposure and in the beginning we only collect the main notions from [22] without any motivation or intuitive explanation. For nice introductions to regularity structures see for example [23, 12].

Definition 6.1. *A regularity structure is a triple $\mathcal{T} = (A, T, G)$, where $A \subset \mathbb{R}$ without accumulation point except possibly at ∞, where*

$$T = \bigoplus_{\alpha \in A} T_\alpha$$

and each T_α is a Banach space, and where G is a group of bounded linear operators on T such that for all $\Gamma \in G$ and all $\tau \in T_\alpha$

$$\Gamma \tau - \tau \in \bigoplus_{\beta < \alpha} T_\beta, \qquad \Gamma \in G, \tau \in T_\alpha.$$

We call T the model space and G the structure group.

A regularity structure is a purely abstract construct that provides a framework in which we can set up new notions of regularity and new function spaces. These function spaces depend on concrete realizations of regularity structures, that are encoded in *models*.

For $\varphi \colon \mathbb{R}^d \to \mathbb{R}$ we define

$$\varphi_x^\lambda := \lambda^{-d} \varphi(\lambda^{-1}(\cdot - x)), \qquad \lambda > 0, x \in \mathbb{R}^d.$$

For $\alpha \in A$ and $\tau \in T$ we write $\mathcal{Q}_\alpha \tau$ for the projection of τ onto T_α, and

$$\|\tau\|_\alpha := \|\mathcal{Q}_\alpha \tau\|$$

Definition 6.2. *Let $\mathcal{T} = (A, T, G)$ be a regularity structure and $d \in \mathbb{N}$. A model for $\mathcal{T}$ on $\mathbb{R}^d$ consists of maps*

$$\Pi: \mathbb{R}^d \to \mathcal{L}(T, \mathscr{S}'(\mathbb{R}^d)), \qquad \Gamma: \mathbb{R}^d \times \mathbb{R}^d \to G,$$

such the algebraic relations

$$\Pi_x \Gamma_{xy} = \Pi_y, \qquad \Gamma_{xy} \Gamma_{yz} = \Gamma_{xz}$$

hold, and such that with $r > -\alpha$ for all $\alpha \in A$ the following analytic relation holds: There exists $C > 0$ such that

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

for all $\varphi \in C_b^r$ with $\|\varphi\|_{C_b^r} \leqslant 1$, for all $\lambda \in (0, 1]$, for all $x, y \in \mathbb{R}^d$ and for all $\alpha, \beta \in A$.

Models provide the framework in which to model regularity, and given a model we can define new function spaces:

Definition 6.3. *Let $\mathcal{T} = (A, T, G)$ be a regularity structure with model (Π, Γ). For $\gamma > 0$ the space of modelled distributions $\mathscr{D}^\gamma$ consists of the maps $f: \mathbb{R}^d \to \bigoplus_{\alpha < \gamma} T_\alpha \subset T$ such that*

$$\|f_x - \Gamma_{xy} f_y\|_\alpha \leqslant C |x - y|^{\gamma - \alpha}, \qquad \|f_x\|_\alpha \leqslant C$$

for all $\alpha < \gamma$ and $x, y \in \mathbb{R}^d$, where $C > 0$. We write $\|f\|_{\mathscr{D}^\gamma}$ for the smallest such constant.

Modelled distributions take values in the abstract Banach space T. But we can associate to each modelled distribution an element of $\mathscr{S}'(\mathbb{R}^d)$:

Theorem 6.4. ([22], Theorem 3.10) *Let $\gamma > 0$. Then there exists a bounded linear operator*

$$\mathscr{R}: \mathscr{D}^\gamma \to \mathscr{C}^{\alpha_0},$$

where $\alpha_0 = \min_{\alpha \in A}$, such that for some $C > 0$

$$|\mathscr{R} f(\varphi_x^\lambda) - \Pi_x f_x(\varphi_x^\lambda)| \leqslant C \lambda^\gamma \tag{6.1}$$

for all $\varphi \in C_b^r$ with $\|\varphi\|_{C_b^r} \leqslant 1$, for all $\lambda \in (0, 1]$, and for all $x \in \mathbb{R}^d$. Here r is as in Definition 6.2. Moreover, $\mathscr{R} f$ is the unique element of $\mathscr{S}'$ that satisfies (6.1).

We now want to link the theory of regularity structures with paracontrolled distributions. More precisely, we give descriptions of modelled distributions based on paraproducts. The material here is from Section 6 of [16] and from [27].

Definition 6.5. *For $f: \mathbb{R}^d \to T$ we define*

$$P(f, \Pi)(x) := \sum_{j \geqslant 1} \iint K_{<j-1}(x - y) K_j(x - z) \Pi_y f_y(z) \mathrm{d}y \mathrm{d}z$$

and

$$P(f, \Gamma)(x) := \sum_{j \geqslant 1} \iint K_{<j-1}(x - y) K_j(x - z) \Gamma_{zy} f_y \mathrm{d}y \mathrm{d}z$$

whenever these are well defined. Note that $P(f, \Pi)$ takes values in $\mathscr{S}'(\mathbb{R}^d)$, while $P(f, \Gamma)$ takes values in T.

We can extend our Besov spaces $\mathscr{C}^\kappa$ easily to distributions with values in a Banach space X, by writing

$$\|u\|_{\mathscr{C}^\kappa(X)} = \sup_{j \geqslant -1} 2^{j\kappa} \|K_j * u\|_{L^\infty(X)} =: \sup_{j \geqslant -1} 2^{j\kappa} \|\Delta_j u\|_{L^\infty(X)}$$

Lemma 6.6. *Let $\gamma > 0$ and $f \in \mathscr{D}^\gamma$. Then for all $\alpha < \gamma$*

$$\|\mathcal{Q}_\alpha(f - P(f, \Gamma))\|_{\mathscr{C}^{\gamma-\alpha}(T_\alpha)} \lesssim \|f\|_{\mathscr{D}^\gamma},$$

and also

$$\|\mathscr{R} f - P(f, \Pi)\|_\gamma \lesssim \|f\|_\gamma.$$

Proof. First observe the trivial estimate

$$\|\mathcal{Q}_\alpha \Delta_{\leqslant 0} f\|_{\mathscr{C}^\gamma(T_\alpha)} \lesssim \|f\|_{L^\infty(T_\alpha)} \leqslant \|f\|_{\mathscr{D}^\gamma}.$$

Moreover, since $\int K_{<j-1}(x) \mathrm{d}x = 1$ for all $j \geqslant 1$:

$$\begin{aligned} & \mathcal{Q}_\alpha(f - \Delta_{\leqslant 0} f - P(f, \Gamma))(x) \\ = & \sum_{j \geqslant 1} \iint K_{<j-1}(x-y) K_j(x-z) \mathcal{Q}_\alpha(f_z - \Gamma_{zy} f_y) \mathrm{d}y \mathrm{d}z. \end{aligned}$$

By Lemma 2.7 it suffices to bound each addend of the series, and using that $f \in \mathscr{D}^\gamma$ we have

$$\begin{aligned} & \left| \iint K_{<j-1}(x-y) K_j(x-z) \mathcal{Q}_\alpha(f_z - \Gamma_{zy} f_y) \mathrm{d}y \mathrm{d}z \right| \\ \leqslant & \iint |K_{<j-1}(x-y) K_j(x-z)| \times |z-y|^{\gamma-\alpha} \|f\|_{\mathscr{D}^\gamma} \mathrm{d}y \mathrm{d}z \\ \lesssim & \|f\|_{\mathscr{D}^\gamma} \iint |K_{<j-1}(x-y) K_j(x-z)| \times (|z-x|^{\gamma-\alpha} + |x-y|^{\gamma-\alpha}) \mathrm{d}y \mathrm{d}z \\ \lesssim & \|f\|_{\mathscr{D}^\gamma} 2^{-j(\gamma-\alpha)}, \end{aligned}$$

which shows that $\|\mathcal{Q}_\alpha(f - P(f, \Gamma))\|_{\mathscr{C}^{\gamma-\alpha}(T_\alpha)} \lesssim \|f\|_{\mathscr{D}^\gamma}$.

The second bound holds by very similar arguments:

$$\|\Delta_{\leqslant 0} \mathscr{R} f\|_\gamma \lesssim \|\mathscr{R} f\|_{\alpha_0} \lesssim \|f\|_{\mathscr{D}^\gamma}$$

by definition of $\Delta_{\leqslant 0}$ and the reconstruction operator, and

$$(\Delta_{>0} \mathscr{R} - P(f, \Pi))(x) = \sum_{j \geqslant 1} \iint K_{<j-1}(x-y) K_j(x-z) (\mathscr{R} f(z) - \Pi_y f_y(z)) \mathrm{d}y \mathrm{d}z.$$

Again it suffices to bound each addend individually, and

$$\begin{aligned} & \left| \iint K_{<j-1}(x-y) K_j(x-z) (\mathscr{R} f(z) - \Pi_y f_y(z)) \mathrm{d}y \mathrm{d}z \right| \\ = & \left| \int K_{<j-1}(x-y) \big(\mathscr{R} f \big(K_x^{2^{-j}} \big) - \Pi_y f_y \big(K_x^{2^{-j}} \big) \big) \mathrm{d}y \right| \\ \leqslant & \left| \int K_{<j-1}(x-y) \big(\mathscr{R} f \big(K_x^{2^{-j}} \big) - \Pi_x f_x \big(K_x^{2^{-j}} \big) \big) \mathrm{d}y \right| \\ + & \left| \int K_{<j-1}(x-y) \Pi_x (f_x - \Gamma_{xy} f_y) \big(K_x^{2^{-j}} \big) \mathrm{d}y \right| \\ \lesssim & \left| \int K_{<j-1}(x-y) 2^{-j\gamma} \|f\|_{\mathscr{D}^\gamma} \mathrm{d}y \right| + \sum_{\alpha < \gamma} \left| \int K_{<j-1}(x-y) 2^{-j\alpha} |x-y|^{\gamma-\alpha} \|f\|_{\mathscr{D}^\gamma} \mathrm{d}y \right| \\ \lesssim & 2^{-j\gamma} \|f\|_{\mathscr{D}^\gamma}, \end{aligned}$$

which completes the proof. □

Note that $\mathscr{R}f$ itself is much more irregular than $\mathscr{C}^\gamma$, in general we only get $\mathscr{R}f \in \mathscr{C}^{\alpha_0}$. So the lemma gives a decomposition of $\mathscr{R}f$ into a paraproduct and a smooth remainder.

To do: paracontrolled distributions are modelled.

7 A nonlinear stochastic wave equation

7.1 Dimension 2

Here we follow Gubinelli, Koch and Oh [18] and study the equation

$$\partial^2_{tt} u = \Delta u + u^2 + \xi$$

on $\mathbb{R}_+ \times \mathbb{T}^2$, where ξ is a space-time white noise. While at first sight this looks very similar to the Φ^3_2 equation, it behaves were differently because the wave equation has much worse regularizing properties than the heat heat equation.

We can bring the equation to first order in time by rewriting it as a system, $u = (u, v)$, with

$$\begin{aligned} \partial_t u &= v \\ \partial_t v &= \Delta v + u^2 + \xi, \end{aligned}$$

or

$$\partial_t \begin{pmatrix} u \\ v \end{pmatrix} = A \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ u^2 + \xi \end{pmatrix},$$

where

$$A = \begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix}.$$

Then we have with $|\nabla| := (-\Delta)^{1/2}$

$$e^{tA} = \begin{pmatrix} \cos(t|\nabla|) & |\nabla|^{-1}\sin(t|\nabla|) \\ -|\nabla|\sin(t|\nabla|) & \cos(t|\nabla|) \end{pmatrix}, \tag{7.1}$$

where $\cos(t|\nabla|)$ and $\sin(t|\nabla|)$ are defined in terms of spectral calculus, or explicitly through the Fourier transform:

$$\cos(t|\nabla|)u = \sum_{k \in \mathbb{Z}^2} \cos(t|2\pi k|)\hat{u}(k) e^{2\pi i k \cdot x}.$$

We can verify this representation for e^{tA} by differentiating the matrix in (7.1) and by showing that the derivative equals Ae^{tA}.

In particular, the variation of constants formula gives

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = e^{tA} \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \int_0^t e^{(t-s)A} \begin{pmatrix} 0 \\ u(s)^2 + \xi(s) \end{pmatrix} \mathrm{d}s,$$

and since we are mainly interested in u:

$$u(t) = \mathcal{S}(t)(u_0, v_0) + \int_0^t \frac{\sin((t-s)|\nabla|)}{|\nabla|} (u(s)^2 + \xi(s)) \mathrm{d}s,$$

where

$$\mathcal{S}(t)(u_0, v_0) := \cos(t|\nabla|)u_0 + \frac{\sin(t|\nabla|)}{|\nabla|} v_0.$$

Here we see an important difference compared to the heat equation: While $P_t = e^{t\Delta} = e^{-t|\nabla|^2}$ is infinitely regularizing (with a blow-up for $t \to 0$), the same is not true for $\sin(t|\nabla|)/|\nabla|$, which only seems to gain one derivative.

Let us proceed anyways, and apply the tools that we developed for dealing with the Φ_2^4 equation: We make the Ansatz

$$u = Z + w,$$

where

$$Z(t) = \int_0^t \frac{\sin((t-s)|\nabla|)}{|\nabla|} \xi(s) \mathrm{d}s.$$

To simplify notation, we also write

$$\mathcal{I} f(t) := \int_0^t \frac{\sin((t-s)|\nabla|)}{|\nabla|} f(s) \mathrm{d}s,$$

so that for example $Z = \mathcal{I}\xi$. Then

$$w = \mathcal{S}(u_0, v_0) + \mathcal{I}(w^2 + 2wZ + Z^2).$$

Of course, there will be problems with defining Z^2 because of the irregularity of Z. We can derive the regularity of Z along the lines of Lemma 3.3, and this gives

$$Z \in C_T \mathscr{C}^{-\kappa}$$

almost surely for all $\kappa > 0$. This may be surprising, because it is the same regularity that we got for the convolution of ξ with the better behaved heat kernel, but at least an estimate in the Sobolev scale is actually very easy to obtain: We have for a family of complex valued standard Brownian motions $(B^k)_{k \in \mathbb{Z}}$ (i.e. both real and imaginary part of B^k are independent standard Brownian motions) such that $\mathbb{E}[B_t^k B_s^\ell] = 2\delta_{k,-\ell} s \wedge t$ the representation

$$Z(t,x) = \sum_{k \in \mathbb{Z}^2} e^{2\pi i k \cdot x} \int_0^t \frac{\sin((t-s)|2\pi k|)}{|2\pi k|} \mathrm{d}B_s^k,$$

and thus

$$\mathbb{E}[\|Z(t)\|_{H^\alpha}^2] = \sum_{k \in \mathbb{Z}^2} (1+|k|^2)^\alpha 2 \int_0^t \frac{\sin((t-s)|2\pi k|)^2}{|2\pi k|^2} \mathrm{d}s \lesssim \sum_{k \in \mathbb{Z}^2} (1+|k|^2)^{\alpha-1},$$

which is finite as soon as $\alpha < 0$. Since also

$$\int_0^t e^{-2(t-s)|2\pi k|^2} \mathrm{d}s = O(|k|^{-2}),$$

this explains to some extent why we see the same regularity as for the heat equation. See Proposition 2.1 of [18] for the precise derivation of the regularity of Z, where it is also shown that for all n

$$Z^{:n:} := \lim_{\varepsilon \to 0} Z_\varepsilon^{:n:} := \lim_{\varepsilon \to 0} H_n(Z_\varepsilon, \mathrm{var}(Z_\varepsilon)),$$

for a suitable mollification Z_ε of Z, satisfies $Z^{:n:} \in C_T \mathscr{C}^{-\kappa}$ for all n and all $\kappa > 0$. Hence, we modify the equation for w to take the renormalization into account and try to solve

$$w = \mathcal{S}(u_0, v_0) + \mathcal{I}(w^2 + 2wZ + Z^{:2:}).$$

For that purpose we have to understand the regularizing properties of $\mathcal{I}$ better, which are provided by the Strichartz estimates for $\mathcal{I}$. To formulate them, we need a definition:

Definition 7.1. ([18], Lemma 3.1)

Let $s\in(0,1)$. A pair $(q,r)\in(2,\infty]\times[2,\infty)$ is s-admissible if

$$\frac{1}{q}+\frac{2}{r}=1-s \qquad \text{and} \qquad 2\leqslant r\leqslant\begin{cases}\frac{6}{3-4s}, & s<\frac{3}{4}\\ \infty, & \text{else}\end{cases}.$$

A pair $(\tilde{q},\tilde{r})\in[1,2)\times(1,2]$ is dual s-admissible if the conjugate exponents $(\tilde{q}',\tilde{r}')$ are $(1-s)$-admissible, or equivalently if

$$\frac{1}{\tilde{q}}+\frac{2}{\tilde{r}}=3-s, \qquad \text{and} \qquad \max\left\{1+,\frac{6}{7-4s}\right\}\leqslant\tilde{r}\leqslant\frac{2}{2-s},$$

where $\tilde{r}\geqslant 1+$ means $\tilde{r}>1$.

We also need suitable function spaces to work in. It turns out that the regularizing properties of $\mathcal{I}$ are best seen in L^2-Sobolev spaces, therefore we set

$$\mathcal{H}^s:=H^s\times H^{s-1}, \qquad s\in\mathbb{R}.$$

We also write

$$L_T^pX=L^p([0,T],X)$$

for a Banach space X

Lemma 7.2. ([18], Lemma 3.2)
Let $s\in(0,1)$ and let (q,r) be s-admissible and $(\tilde{q},\tilde{r})$ be dual s-admissible. Then we have for

$$u=\mathcal{S}(u_0,v_0)+\mathcal{I}f$$

and $T\in(0,1]$ the following Strichartz estimates:

$$\|(u,\partial_t u)\|_{L_T^\infty\mathcal{H}^s}+\|u\|_{L_T^qL^r}\lesssim\|(u_0,v_0)\|_{\mathcal{H}^s}+\min\left\{\|f\|_{L_T^{\tilde{q}}L^{\tilde{r}}},\|f\|_{L_T^1H^{s-1}}\right\}.$$

In other words, we gain one derivative on the Sobolev scale H^α for u, but we can also gain integrability instead.

Now let us try to set up a Picard iteration for w in $L_T^\infty\mathcal{H}^s$ with $s\in(0,1)$. Then we have to control the right hand side of the equation, which is done in the following lemma:

Lemma 7.3. *We have for $s\in(0,1)$ and $\kappa\in(0,s\wedge(1-s))$*

$$\|w^2+2wZ+Z^{:2:}\|_{L_T^\infty H^{s-1}}\lesssim\|w\|_{L_T^\infty H^s}^2+\|w\|_{L_T^\infty H^s}\|Z\|_{L_T^\infty\mathscr{C}^{-\kappa}}+\|Z^{:2:}\|_{L_T^\infty\mathscr{C}^{-\kappa}}.$$

Proof. We decompose

$$w^2=w\odot w+2w\otimes w$$

and estimate for $\kappa\in(0,s)$ via Besov embedding (and using that $d=2$)

$$\|w\odot w\|_{L_T^\infty H^{s-1}}\lesssim\|w\odot w\|_{L_T^\infty\mathscr{C}_2^{s-1+\kappa}}\lesssim\|w\odot w\|_{L_T^\infty\mathscr{C}_1^{s+\kappa}}\lesssim\|w\|_{L_T^\infty\mathscr{C}_2^{(s+\kappa)/2}}^2\lesssim\|w\|_{L_T^\infty H^s}^2,$$

as well as

$$\begin{aligned}\|w\otimes w\|_{L_T^\infty H^{s-1}}&\lesssim\|w\otimes w\|_{L_T^\infty\mathscr{C}_2^{s-1+\kappa}}\\&\lesssim\|w\|_{L_T^\infty\mathscr{C}_\infty^{-1+\kappa}}\|w\|_{L_T^\infty\mathscr{C}_2^s}\\&\lesssim\|w\|_{L_T^\infty\mathscr{C}_2^\kappa}\|w\|_{L_T^\infty\mathscr{C}_2^s}\\&\lesssim\|w\|_{L_T^\infty H^s}^2.\end{aligned}$$

Next, we let $\kappa \in (0, (1-s) \wedge s)$ and get

$$\|wZ\|_{L_T^\infty H^{s-1}} \lesssim \|wZ\|_{L_T^\infty \mathscr{C}_2^{-\kappa}} \lesssim \|w\|_{L_T^\infty \mathscr{C}_2^s} \|Z\|_{L_T^\infty \mathscr{C}^{-\kappa}} \lesssim \|w\|_{L_T^\infty H^s} \|Z\|_{L_T^\infty \mathscr{C}^{-\kappa}}.$$

And finally

$$\|Z^{:2:}\|_{L_T^\infty H^{s-1}} \lesssim \|Z^{:2:}\|_{L_T^\infty \mathscr{C}^{-\kappa}}.$$

□

Combining this with the Strichartz estimate, we can set up a Picard iteration for $(u, \partial_t u)$ in $C_T \mathcal{H}^s$ and get existence and uniqueness of solutions for small $T > 0$.

It may seem surprising that we stated the Strichartz estimates in such a complicated way, when the regularization effect in $\mathcal{H}^s$ was all we needed. But that is very much due to the fact that here we only considered the nonlinearity u^2. If we replace u^2 by u^k for $k > 2$, then things become more complicated because we lose more integrability, and then we need to start keeping track also of the integrability. Moreover, we pick up additional constraints for the regularity s of the initial condition.

7.2 Dimension 3

Now let us try to see what could be done for the same wave equation

$$\partial_{tt}^2 u = \Delta u + u^2 + \xi$$

in $d = 3$, i.e. $u: \mathbb{R}_+ \times \mathbb{T}^3 \to \mathbb{R}$. Here we follow Gubinelli, Koch and Oh [17]. As before, we start with the ansatz

$$u = \Psi + w,$$

where we use tree notation now and

$$\Psi = \mathcal{I}\xi.$$

The same simple computation we did above to derive the regularity of Z also works in this setting, and it suggests (correctly) that $\Psi \in C_T \mathscr{C}^{-1/2-}$. The next term in the expansion is

$$\Upsilon = \mathcal{I}(\Psi^{:2:}),$$

and if we apply the usual heuristic for guessing the regularity, we get $\Psi^{:2:} \in C_T \mathscr{C}^{-1-}$ and then, since $\mathcal{I}$ should gain one derivative because of the factor $|\nabla|^{-1}$, we guess

$$\Upsilon = \mathcal{I}(\Psi^{:2:}) \in C_T \mathscr{C}^{0-}.$$

But unlike in the parabolic setting, this guess is in fact suboptimal, and we can show that

$$\Upsilon \in C_T \mathscr{C}^{1/2-},$$

i.e. it is “half a derivative” better than expected. But to see this we have to estimate Υ directly and cannot first construct $\Psi^{:2:}$ and then apply Strichartz estimates. See [17], Proposition 1.6.

We thus make the ansatz

$$u = \Psi + \Upsilon + w$$

and obtain the following equation for w:

$$w = \mathcal{S}(u_0, v_0) + \mathcal{I}((w + \Upsilon)^2 + 2(w + \Upsilon)\Psi).$$

Since $\dot{\vert} \in C_T \mathscr{C}^{-1/2-}$, we expect w at best to have regularity $1/2-$, so that $w \dot{\vert}$ is not well defined. To proceed, we use the paracontrolled ansatz, although in a slightly different formulation that goes back to Mourrat and Weber [29]. Namely, we make the Ansatz

$$w = w_1 + w_2$$

with

$$\begin{aligned} w_1 &= \mathcal{I}(2(w_1 + w_2 + \mathsf{Y}) \olessthan \dot{\vert}) \\ w_2 &= \mathcal{S}(u_0, v_0) + \mathcal{I}((w_1 + w_2 + \mathsf{Y})^2 + 2(w + \mathsf{Y}) \odot \dot{\vert} + 2(w + \mathsf{Y}) \olessthan \dot{\vert}). \end{aligned}$$

By the usual power counting we would guess the regularities $w_1 \in C_T B_?^{1/2-}$ and $w_2 \in C_T B_?^{1-}$, where $B_?$ denotes a suitable Besov space that will be determined later. We also see that we would expect having to control

$$\mathsf{Y} \odot \dot{\vert} \in C_T \mathscr{C}^{0-},$$

and indeed this is possible. Now it looks like we are in good shape to put our usual paracontrolled machinery in place, provided that (u_0, v_0) are regular enough and we control $(\mathcal{I}\dot{\vert}) \odot \dot{\vert}$. However, there is one major problem: When solving parabolic paracontrolled equations, we essentially used that the heat semigroup commutes with the paraproduct up to a smoother remainder (this was hidden in Corollary 4.9), and that is no longer true for the operator $\mathcal{I}$. We can however control the relevant resonant product with stochastic computations, i.e. show that

$$f \mapsto \mathcal{I}(\mathcal{I}(f \olessthan \dot{\vert}) \odot \dot{\vert})$$

is a bounded random operator between suitable Sobolev spaces. From here the analysis is similar to what we have seen before.

Bibliography

[1] Sergio Albeverio and Michael Röckner. Stochastic differential equations in infinite dimensions: solutions via Dirichlet forms. *Probab. Theory Related Fields*, 89(3):347–386, 1991.

[2] Romain Allez and Khalil Chouk. The continuous Anderson hamiltonian in dimension two. *ArXiv preprint arXiv:1511.02718*, 2015.

[3] Hajer Bahouri, Jean-Yves Chemin, and Raphael Danchin. *Fourier analysis and nonlinear partial differential equations*. Springer, 2011.

[4] Ismaël Bailleul and Frédéric Bernicot. Higher order paracontrolled calculus. *ArXiv preprint arXiv:1609.06966*, 2016.

[5] Ismaël Bailleul, Frédéric Bernicot, and Dorothee Frey. Space-time paraproducts for paracontrolled calculus, 3d-PAM and multiplicative Burgers equations. *ArXiv preprint arXiv:1506.08773*, 2015.

[6] Jean-Michel Bony. Calcul symbolique et propagation des singularites pour les équations aux dérivées partielles non linéaires. *Ann. Sci. Éc. Norm. Supér. (4)*, 14:209–246, 1981.

[7] Yvain Bruned, Ajay Chandra, Ilya Chevyrev, and Martin Hairer. Renormalising SPDEs in regularity structures. *ArXiv preprint arXiv:1711.10239*, 2017.

[8] Yvain Bruned, Martin Hairer, and Lorenzo Zambotti. Algebraic renormalisation of regularity structures. *ArXiv preprint arXiv:1610.08468*, 2016.

[9] Giuseppe Cannizzaro, Peter K Friz, and Paul Gassiat. Malliavin calculus for regularity structures: the case of gPAM. *J. Funct. Anal.*, 272(1):363–419, 2017.

[10] Rémi Catellier and Khalil Chouk. Paracontrolled distributions and the 3-dimensional stochastic quantization equation. *Ann. Probab.*, 46(5):2621–2679, 2018.

[11] Ajay Chandra and Martin Hairer. An analytic BPHZ theorem for regularity structures. *ArXiv preprint arXiv:1612.08138*, 2016.

[12] Ajay Chandra and Hendrik Weber. Stochastic PDEs, regularity structures, and interacting particle systems. *Ann. Fac. Sci. Toulouse Math. (6)*, 26(4):847–909, 2017.

[13] Giuseppe Da Prato and Arnaud Debussche. Strong solutions to the stochastic quantization equations. *Ann. Probab.*, 31(4):1900–1916, 2003.

[14] Klaus Deimling. *Nonlinear functional analysis*. Springer-Verlag, Berlin, 1985.

[15] James Glimm and Arthur Jaffe. *Quantum physics. A functional integral point of view*. Springer-Verlag, New York, Second edition, 1987.

[16] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics, Pi*, 3(e6), 2015.

[17] Massimiliano Gubinelli, Herbert Koch, and Tadahiro Oh. Paracontrolled approach to the three-dimensional stochastic nonlinear wave equation with quadratic nonlinearity. *ArXiv preprint arXiv:1811.07808*, 2018.

[18] Massimiliano Gubinelli, Herbert Koch, and Tadahiro Oh. Renormalization of the two-dimensional stochastic nonlinear wave equations. *Trans. Amer. Math. Soc.*, 370(10):7335–7359, 2018.

[19] Massimiliano Gubinelli and Nicolas Perkowski. Lectures on singular stochastic PDEs. *Ensaios Mat.*, 29, 2015.

[20] Massimiliano Gubinelli and Nicolas Perkowski. KPZ reloaded. *Comm. Math. Phys.*, 349(1):165–269, 2017.

[21] Massimiliano Gubinelli, Baris Evren Ugurcan, and Immanuel Zachhuber. Semilinear evolution equations for the anderson hamiltonian in two and three dimensions. *ArXiv preprint arXiv:1807.06825*, 2018.

[22] Martin Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[23] Martin Hairer. Introduction to regularity structures. *Braz. J. Probab. Stat.*, 29(2):175–210, 2015.

[24] Martin Hairer and Cyril Labbé. A simple construction of the continuum parabolic Anderson model on $\boldsymbol{R}^2$. *Electron. Commun. Probab.*, 20:0, 2015.

[25] Svante Janson. *Gaussian Hilbert spaces*, volume 129 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1997.

[26] Jörg Martin and Nicolas Perkowski. Paracontrolled distributions on Bravais lattices and weak universality of the 2d parabolic Anderson model. *ArXiv preprint arXiv:1704.08653*, 2017.

[27] Jörg Martin and Nicolas Perkowski. A Littlewood-Paley description of modelled distributions. *ArXiv preprint arXiv:1808.00500*, 2018.

[28] Jean-Christophe Mourrat and Hendrik Weber. Convergence of the two-dimensional dynamic Ising-Kac model to Φ^4_2. *Comm. Pure Appl. Math.*, 70(4):717–812, 2017.

[29] Jean-Christophe Mourrat and Hendrik Weber. The dynamic Φ^4_3 model comes down from infinity. *Comm. Math. Phys.*, 356(3):673–753, 2017.

[30] Jean-Christophe Mourrat, Hendrik Weber, and Weijun Xu. Construction of Φ^4_3 diagrams for pedestrians. In *From particle systems to partial differential equations*, volume 209 of *Springer Proc. Math. Stat.*, pages 1–46. Springer, Cham, 2017.

[31] Carl Mueller. On the support of solutions to the heat equation with noise. *Stochastics Stochastics Rep.*, 37(4):225–245, 1991.

[32] David Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, Second edition, 2006.

[33] G. Parisi and Yong Shi Wu. Perturbation theory without gauge fixing. *Sci. Sinica*, 24(4):483–496, 1981.

[34] Nicolas Simon Perkowski. *Studies of robustness in stochastic analysis and mathematical finance*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, 2014.

[35] David J. Prömel and Mathias Trabs. Rough differential equations driven by signals in Besov spaces. *J. Differential Equations*, 260(6):5202–5249, 2016.