Lectures on singular stochastic PDEs

M. Gubinelli
CEREMADE & CNRS UMR 7534
Université Paris Dauphine and IUF, France
Email: gubinelli@ceremade.dauphine.fr

N. Perkowski
CEREMADE & CNRS UMR 7534
Université Paris Dauphine
Email: perkowski@ceremade.dauphine.fr

Version 0.981 – DRAFT – August 2014

Abstract

These are the notes for a course at the 18th Brazilian School of Probability held from 3th to 9th August 2014 in Mambucaba. The aim of the course is to introduce the basic problems of non–linear PDEs with stochastic and irregular terms. We explain how it is possible to handle them using two main techniques: the notion of energy solutions [GJ10, GJ13] and that of paracontrolled distributions, recently introduced in [GIP13]. In order to maintain a link with physical intuitions we motivate such singular SPDEs via an homogeneisation problem for a random potential.

Table of contents

1 Introduction

The aim of these lectures is to explain how to apply controlled path ideas [Gub04] to solve simple problems in singular PDEs. The hope is that the insight gained by doing so can inspire new applications or the construction of other more powerful tools to analyze a wider class of problems.

We discuss some problems involving singular stochastic non–linear parabolic equations from the point of view of controlled paths. To understand the origin of such singular equations we have chosen to present the example of an homogeneisation problem of a singular potential in a linear parabolic equation. This point of view have the added benefit to be able to track back the renormalization needed to handle the singularities as effects living on other scales than those of interest. The basic problem is that of having to handle effects of the microscopic scales and their interaction via the non–linearities on the macroscopic behaviour of the solution.

Mathematically this problem translates in the attempt to make coexists Schwartz theory of distribution with non–linear operations which are notoriously not continuous in the usual topologies on distributions. This is a very old problem of analysis and has been widely studied. The additional input which is not present in the usual approaches is that the singularities which force to treat the problem in the setting of Schwartz's distributions are of a stochastic nature. So we dispose of two handles on the problem: the analytical one and the probabilistic one. The right mix of the two will provide an effective solution to a wide class of problems.

A first and deep understanding of these problems have been obtained starting from the late '90 by T. Lyons [Lyo98] which introduced a theory of *rough paths* in order to settle the conflicts of topology and non–linearity in the context of driven differential equations or more in general in the context of the non–linear analysis of time–varying signals. Nowadays there are a lot of expositions of this theory [LQ02, FV10, LCL07] and we refer the reader to the literature for more details.

In [Gub04, Gub10] the notion of *controlled path* has been introduced in order to extend the applicability of the rough path ideas to a larger class of problems which are not necessarily related to integration of ODEs but which still retained the one–dimensional nature of the directions in which the irregularity manifest itself. The controlled path approach has been used to define some evolution of irregular objects like vortex filaments and some SPDEs. Later Hairer understood how to apply these ideas to the long standing problem of the Kardar–Parisi–Zhang equation [Hai13] and his insights prompted the researchers to trying more ambitious approaches to extend rough path ideas to multidimensional setting.

In [GIP13], in collaboration with P. Imkeller, we introduced a notion of *paracontrolled distributions* suitable to handle a wide class of SPDEs. At the same time Hairer managed to devise a vast generalization of the basic construction of controlled rough paths in the multidimensional and distributional setting which he called the theory of *regularity structures* [Hai14] and which subsumes standard analysis based on Hölder spaces and controlled rough path theory but goes well beyond.

At this date it seems that the theory of regularity structures has a wider range of applicability than the paracontrolled approach described in [GIP13] but also at the expense of a very deep conceptual sophistication. There are problems (like the 1d heat equation with multiplicative noise and general nonlinearity) that cannot be solved via paracontrolled distributions but these problems seems also quite difficult (even if doable and work in progress) also via regularity structures. Moreover equations of more general kind (dispersive equations, wave equations) are still poorly (or not at all) understood in these approaches.

Just few days after the lectures at Mambucaba took place it was annouced that Martin Hairer was awared a Fields Medal for his work on SPDEs and in particular for his theory of regularity structures [Hai14] for dealing with singular SPDEs. This prize witness the exciting period we are experiencing: we now understand sound lines of attack to old standing problems and new opportunities to apply similar ideas to new problems.

The plan of the lectures is the following. We start by explaining the notion of "energy solutions" [GJ10, GJ13] which is a notion of solution to (a particular class of) singular PDEs which has the advantage to be quite easy to handle but also that has the inconvenient not the have a comprehensive uniqueness theory to this date. This will allows us to introduce the readed to SPDEs in a quite progressive way and also to introduce Gaussian tools (Wick's products, hypercontractivity) and some of the basic phenomena appearing when dealing with singular SPDEs. Next we set up the analytical tools we need in the rest of the lectures: Besov spaces and some basic harmonic analysis via the Littlewood–Paley decomposition. Next, in order to motivate the readers and provide a physical ground for the intuition to stand on, we discuss the homogeneisation problem for the linear heat equation with random potential. This will allow us to track the need of the weak topologies and of irregular objects like white–noise from first principles and "concrete" applications. The homogeneisation problem allows also to see naturally appear the renormalization effects into the picture and track their mathematical meaning. Starting from these problems we introduce the 2d parabolic Anderson model which is the simpler SPDEs in which most of the features of more difficult problems are already present and explain how to us paraproducts and the paracontrolled Ansatz in order to keep under control the non–linear effect of the singular data. Then we go on to discuss the more involved situation of the Stochastic Burgers equation in 1d which is one of the avatars of the Kardar–Parisi–Zhang equation.

Conventions and notations We write $a \lesssim b$ if there exists a constant $C > 0$, independent of the variables under consideration, such that $a \leqslant Cb$. Similarly we define $\gtrsim$. We write $a \simeq b$ if $a \lesssim b$ and $b \lesssim a$. If we want to emphasize the dependence of C on the variable x, then we write $a(x) \lesssim_x b(x)$.

If i and j are index variables of Littlewood-Paley blocks (to be defined below), then $i \lesssim j$ is to be interpreted as $2^i \lesssim 2^j$, and similarly for $\simeq$ and $\lesssim$. In other words, $i \lesssim j$ means $i \leqslant j + N$ for some fixed $N \in \mathbb{N}$ that does not depend on i or j.

We use standard multi-index notation: for $\mu \in \mathbb{N}_0^d$ we write $|\mu| = \mu_1 + \ldots + \mu_d$ and $\partial^\mu = \partial^{|\mu|} / \partial_{x_1}^{\mu_1} \ldots \partial_{x_d}^{\mu_d}$, as well as $x^\mu = x_1 \cdot \ldots \cdot x_d$ for $x \in \mathbb{R}^d$.

For $\alpha > 0$ we write C_b^α for the functions $F: \mathbb{R} \to \mathbb{R}$ which are $\lfloor \alpha \rfloor$ times continuously differentiable with $(\alpha - \lfloor \alpha \rfloor)$–Hölder continuous derivatives of order $\lfloor \alpha \rfloor$.

If we write $u \in \mathscr{C}^{\alpha -}$, then that means that u is in $\mathscr{C}^{\alpha - \varepsilon}$ for all $\varepsilon > 0$. The $\mathscr{C}^\alpha$ spaces will be defined below.

2 Energy solutions

The first issue one encounters dealing with singular SPDEs has to do with the not–well posed character of the equation, even in a weak sense. Typically the non–linearity does not make sense in the natural spaces were solutions lives and one has to provide a suitable smaller space which allow to identify the correct meaning to give to "ambiguous quantities" featuring in the equation.

Energy solution [GJ10, GJ13] are a simple tool in order to come up with a well–defined non–linearities. The drawback is that currently the issue of uniqueness, in the interesting cases, is open. It is not clear if uniqueness of energy solutions holds or even what to do in order to find conditions which ensure uniqueness. On the other hand proving existence of energy solution or even convergence to energy solutions is usually a quite simple problem, at least compared to the other approaches like paracontrolled solutions or regularity structures where existence require already quite a large amount of computations but where uniqueness can be established quite easily afterwards.

Our aim is to motivate the ideas leading to the notion of energy solutions. We will not insist on a detailed formulation of all the available results. The reader can always refer to the original paper [GJ13] for all the missing details. Applications to the large scale behavior of particle systems are studied here [GJ10].

We will study energy solutions for the stochastic Burgers equation on $\mathbb{T}$: the unknown u: $\mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}$ satisfy

$$\partial_t u = \Delta u + \partial_x u^2 + \partial_x \xi$$

where $\xi: \mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}$ is a space–time white noise defined on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ fixed once and for all. The equation has to be understood as a relation for processes which are distributions in space with regular enough time dependence. In particular if we test the above relation with $\varphi \in \mathscr{S}(\mathbb{T})$, denote with $u_t(\varphi)$ the pairing of the distribution $u(t, \cdot)$ with φ and integrate in time in the interval $[0, t]$ we get

$$u_t(\varphi) = u_0(\varphi) + \int_0^t u_s(\Delta\varphi)\mathrm{d}s - \int_0^t \langle u_s^2, \partial_x\varphi\rangle \mathrm{d}s - \int_0^t \xi_s(\partial_x\varphi)\mathrm{d}s$$

Let us discuss the various terms in this equation. In order to make sense of $u_t(\varphi)$ and $\int_0^t u_s(\Delta\varphi)\mathrm{d}s$ it is enough to assume that for all $\varphi \in \mathscr{S}(\mathbb{T})$ the mapping $(t, \omega) \mapsto u_t(\varphi)$ is a stochastic process with continuous trajectories. Next, if we denote $M_t(\varphi) = \int_0^t \xi_s(\partial_x\varphi)\mathrm{d}s$ then, at least by a formal computation, we have that $(M_t(\varphi))_{t \geqslant 0, \varphi \in \mathscr{S}(\mathbb{T})}$ is a Gaussian random field with covariance

$$\mathbb{E}[M_t(\varphi) M_s(\varphi)] = (t \wedge s) \langle \partial_x\varphi, \partial_x\psi \rangle_{L^2(\mathbb{T})}$$

In particular, for every $\varphi \in \mathscr{S}$ the stochastic process $(M_t(\varphi))_{t \geqslant 0}$ is a Brownian motion with covariance $\|\varphi\|^2_{H^1(\mathbb{T})} = \langle \partial_x\varphi, \partial_x\varphi \rangle_{L^2(\mathbb{T})}$. We used the notation M in order to stress the fact that $M_t(\varphi)$ is a martingale for its natural filtration and more in general for the filtration given by $\mathcal{F}_t = \sigma(M_s(\varphi): s \leqslant t, \varphi \in H^1(\mathbb{T}))$. (The quantification over $\varphi \in H^1(\mathbb{T})$ is not allowed but we can use a dense countable $(\varphi_n)_{n \geqslant 0}$ subset of $H^1(\mathbb{T})$).

The most difficult term is of course the nonlinear one: $\int_0^t \langle u_s^2, \partial_x\varphi\rangle \mathrm{d}s$. In order to define it indeed we need to square the distribution u_t, operation which is quite dangerous in general. One natural approach would be to define it as the limit of some regularizations. For example, if we let $\rho: \mathbb{R} \to \mathbb{R}_+$ a compactly supported C^∞ positive function such that $\int_{\mathbb{R}} \rho = 1$ and let $\rho_\varepsilon(\cdot) = \rho(\cdot/\varepsilon)/\varepsilon$ then we can let $\mathcal{N}_{t,\varepsilon}(u)(x) = \int_0^t (\rho_\varepsilon * u_s)(x)^2 \mathrm{d}s$ and define the distribution $\mathcal{N}_t(u) = \lim_{\varepsilon \to 0} \mathcal{N}_{t,\varepsilon}(u)$ whenever the limit exists in $\mathscr{S}'(\mathbb{T})$. Which properties u should have in order for this to occur is the question.

2.1 The Ornstein–Uhlenbeck process

Let us simplify the problem and look at solutions of the linearized equation obtained by neglecting the non–linear term. Let X be a solution to

$$X_t(\varphi) = X_0(\varphi) + \int_0^t X_s(\Delta\varphi)\mathrm{d}s + M_t(\varphi) \tag{1}$$

for all $t \geqslant 0$ and $\varphi \in \mathscr{S}(\mathbb{T})$. This equation has a unique solution (for fixed X_0 and M), indeed the difference D between two solutions should satisfy $D_t(\varphi) = \int_0^t D_s(\Delta\varphi)\mathrm{d}s$ which means that D is a distributional solution to the heat equation. Taking $\varphi(x) = \exp(ikx)/\sqrt{2\pi} = e_k(x)$ for $k \in \mathbb{Z}$ we get $D_t(e_k) = -k^2 \int_0^t D_s(e_k)\mathrm{d}s$ and by Gronwall $D_t(e_k) = 0$ for all $t \geqslant 0$ which easily implies $D_t = 0$ in $\mathscr{S}'$ for all $t \geqslant 0$.

To obtain a solution of the equation observe that

$$X_t(e_k) = X_0(e_k) - k^2 \int_0^t X_s(e_k) \mathrm{d}s + M_t(e_k)$$

and that $M_t(e_0) = 0$ while for all $k \neq 0$ the process $\beta_t(k) = M_t(e_k) / (i\, k)$ is a complex Brownian motion with covariance

$$\mathbb{E}[\beta_t(k) \beta_s(m)] = (t \wedge s) \delta_{k+m=0}$$

and satisfying $\beta_t(k)^* = \beta_t(-k)$ for all $k \neq 0$ and $\beta_t(0) = 0$. Then $X_t(e_k)$ is a 1d Ornstein–Ulhenbeck process which solves a standard linear 1d SDE and has an explicit representation given by the variation of constants formula

$$X_t(e_k) = e^{-k^2 t} X_0(e_k) + i\, k \int_0^t e^{-k^2 (t-s)} \mathrm{d}_s \beta_s(k)$$

and this is enough to determine completely $X_t(\varphi)$ for all $t \geqslant 0$ and $\varphi \in \mathscr{S}$. In particular X is a complex Gaussian random field with mean

$$\mathbb{E}[X_t(e_k)] = e^{-k^2 t} X_0(e_k)$$

and covariance

$$\mathrm{Cov}(X_t(e_k), X_s(e_m)) = k^2 \delta_{k+m=0} \int_0^{t \wedge s} e^{-k^2 (t-r) - k^2 (s-r)} \mathrm{d}r$$

so that

$$X_t(e_k) \sim \mathcal{N}_{\mathbb{C}}(e^{-k^2 t} X_0(e_k), (1 - e^{-2k^2 t})).$$

Sobolev regularity of X is the object of the following lemma.

Lemma 1. *Let $\varepsilon > 0$ and assume that $X_0 \in H^{-1/2-\kappa}(\mathbb{T})$. Then almost surely $X \in CH^{-1/2-\varepsilon}(\mathbb{T})$.*

Proof. Let $\alpha = -1/2 - \varepsilon$ and consider that

$$\|X_t - X_s\|_{H^\alpha(\mathbb{T})}^2 = \sum_{k \in \mathbb{Z}} (1 + |k|^2)^\alpha |X_t(e_k) - X_s(e_k)|^2.$$

Let us estimate the $L^{2p}(\Omega)$ norm of this quantity for $p \in \mathbb{N}$ by writing

$$\mathbb{E}\|X_t - X_s\|_{H^\alpha(\mathbb{T})}^{2p} = \sum_{k_1, \ldots, k_p \in \mathbb{Z}} \prod_{i=1}^p (1 + |k_i|^2)^\alpha \mathbb{E} \prod_{i=1}^p |X_t(e_{k_i}) - X_s(e_{k_i})|^2.$$

By Cauchy–Schwartz

$$\lesssim \sum_{k_1, \ldots, k_p \in \mathbb{Z}} \prod_{i=1}^p (1 + |k_i|^2)^\alpha \prod_{i=1}^p (\mathbb{E}|X_t(e_{k_i}) - X_s(e_{k_i})|^{2p})^{1/p}.$$

Note now that $X_t(e_{k_i}) - X_s(e_{k_i})$ is a Gaussian random variable, so that there exists a universal constant C_p for which

$$\mathbb{E}|X_t(e_{k_i}) - X_s(e_{k_i})|^{2p} \leqslant C_p (\mathbb{E}|X_t(e_{k_i}) - X_s(e_{k_i})|^2)^p$$

and that

$$X_t(e_k) - X_s(e_k) = (e^{-k^2 (t-s)} - 1) X_s(e_k) + i\, k \int_s^t e^{-k^2 (t-r)} \mathrm{d}_r \beta_r(k)$$

$$\mathbb{E}|X_t(e_k) - X_s(e_k)|^2 = (e^{-k^2 (t-s)} - 1)^2 \mathbb{E}|X_s(e_k)|^2 + k^2 \int_s^t e^{-2k^2 (t-r)} \mathrm{d}r$$

$$= (e^{-k^2 (t-s)} - 1)^2 e^{-2k^2 s} |X_0(e_k)|^2 + (e^{-k^2 (t-s)} - 1)^2 k^2 \int_0^s e^{-2k^2 (s-r)} \mathrm{d}r + k^2 \int_s^t e^{-2k^2 (t-r)} \mathrm{d}r$$

$$= (e^{-k^2 t} - e^{-k^2 s})^2 |X_0(e_k)|^2 + \frac{1}{2} (e^{-k^2 (t-s)} - 1)^2 (1 - e^{-2k^2 s}) + \frac{1}{2} (1 - e^{-2k^2 (t-s)})$$

and for any $\kappa > 0$ and $k \neq 0$ we have

$$\mathbb{E}|X_t(e_k) - X_s(e_k)|^2 \lesssim (k^2 (t-s))^\kappa (|X_0(e_k)|^2 + 1)$$

while for $k=0$ we have $\mathbb{E}|X_t(e_0)-X_s(e_0)|^2=0$ then

$$\mathbb{E}\|X_t-X_s\|_{H^\alpha(\mathbb{T})}^{2p}\lesssim\sum_{k_1,\ldots,k_p\in\mathbb{Z}_0}\prod_{i=1}^p(1+|k_i|^2)^\alpha\prod_{i=1}^p\mathbb{E}|X_t(e_{k_i})-X_s(e_{k_i})|^2$$

$$\lesssim(t-s)^{\kappa p}\sum_{k_1,\ldots,k_p\in\mathbb{Z}_0}\prod_{i=1}^p(1+|k_i|^2)^\alpha(k_i^2)^\kappa(|X_0(e_{k_i})|^2+1)$$

$$\lesssim(t-s)^{\kappa p}\left[\sum_{k\in\mathbb{Z}_0}(1+|k|^2)^\alpha(k^2)^\kappa(|X_0(e_k)|^2+1)\right]^p$$

$$\lesssim(t-s)^{\kappa p}(\|X_0\|_{H^{\alpha+\kappa}(\mathbb{T})}^{2p}+[\sum_{k\in\mathbb{Z}_0}(1+|k|^2)^\alpha(k^2)^\kappa]^p)$$

so if $\alpha<-1/2-\kappa$ the sum in the r.h.s. is finite and we obtain an estimation of the modulus of continuity of $t\mapsto X_t$ in $L^{2p}(\Omega;H^\alpha)$:

$$\mathbb{E}\|X_t-X_s\|_{H^\alpha(\mathbb{T})}^{2p}\lesssim(t-s)^{\kappa p}[1+\|X_0\|_{H^{\alpha+\kappa}(\mathbb{T})}^{2p}].$$

Now, by Kolmogorov lemma, we can conclude that for some small $\varepsilon>0$ $X\in CH^{-1/2-\varepsilon}(\mathbb{T})$ almost surely if $X_0\in H^{-1/2-\varepsilon}(\mathbb{T})$. □

Now note that the regularity of the OU process does not allow to form the quantity X_t^2 point–wise in time since by Fourier transform we have $X_t(x)=\sum_k X_t(e_k)e_k^*(x)$ and we should have

$$X_t^2(e_k)=\sum_{\ell+m=k}X_t(e_\ell)X_t(e_m).$$

Of course this expression is formal at this point since we cannot guarantee that the infinite sum converges. A reasonable thing to try is to approximate the square by regularizing the distribution via a convolution with a smooth kernel and then try to remove the regularization. Let Π_N the projector of a distribution on a finite number of Fourier modes:

$$(\Pi_N\rho)(x)=\sum_{|k|\leqslant N}\rho(e_k)e_k^*(x)$$

Note that $(\Pi_N\rho)(\varphi)=\sum_{|k|\leqslant N}\rho(e_k)\hat{\varphi}(k)$. Then $\Pi_N X_t(x)$ is a nice smooth function of x and we can consider $[(\Pi_N X_t)^2](x)$ which satisfies

$$(\Pi_N X_t)^2(e_k)=\sum_{\ell+m=k}\mathbb{I}_{|\ell|\leqslant N,|m|\leqslant N}X_t(e_\ell)X_t(e_m)$$

and then try to take the limit $N\to+\infty$. Below for convenience we will do computations already in the limit $N=+\infty$ but one has to come back to the finite N case in order to make it rigorous.

Now,

$$\mathbb{E}[X_t^2(e_k)]=\delta_{k=0}\sum_{m\in\mathbb{Z}_0}\mathbb{E}[X_t(e_{-m})X_t(e_m)]$$

$$=\delta_{k=0}\sum_{m\in\mathbb{Z}_0}e^{-2m^2t}|X_0(e_m)|^2+\delta_{k=0}\sum_{m\in\mathbb{Z}_0}m^2\int_0^t e^{-2m^2(t-s)}\mathrm{d}s$$

but

$$\sum_{m\in\mathbb{Z}_0}m^2\int_0^t e^{-2m^2(t-s)}\mathrm{d}s=\frac{1}{2}\sum_{m\in\mathbb{Z}_0}\left(1-e^{-2m^2t}\right)=+\infty.$$

This is not really a problem since in the equation only the components with $k\neq0$ of $u_t^2(e_k)$ appears. However $X_t^2(e_k)$ is not even a well–defined random variable. For a moment let us assume that $X_0=0$, this will simplify a bit the computation. Next note that if $k\neq0$ we have

$$\mathbb{E}[|X_t^2(e_k)|^2]=\mathbb{E}[X_t^2(e_k)X_t^2(e_{-k})]=\sum_{\ell+m=k}\sum_{\ell'+m'=k}\mathbb{E}[X_t(e_\ell)X_t(e_m)X_t(e_{\ell'})X_t(e_{m'})]$$

and by Wick's theorem the expectation can be computed in terms of the covariances of all possible pairings of the four Gaussian random variables (3 possible combinations)

$$\mathbb{E}[X_t(e_\ell)X_t(e_m)X_t(e_{\ell'})X_t(e_{m'})] = \mathbb{E}[X_t(e_\ell)X_t(e_m)]\mathbb{E}[X_t(e_{\ell'})X_t(e_{m'})]$$

$$\mathbb{E}[X_t(e_\ell)X_t(e_{\ell'})]\mathbb{E}[X_t(e_m)X_t(e_{m'})] + \mathbb{E}[X_t(e_\ell)X_t(e_{m'})]\mathbb{E}[X_t(e_m)X_t(e_{\ell'})]$$

Since $k \neq 0$ we have $\ell + m \neq 0$ and $\ell' + m' \neq 0$ which allow to neglect the first term since it is zero, by symmetry of the summations the two other give the same contribution so we remain with

$$\begin{aligned}\mathbb{E}[|X_t^2(e_k)|^2] = 2\sum_{\ell+m=k}\sum_{\ell'+m'=k}& \mathbb{E}[X_t(e_\ell)X_t(e_{\ell'})]\mathbb{E}[X_t(e_m)X_t(e_{m'})]\\ =2\sum_{\ell+m=k}& \mathbb{E}[X_t(e_\ell)X_t(e_{-\ell})]\mathbb{E}[X_t(e_m)X_t(e_{-m})]\\ =\frac{1}{2}\sum_{\ell+m=k}& \left(1-e^{-2\ell^2 t}\right)\left(1-e^{-2m^2 t}\right) = +\infty\end{aligned}$$

Showing, at least at the heuristic level, that there will indeed problems with X_t^2.

The OU process can be decomposed as

$$X_t(e_k) = i\,k\int_\infty^t e^{-k^2(t-s)}\mathrm{d}\beta_s(k) - i\,k e^{-k^2 t}\int_{-\infty}^0 e^{k^2 s}\mathrm{d}\beta_s(k)$$

by extending the Brownian motions $(\beta_s(k))_{s\geqslant 0}$ to a two sided complex BM via independent copies. It is not difficult to show that the second term give rise to a smooth function if $t > 0$, so all the irregularity of X_t is described by the first one which we call $Y_t(e_k)$ and then we note that $Y_t(e_k) \sim \mathcal{N}_{\mathbb{C}}(0, 1/2)$ for all $k \in \mathbb{Z}_0$ and $t \in \mathbb{R}$. The random distribution Y_t satisfy then $Y_t(\varphi) \sim \mathcal{N}\big(0, \|\varphi\|_{L^2(\mathbb{T})}^2/2\big)$ that is, it is the white noise on $\mathbb{T}$. It is also possible to deduce that the white noise on $\mathbb{T}$ is really the invariant measure of the OU process and that it is, indeed, the only one and it is approached quite fast.

So we should expect that, at fixed time, the regularity of the OU process is like that of the space white noise and this is a way to understand our difficulties in defining X_t^2 since this will be, modulo smooth terms, the square of the space white noise.

A different matter is to make sense of the time–integral of $\partial_x X_t^2$, let us give it a name and call it $J_t(\varphi) = \int_0^t \partial_x X_s^2(\varphi)\mathrm{d}s$. For $J_t(e_k)$ the computation of its variance gives a quite different result. Proceeding as above we have now

$$\mathbb{E}[|J_t(e_k)|^2] = 2k^2\int_0^t\int_0^t\sum_{\ell+m=k}\mathbb{E}[X_s(e_\ell)X_{s'}(e_{-\ell})]\mathbb{E}[X_s(e_m)X_{s'}(e_{-m})]\mathrm{d}s\mathrm{d}s'$$

and, if $s > s'$,

$$\mathbb{E}[X_s(e_\ell)X_{s'}(e_{-\ell})] = \frac{1}{2}e^{-\ell^2(s-s')}\left(1-e^{-2\ell^2 s'}\right)$$

so

$$\mathbb{E}[|J_t(e_k)|^2] = \frac{k^2}{2}\int_0^t\int_0^t\sum_{\ell+m=k}e^{-(\ell^2+m^2)|s-s'|}\left(1-e^{-2\ell^2(s'\wedge s)}\right)\left(1-e^{-2m^2(s'\wedge s)}\right)\mathrm{d}s\mathrm{d}s'$$

$$\leqslant \frac{k^2}{2}\int_0^t\int_0^t\sum_{\ell+m=k}e^{-(\ell^2+m^2)|s-s'|}\mathrm{d}s\mathrm{d}s' \leqslant k^2 t\sum_{\ell+m=k}\int_0^\infty e^{-(\ell^2+m^2)r}\mathrm{d}r = k^2\,t\sum_{\ell+m=k}\frac{1}{\ell^2+m^2}$$

and now for $k \neq 0$:

$$\sum_{\ell+m=k}\frac{1}{\ell^2+m^2} \lesssim \int_{\mathbb{R}}\frac{\mathrm{d}x}{x^2+(x+k)^2} \lesssim \frac{1}{|k|}$$

so finally $\mathbb{E}[|J_t(e_k)|^2] \lesssim |k|\, t$. Redoing a similar computation in the case $J_t(e_k) - J_s(e_k)$ we obtain $\mathbb{E}[|J_t(e_k) - J_s(e_k)|^2] \lesssim |k|\ (t - s)$. From this estimate to a path–wise regularity result of the distribution $(J_t)_t$, following the line of reasoning of Lemma 1, we need to estimate the p-th moment of $J_t(e_k) - J_s(e_k)$. Gaussian hypercontractivity tells us that all the L^p moments of polynomials for gaussian random variables are equivalent and in particular that

$$\mathbb{E}[|J_t(e_k) - J_s(e_k)|^{2p}] \lesssim_p (\mathbb{E}[|J_t(e_k) - J_s(e_k)|^2])^p$$

so by redoing the estimates of the Lemma we discover that almost surely $J \in C^{1/2-} H^{-1/2-}(\mathbb{T})$. This shows that $\partial_x X_t^2$ exists as a space–time distribution but not as a continuous function of time with values in distributions in space. The key point of this computations is the fact that the OU process decorrelates quite rapidly in time.

The construction of the process J sketched in the computations above does not solve our problem since we need similar properties for the full solution u of the non–linear dynamics (or for some approximations thereof) and all we have done relies on explicit computations and the specific Gaussian features of the OU process. But at least give us an hint that indeed there could a way to make sense of the term $\partial_x u(t, x)^2$ even if only as a space–time distribution and that in doing this we should exploit some decorrelation properties of the dynamics.

We need a replacement for the Gaussian computations used above. This will be provided, in our case, by the stochastic calculus along the time direction. Indeed note that for each $\varphi \in \mathscr{S}$ the process $(X_t(\varphi))_{t \geqslant 0}$ is a semimartingale for the filtration $(\mathcal{F}_t)_{t \geqslant 0}$.

Before proceeding with these computations we need to develop some tools to describe Itô the formula along the OU process. This will be also the occasion to set up some analysis of Gaussian spaces.

2.2 Gaussian computations

For cylindrical functions $F: \mathscr{S}' \to \mathbb{R}$ of the form $F(\rho) = f(\rho(\varphi_1), \ldots, \rho(\varphi_n))$ with $\varphi_1, \ldots, \varphi_n \in \mathscr{S}$ and $f: \mathbb{R}^n \to \mathbb{R}$ at least C_b^2, we have by Itô's formula

$$d_t F(X_t) = \sum_{i=1}^n F_i(X_t) \mathrm{d}X_t(\varphi_i) + \frac{1}{2} \sum_{i,j=1}^n F_{i,j}(X_t) \mathrm{d}\langle X(\varphi_i), X(\varphi_j) \rangle_t$$

where $\langle \rangle_t$ denotes the quadratic covariation of two semimartingales and where $F_i(\rho) = \partial_i f(\rho(\varphi_1), \ldots, \rho(\varphi_n))$ and $F_{i,j}(\rho) = \partial_{i,j}^2 f(\rho(\varphi_1), \ldots, \rho(\varphi_n))$ with ∂_i the derivative with respect to the i-th argument. Now

$$\mathrm{d}\langle X(\varphi_i), X(\varphi_j) \rangle_t = \mathrm{d}\langle M(\varphi_i), M(\varphi_j) \rangle_t = \langle \partial_x \varphi_i, \partial_x \varphi_j \rangle_{L^2(\mathbb{T})} \mathrm{d}t$$

and then

$$\mathrm{d}_t F(X_t) = \sum_{i=1}^n F_i(X_t) \mathrm{d}M_t(\varphi_i) + L_0 F(X_t) \mathrm{d}t$$

Where L_0 is the second–order differential operator defined on cylindrical functions F as

$$L_0 F(\rho) = \sum_{i=1}^n F_i(\rho) \rho(\Delta \varphi_i) + \sum_{i,j=1}^n \frac{1}{2} F_{i,j}(\rho) \langle \partial_x \varphi_i, \partial_x \varphi_j \rangle_{L^2(\mathbb{T})}.$$

Another way to describe the generator is to give its value on the functions $\exp(\eta(\psi))$ which is

$$L_0 e^{\rho(\psi)} = e^{\rho(\psi)} (\rho(\Delta \psi) - \frac{1}{2} \langle \psi, \Delta \psi \rangle_{L^2(\mathbb{T})}).$$

If F, G are two cylindrical functions (which we can take of the form $F(\rho) = f(\rho(\varphi_1), \ldots, \rho(\varphi_n))$ and $G(\rho) = g(\rho(\varphi_1), \ldots, \rho(\varphi_n))$ for the same $\varphi_1, \ldots, \varphi_n \in \mathscr{S}$) we can check that

$$L_0(FG) = (L_0 F) G + F(L_0 G) + \mathcal{E}(F, G)$$

where the quadratic form $\mathcal{E}$ is given by

$$\mathcal{E}(F,G)(\rho)=\sum_{i,j}F_i(\rho)G_j(\rho)\langle\partial_x\varphi_i,\partial_x\varphi_j\rangle_{L^2(\mathbb{T})}.$$

Assume that $X_0=\eta$ where $\eta\sim\mathcal{N}(0,\|\varphi\|^2_{L^2(\mathbb{T})}/2)$ is a space–white noise and recall that we already established that white noise is invariant for the OU dynamics so that $X_t\sim\eta$ for all $t\geqslant 0$.

Lemma 2. *(Gaussian integration by parts) Let $(Z_i)_{i=1,\ldots,M}$ be a M-dimensional Gaussian vector with zero mean and covariance $(C_{i,j})_{i,j=1,\ldots,M}$. Then for all $g\in C^1_b(\mathbb{R}^M)$ we have*

$$\mathbb{E}[Z_k g(Z)]=\sum_\ell C_{k,\ell}\mathbb{E}\left[\frac{\partial g(Z)}{\partial Z_\ell}\right]$$

Proof. Use that $\mathbb{E}[e^{\iota\langle Z,\lambda\rangle}]=e^{-\langle\lambda,C\lambda\rangle/2}$ and moreover that

$$\begin{aligned}\mathbb{E}[Z_k e^{\iota\langle Z,\lambda\rangle}]&=(-\iota)\frac{\partial}{\partial\lambda_k}\mathbb{E}[e^{\iota\langle Z,\lambda\rangle}]=(-\iota)\frac{\partial}{\partial\lambda_k}e^{-\langle\lambda,C\lambda\rangle/2}=\iota(C\lambda)_k e^{-\langle\lambda,C\lambda\rangle/2}\\&=\iota\sum_\ell C_{k,\ell}\lambda_\ell\mathbb{E}[e^{\iota\langle Z,\lambda\rangle}]=\mathbb{E}[\sum_\ell C_{k,\ell}\frac{\partial}{\partial Z_\ell}e^{\iota\langle Z,\lambda\rangle}].\end{aligned}$$

The relation is true for trigonometric functions and by Fourier transform for all $g\in\mathscr{S}$. Is then a matter to take limits to show that we can extend it to any $g\in C^1_b(\mathbb{R}^M)$. □

Note that $\mathbb{E}[\eta(\varphi_i)\eta(\Delta\varphi_j)]=\frac{1}{2}\langle\varphi_i,\Delta\varphi_j\rangle_{L^2(\mathbb{T})}$ so

$$\begin{aligned}\mathbb{E}\sum_{i,j=1}^n\frac{1}{2}F_{i,j}(\eta)\langle\partial_x\varphi_i,\partial_x\varphi_j\rangle_{L^2(\mathbb{T})}&=-\mathbb{E}\sum_{i,j=1}^n\frac{1}{2}F_{i,j}(\eta)\langle\varphi_i,\Delta\varphi_j\rangle_{L^2(\mathbb{T})}\\&=-\frac{1}{2}\sum_{i,j=1}^n\langle\varphi_i,\Delta\varphi_j\rangle_{L^2(\mathbb{T})}\mathbb{E}\frac{\partial}{\partial\eta(\varphi_i)}F_j(\eta)=-\sum_{j=1}^n\mathbb{E}[\eta(\Delta\varphi_j)F_j(\eta)]\end{aligned}$$

which gives again that $\mathbb{E}[L_0F]=0$. And

$$\frac{1}{2}\mathbb{E}[\mathcal{E}(F,G)(\eta)]=-\frac{1}{2}\sum_{i,j}\mathbb{E}(F_i(\eta)G_j(\eta))\langle\varphi_i,\Delta\varphi_j\rangle_{L^2(\mathbb{T})}.$$

$$=-\frac{1}{2}\sum_{i,j}\mathbb{E}((F(\eta)G_j(\eta))_i)\langle\varphi_i,\Delta\varphi_j\rangle_{L^2(\mathbb{T})}+\frac{1}{2}\sum_{i,j}\mathbb{E}(F(\eta)G_{ij}(\eta))\langle\varphi_i,\Delta\varphi_j\rangle_{L^2(\mathbb{T})}$$

$$=-\sum_j\mathbb{E}(F(\eta)G_j(\eta)\eta(\Delta\varphi_j))+\frac{1}{2}\sum_{i,j}\mathbb{E}(F(\eta)G_{ij}(\eta))\langle\varphi_i,\Delta\varphi_j\rangle_{L^2(\mathbb{T})}$$

$$=-\mathbb{E}[(FL_0G)(\eta)]$$

so $\mathbb{E}[(FL_0G)(\eta)]=\mathbb{E}[(G\,L_0F)(\eta)]$, that is L_0 is a symmetric operator with respect to the law of η.

Let D be the operator defined as $\mathrm{D}F(\rho)=\sum_i F_i(\rho)(\varphi_i)$ and note that

$$\mathbb{E}[F(\eta)\langle\psi,\mathrm{D}G(\eta)\rangle]+\mathbb{E}[G(\eta)\langle\psi,\mathrm{D}F(\eta)\rangle]=\sum_i\mathbb{E}[(FG)_i(\eta)\langle\psi,\varphi_i\rangle]=2\,\mathbb{E}[\eta(\psi)(FG)(\eta)]$$

so

$$\mathbb{E}[F(\eta)\langle\psi,\mathrm{D}G(\eta)\rangle]=\mathbb{E}[G(\eta)\langle\psi,\mathrm{D}^*F(\eta)\rangle]$$

with $\mathrm{D}^*F(\rho)=-\mathrm{D}F(\rho)+2\rho$ being the adjoint of D for the $L^2(\mathrm{Law}(\eta))$ scalar product. Let $\mathrm{D}_\psi F=\langle\psi,\mathrm{D}F\rangle$ and similarly for $\mathrm{D}^*_\psi F=-\mathrm{D}_\psi F+2\rho(\psi)$. Then $L_0=\frac{1}{2}\sum_k\mathrm{D}^*_{\Delta^{1/2}e_k}\mathrm{D}_{\Delta^{1/2}e_k}$ for an orthonormal basis $(e_n)_{n\geqslant1}$ of $L^2(\mathbb{T})$ and $[\mathrm{D}^*_\psi,\mathrm{D}_\theta]=2\langle\psi,\theta\rangle_{L^2(\mathbb{T})}$. Note that

$$\begin{aligned}[L_0,\mathrm{D}^*_\psi]&=\frac{1}{2}\sum_k\left[\mathrm{D}^*_{\Delta^{1/2}e_k}\mathrm{D}_{\Delta^{1/2}e_k},\mathrm{D}^*_\psi\right]=\frac{1}{2}\sum_k\mathrm{D}^*_{\Delta^{1/2}e_k}[\mathrm{D}_{\Delta^{1/2}e_k},\mathrm{D}^*_\psi]+\frac{1}{2}\sum_k\left[\mathrm{D}^*_{\Delta^{1/2}e_k},\mathrm{D}^*_\psi\right]\mathrm{D}_{\Delta^{1/2}e_k}\\&=-\sum_k\mathrm{D}^*_{\Delta^{1/2}e_k}\left\langle\psi,\Delta^{1/2}e_k\right\rangle_{L^2(\mathbb{T})}=-\mathrm{D}^*_{\Delta\psi}\end{aligned}$$

so if ψ is an eigenvector of Δ with eigenvalue λ :

$$[L_0, D_{\psi}^*] = -\lambda \mathrm{D}_{\psi}^*.$$

Let now $(\psi_n)_{n\geqslant 0}$ be an orthonormal eigenbasis for Δ with eigenvalues $\Delta\psi_n=\lambda_n$ and consider the functions $H(\psi_{i_1}, ..., \psi_{i_n}) = \mathrm{D}_{\psi_{i_1}}^* \cdots \mathrm{D}_{\psi_{i_n}}^* 1$. Then

$$L_0 H(\psi_{i_1}, ..., \psi_{i_n}) = L_0 \mathrm{D}_{\psi_{i_1}}^* \cdots \mathrm{D}_{\psi_{i_n}}^* 1 = \mathrm{D}_{\psi_{i_1}}^* L_0 \mathrm{D}_{\psi_{i_2}}^* \cdots \mathrm{D}_{\psi_{i_n}}^* 1 - \lambda_{i_1} \mathrm{D}_{\psi_{i_1}}^* \cdots \mathrm{D}_{\psi_{i_n}}^* 1$$
$$= \cdots = -(\lambda_{i_1} + \cdots + \lambda_{i_n}) H(\psi_{i_1}, ..., \psi_{i_n})$$

since $L_0 1 = 0$. Then these functions are eigenfunction for L_0 and the eigenvalues are all the possible combinations of $\lambda_{i_1} + \cdots + \lambda_{i_n}$ for $i_1, ..., i_n \in \mathbb{N}$. We have immediately that these functions are orthogonal for different n. They are actually orthogonal as soon as the indexes i differ since in that case there is an index j which is in one but not in the other and using the fact that $\mathrm{D}_{\psi_j}^*$ is adjoint to D_{ψ_j} and that $D_{\psi_j} G = 0$ if G does not contain $\mathrm{D}_{\psi_j}^*$ we get the orthogonality. These functions are polynomials and they are called Wick polynomials. Note also that

$$\mathbb{E}\big(F(\rho)\, e^{\mathrm{D}_{\psi}^*} 1\big) = \mathbb{E}(e^{\mathrm{D}_{\psi}} F(\rho)\,) = \mathbb{E}[F(\rho+\psi)] = \mathbb{E}\Big[F(\rho) e^{\rho(\psi) - \frac{1}{2}\|\psi\|^2}\Big]$$

so taking $\psi = \sum_i \sigma_i \psi_i$ we get

$$e^{\sum_i \sigma_i \rho(\psi_i) - \sum_i \frac{\sigma_i^2}{2}\|\psi_i\|^2} = e^{\mathrm{D}_{\psi}^*} 1 = \sum_{n\geqslant 0} \frac{(\mathrm{D}_{\psi}^*)^n}{n!} = \sum_{n\geqslant 0} \sum_{i_1, \ldots, i_n} \frac{\sigma_{i_1} \cdots \sigma_{i_n}}{n!} H(\underbrace{\psi_{i_1}, ..., \psi_{i_n}}_{n \text{ times}})$$

which is enough to show that any random variable in L^2 can be expanded in a series of Wick polynomials showing that indeed Wick polynomials are an orthogonal basis of $L^2(\mathrm{Law}(\eta))$ (but they are still not normalized). Indeed assume that $Z \in L^2(\mathrm{Law}(\eta))$ but $Z \perp H(\psi_{i_1}, ..., \psi_{i_n})$ for all $n \geqslant 0$, $i_1, ..., i_n \in \mathbb{N}$, then

$$0 = e^{\sum_i \frac{\sigma_i^2}{2}\|\psi_i\|^2} \mathbb{E}\big[Z(\eta)\, e^{\mathrm{D}_{\psi}^*} 1\big] = e^{\sum_i \frac{\sigma_i^2}{2}\|\psi_i\|^2} \mathbb{E}[Z(\eta) e^{\sum_i \sigma_i \eta(\psi_i) - \sum_i \frac{\sigma_i^2}{2}\|\psi_i\|^2}] = \mathbb{E}[Z(\eta) e^{\sum_i \sigma_i \eta(\psi_i)}]$$

Since the σ_i are arbitrary this means that $Z(\eta)$ is orthogonal to any polynomial in η and then that is orthogonal also to $\exp(\iota \sum_i \sigma_i \eta(\psi_i))$. But then take $\hat{q} \in \mathscr{S}(\mathbb{R}^M)$ and $\sigma_i = 0$ if $i > M$, and observe that

$$0 = \int \mathrm{d}\sigma_1 \cdots \mathrm{d}\sigma_m \hat{q}(\sigma_1, ..., \sigma_m)\, \mathbb{E}[Z(\eta) e^{\iota \sum_i \sigma_i \eta(\psi_i)}] = \mathbb{E}[Z(\eta) q(\eta(\psi_1), ..., \eta(\psi_M))]$$

which means that $Z(\rho)$ is orthogonal to all the random variables in L^2 which are measurable with respect to the σ–field generated by $(\rho(\psi_n))_{n\geqslant 0}$. This implies $Z(\rho) = 0$. That is, Wick polynomials form a basis for L^2.

The first few (un–normalized) Wick polynomials are

$$H(\psi_i) = \mathrm{D}_{\psi_i}^* 1 = 2\rho(\psi_i) = 2\rho(\psi_i)$$

and

$$H(\psi_i, \psi_j) = \mathrm{D}_{\psi_i}^* \mathrm{D}_{\psi_j}^* 1 = 2\mathrm{D}_{\psi_i}^* \rho(\psi_j) = -2\delta_{i,j} + 4\rho(\psi_i)\rho(\psi_j)$$
$$H(\psi_i, \psi_j, \psi_k) = \mathrm{D}_{\psi_i}^*(-2\delta_{j,k} + 4\rho(\psi_j)\rho(\psi_k))$$
$$= -4\delta_{j,k}\rho(\psi_i) - 4\delta_{i,j}\rho(\psi_k) - 4\delta_{i,k}\rho(\psi_j) + 8\rho(\psi_i)\rho(\psi_j)\rho(\psi_k)$$

and so on.

Some other properties of Wick's polynomials can be derived using the commutation relation between D and D*. By linearity $\mathrm{D}_{\varphi+\psi}^* = \mathrm{D}_{\varphi}^* + \mathrm{D}_{\psi}^*$ so

$$H_n(\varphi+\psi) = H(\varphi+\psi, ..., \varphi+\psi) = \sum_{0\leqslant k\leqslant n} \binom{k}{n} H(\underbrace{\varphi, ..., \varphi}_{k}, \underbrace{\psi, ..., \psi}_{n-k})$$

Then note that

$$e^{\rho(\psi)-\|\psi\|^2/2}e^{\rho(\varphi)-\|\varphi\|^2/2}=e^{\rho(\varphi+\psi)-\|\varphi+\psi\|^2/2+\langle\varphi,\psi\rangle}.$$

By expanding the exponentials we have

$$\sum_{n,m}\frac{H_n(\psi)}{n!}\frac{H_m(\varphi)}{m!}=\sum_{r,\ell}\frac{H_r(\varphi+\psi)}{r!}\frac{(\langle\varphi,\psi\rangle)^\ell}{\ell!}=\sum_{n,m,\ell}\frac{H(\overbrace{\varphi,\ldots,\varphi}^{k},\overbrace{\psi,\ldots,\psi}^{m})}{n!m!}\frac{(\langle\varphi,\psi\rangle)^\ell}{\ell!}.$$

Identifying the terms of the same homogeneity respectively in φ and ψ we get

$$H_n(\psi)H_m(\varphi)=\sum_{p+\ell=n}\sum_{q+\ell=m}\frac{n!\,m!}{p!q!\ell!}H(\overbrace{\varphi,\ldots,\varphi}^{p},\overbrace{\psi,\ldots,\psi}^{q})(\langle\varphi,\psi\rangle)^\ell$$

which gives a general formula for the products. By polarization of this multilinear form we can get also a general formula for the products of general Wick polynomials. Indeed taking $\psi=\sum_{i=1}^n\lambda_i\psi_i$ and $\varphi=\sum_{j=1}^m\rho_j\varphi_j$ for arbitrary real coefficients $\lambda_1,\ldots,\lambda_n$ and $\rho_1,\ldots,\rho_m$ we have

$$H_n(\sum_{i=1}^n\lambda_i\psi_i)H_m(\sum_{j=1}^m\rho_j\varphi_j)=\sum_{i_1,\ldots,i_n}\sum_{j_1,\ldots,j_m}\lambda_{i_1}\cdots\lambda_{i_m}\rho_{j_1}\cdots\rho_{j_m}H(\psi_{i_1},\ldots,\psi_{i_n})H(\psi_{j_1},\ldots,\psi_{j_m})$$

and then deriving this wrt all the λ, ρ parameters and setting them to zero we single out the term

$$\sum_{\sigma\in S_n,\omega\in S_m}H(\psi_{\sigma(1)},\ldots,\psi_{\sigma(n)})H(\psi_{\omega(1)},\ldots,\psi_{\omega(m)})=(n!)(m!)H(\psi_1,\ldots,\psi_n)H(\varphi_1,\ldots,\varphi_m)$$

using symmetry of the Wick polynomials. Doing the same also in the r.h.s. we get

$$H(\psi_1,\ldots,\psi_n)H(\varphi_1,\ldots,\varphi_m)=\sum_{p+\ell=n}\sum_{q+\ell=m}\frac{1}{p!q!\ell!}\sum_{i,j}H(\overbrace{\varphi_{i_1},\ldots,\varphi_{i_p}}^{p},\overbrace{\psi_{j_1},\ldots,\psi_{j_q}}^{q})\prod_{r=1}^{\ell}\langle\varphi_{i_{p+r}},\psi_{j_{q+r}}\rangle$$

where the sum over i,j runs over $i_1,\ldots,i_n$ permutation of $1,\ldots,n$ and similarly for $j_1,\ldots,j_m$.

In particular

$$\mathbb{E}[H(\psi_1,\ldots,\psi_n)H(\psi_1,\ldots,\psi_n)]=\frac{1}{n!}\sum_{i,j}\prod_{r=1}^n\langle\psi_{i_r},\psi_{j_r}\rangle=\sum_{\sigma\in S_n}\prod_{r=1}^n\langle\psi_r,\psi_{\sigma(r)}\rangle.$$

Some remarks about complex valued bases. In our problem it will be convenient to take the Fourier basis as basis in the above computations. Let $e_k(x)=\exp(\iota k\,x)/\sqrt{2\pi}=a_k(x)+i\,b_k(x)$ where a_k, b_k with $k\in\mathbb{N}$ are a real ONB for $L^2(\mathbb{T})$. Then $\eta(e_k)^*=\eta(e_{-k})$ and we will denote $\mathrm{D}_k=\mathrm{D}_{e_k}=\mathrm{D}_{a_k}+i\mathrm{D}_{b_k}$ and similarly for $\mathrm{D}_k^*=\mathrm{D}_{a_k}^*-i\mathrm{D}_{b_k}^*=-\mathrm{D}_{-k}+\rho(e_{-k})$. In this way D_k^* is the adjoint of D_k with respect to the hermitian scalar product on $L^2(\Omega;\mathbb{C})$ and the OU generator takes the form

$$L_0=\sum_{k\in\mathbb{N}}\mathrm{D}_{\partial_x a_k}^*\mathrm{D}_{\partial_x a_k}+\mathrm{D}_{\partial_x b_k}^*\mathrm{D}_{\partial_x b_k}=\frac{1}{2}\sum_{k\in\mathbb{Z}}k^2\mathrm{D}_k^*\mathrm{D}_k$$

and

$$\mathcal{E}(F,G)=\frac{1}{2}\sum_{k\in\mathbb{Z}}k^2(\mathrm{D}_kF)^*(\mathrm{D}_kG).$$

2.3 The Itô trick

We are ready now to start our computations. Recall that we want to analyse $J_t(\varphi)=\int_0^t\partial_xX_s^2(\varphi)\mathrm{d}s$ using Itô calculus over the OU process. We want to see J_t as a correction term in an Itô formula so we have to find a function F such that $\frac{1}{2}L_0F(X_t)(e_k)=\partial_xX_t^2(e_k)$.Note that

$$\partial_xX_t^2(e_k)=i\,k\sum_{\ell+m=k}X_t(e_\ell)X_t(e_m)=i\,k\sum_{\ell+m=k}H_{\ell,m}(X_t)$$

where $H_{\ell,m}(\rho)=\frac{1}{4}\mathrm{D}^*_{-\ell}\mathrm{D}^*_{-m}1=\rho(e_\ell)\rho(e_m)-\frac{1}{2}\delta_{\ell+m=0}$ is a second order Hermite polynomial so that $L_0 H_{\ell,m}=-(\ell^2+m^2)H_{\ell,m}$. So it is enough to take

$$F(X_t)(e_k)=-2\,\iota\,k\sum_{\ell+m=k}\frac{H_{\ell,m}(X_t)}{\ell^2+m^2}$$

Note that this corresponds to the distribution: $F(X_t)(\varphi)=-2\int_0^\infty \partial_x(e^{\Delta s}X_t)^2(\varphi)\mathrm{d}s$. Then

$$F(X_t)(\varphi)=F(X_0)(\varphi)+M_{F,t}(\varphi)+J_t(\varphi)$$

where $M_{F,t}(\varphi)$ is a martingale with quadratic variation

$$\langle M_{F,*}(\varphi),M_{F,*}(\varphi)\rangle_t=\mathcal{E}(F(*)(\varphi),F(*)(\varphi))(X_t)\mathrm{d}t.$$

We can estimate

$$\mathbb{E}[|J_t(\varphi)-J_s(\varphi)|^{2p}]\lesssim_p\mathbb{E}[|M_{F,t}(\varphi)-M_{F,s}(\varphi)|^{2p}]+\mathbb{E}[|F(X_t)(\varphi)-F(X_s)(\varphi)|^{2p}]$$

Note moreover that if m_t is a martingale we have

$$\mathrm{d}_t|m_t|^{2p}=(2p)|m_t|^{2p-1}\mathrm{d}m_t+\frac{1}{2}(2p)(2p-1)|m_t|^{2p-2}\mathrm{d}\langle m\rangle_t$$

and

$$\mathbb{E}[|m_t|^{2p}]=C_p\int_0^t\mathbb{E}(|m_s|^{2p-2}\mathrm{d}\langle m\rangle_s)\leqslant C_p\mathbb{E}(|m_t|^{2p-2}\langle m\rangle_t)$$

by Cauchy–Schwartz:

$$\leqslant C_p\mathbb{E}(|m_t|^{2p})^{(2p-2)/2p}(\mathbb{E}[\langle m_t\rangle^p])^{1/p}$$

which implies that $\mathbb{E}[|m_t|^{2p}]\leqslant C_p\mathbb{E}[\langle m_t\rangle^p]$. So

$$\begin{aligned}\mathbb{E}[|J_t(\varphi)-J_s(\varphi)|^{2p}]&\lesssim_p\mathbb{E}\left[\left|\int_s^t\mathcal{E}(F(*)(\varphi),F(*)(\varphi))(X_r)\mathrm{d}r\right|^p\right]+\mathbb{E}[|F(X_t)(\varphi)-F(X_s)(\varphi)|^{2p}]\\&\lesssim_p(t-s)^{p-1}\int_s^t\mathbb{E}[|\mathcal{E}(F(*)(\varphi),F(*)(\varphi))(X_r)|^p]\mathrm{d}r+\mathbb{E}[|F(X_t)(\varphi)-F(X_s)(\varphi)|^{2p}]\\&\lesssim_p(t-s)^p\mathbb{E}[|\mathcal{E}(F(*)(\varphi),F(*)(\varphi))(\eta)|^p]+\mathbb{E}[|F(X_t)(\varphi)-F(X_s)(\varphi)|^{2p}]\end{aligned}$$

since $X_t\sim\eta$. Now

$$\mathrm{D}_mF(\rho)(e_k)=-2\,\iota\,k\sum_{\ell+m=k}\frac{\rho(e_\ell)}{\ell^2+m^2}$$

$$\mathcal{E}(F(*)(e_k),F(*)(e_k))(\rho)=\sum_m m^2\mathrm{D}_{-m}F(\rho)(e_{-k})\mathrm{D}_mF(\rho)(e_k)$$

$$=4k^2\sum_{\ell+m=k}m^2\frac{|\rho(e_\ell)|^2}{(\ell^2+m^2)^2}\lesssim k^2\sum_{\ell+m=k}\frac{|\rho(e_\ell)|^2}{\ell^2+m^2}$$

Which implies that

$$\mathbb{E}[|\mathcal{E}(F(*)(e_k),F(*)(e_k))(\eta)|]\lesssim k^2\mathbb{E}\sum_{\ell+m=k}\frac{|\eta(e_\ell)|^2}{\ell^2+m^2}\lesssim k^2\sum_{\ell+m=k}\frac{1}{\ell^2+m^2}\lesssim|k|.$$

A similar computation gives also that

$$\mathbb{E}[|\mathcal{E}(F(*)(e_k),F(*)(e_k))(\eta)|^p]\lesssim|k|^p$$

Note that we have also

$$\mathbb{E}[|F(X_t)(e_k)-F(X_s)(e_k)|^2]\lesssim k^2\sum_{\ell+m=k}\mathbb{E}\left[\frac{(H_{\ell,m}(X_t)-H_{\ell,m}(X_s))^2}{(\ell^2+m^2)^2}\right]$$

$$\lesssim k^2|t-s|\sum_{\ell+m=k}\frac{m^2}{(\ell^2+m^2)^2}\lesssim|k|\,|t-s|$$

And finally, this computation let us recover the result that

$$\mathbb{E}[|J_t(e_k) - J_s(e_k)|^{2p}] \lesssim_p (t-s)^p |k|^p.$$

The advantage of the Itô trick with respect to the explicit Gaussian computation is that it goes over to the non-Gaussian case. Note indeed that u satisfy the Itô formula

$$\mathrm{d}_t F(u_t) = \sum_{i=1}^n F_i(u_t) \mathrm{d}M_t(\varphi_i) + LF(u_t)\mathrm{d}t$$

where L is now the full generator of the non–linear dynamics given by

$$LF(\rho) = L_0 F(\rho) + \sum_i F_i(\rho) \langle \partial_x \rho^2, \varphi_i \rangle = L_0 F(\rho) + BF(\rho)$$

where

$$BF(\rho) = \sum_k (\partial_x \rho^2)(e_k) \mathrm{D}_k F(\rho).$$

The non–linear term is antisymmetric with respect to the invariant measure of L_0:

$$\begin{aligned}\sum_i \mathbb{E}[G(\eta) F_i(\eta) \langle \partial_x \eta^2, \varphi_i \rangle] &= \sum_i \mathbb{E}[(G F)_i(\eta) \langle \partial_x \eta^2, \varphi_i \rangle] - \sum_i \mathbb{E}[G_i(\eta) F(\eta) \langle \partial_x \eta^2, \varphi_i \rangle] \\ &= \frac{1}{2} \mathbb{E}[(G F)(\eta) \langle \partial_x \eta^2, \eta \rangle] = \frac{1}{2} \mathbb{E}[(G F)(\eta) \langle \partial_x \eta^3, 1 \rangle] = 0\end{aligned}$$

Moreover if we reverse the process in time letting $\hat{u}_t = u_{T-t}$ we have

$$\mathbb{E}[F(\hat{u}_t) G(\hat{u}_0)] = \mathbb{E}[F(u_{T-t}) G(u_T)] = \mathbb{E}[F(u_0) G(u_t)].$$

So if we denote by $\hat{L}$ the generator of $\hat{u}$ we have

$$\mathbb{E}\big[\hat{L} F(\hat{u}_0) G(\hat{u}_0)\big] = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathbb{E}[F(\hat{u}_t) G(\hat{u}_0)] = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathbb{E}[F(u_0) G(u_t)] = \mathbb{E}[L G(u_0) F(u_0)]$$

which means that $\hat{L}$ is the adjoint of L, that is

$$\hat{L} = L_0 F(\rho) - \sum_k (\partial_x \rho^2)(e_k) \mathrm{D}_k F(\rho).$$

Then the Itô formula gives

$$\mathrm{d}_t F(\hat{u}_t) = \sum_{i=1}^n F_i(\hat{u}_t) \mathrm{d}\hat{M}_t(\varphi_i) + \hat{L} F(\hat{u}_t) \mathrm{d}t.$$

So

$$\begin{aligned} F(u_T)(\varphi) &= F(u_0)(\varphi) + M_{F,T}(\varphi) + \int_0^T LF(u_s)(\varphi) \mathrm{d}s \\ F(u_0)(\varphi) = F(\hat{u}_T)(\varphi) &= F(\hat{u}_0)(\varphi) + \hat{M}_{F,T}(\varphi) + \int_0^T \hat{L} F(\hat{u}_s)(\varphi) \mathrm{d}s \\ &= F(u_T)(\varphi) + \hat{M}_{F,T}(\varphi) + \int_0^T \hat{L} F(u_s)(\varphi) \mathrm{d}s \end{aligned}$$

summing up these two equalities we get

$$0 = M_{F,T}(\varphi) + \hat{M}_{F,T}(\varphi) + \int_0^T (\hat{L} + L) F(u_s)(\varphi) \mathrm{d}s$$

That is

$$2 \int_0^T L_0 F(u_s)(\varphi) \mathrm{d}s = -M_{F,T}(\varphi) - \hat{M}_{F,T}(\varphi)$$

And as above if $L_0 F(\rho) = \partial_x \rho^2$ we end up with

$$\int_0^T \partial_x u_s^2(\varphi) \mathrm{d}s = -M_{F,T}(\varphi) - \hat{M}_{F,T}(\varphi).$$

A similar computation the allow to establish that even in the non–linear case if we set

$$\mathcal{N}_t^N(\varphi) = \int_0^t \partial_x(\Pi_N u_s)^2(\varphi)\mathrm{d}s$$

then

$$\mathbb{E}[|\mathcal{N}_t^N(e_k) - \mathcal{N}_s^N(e_k)|^{2p}] \lesssim_p (t-s)^p |k|^p.$$

and moreover, adapting the computation one can also show that letting $\mathcal{N}_t^{N,M} = \mathcal{N}_t^N - \mathcal{N}_t^M$ we have

$$\mathbb{E}\big[\big|\mathcal{N}_t^{N,M}(e_k) - \mathcal{N}_s^{N,M}(e_k)\big|^{2p}\big] \lesssim_p (|k|/N)^{\varepsilon p}(t-s)^p |k|^p.$$

for all $1 \leqslant N \leqslant M$ from which we can derive that

$$\big(\mathbb{E}\big[\|\mathcal{N}_t^{N,M} - \mathcal{N}_s^{N,M}\|_{H^\alpha}^{2p}\big]\big)^{1/2p} \lesssim_{p,\alpha} N^{-\varepsilon/2}(t-s)^{1/2}$$

for all $\alpha < -1 - \varepsilon$. Realize that this estimate allows you to prove compactness of the approximations $\mathcal{N}^N$ and then convergence to a limit in $L^{2p}(\Omega; C^{1/2-}H^{-1-})$ which we call $\mathcal{N}$.

2.4 Controlled distributions

Let us cook up a definition which will allow us to perform the computations above in a general setting.

Definition 3. *Let $u, \mathcal{A}: \mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}$ a couple of generalized process such that*

i. For all $\varphi \in \mathscr{S}(\mathbb{T})$ the process $t \mapsto u_t(\varphi)$ is a continuous semi–martingale satisfying

$$u_t(\varphi) = u_0(\varphi) + \int_0^t u_s(\Delta\varphi)\mathrm{d}s + \mathcal{A}_t(\varphi) + M_t(\varphi)$$

where $t \mapsto M_t(\varphi)$ is a martingale with quadratic variation $\langle M_t(\varphi), M_t(\psi)\rangle = \langle \partial_x\varphi, \partial_x\psi\rangle_{L^2}$ and $t \mapsto \mathcal{A}_t(\varphi)$ is a finite variation process.

ii. For all $t \geqslant 0$ the random distribution $\varphi \mapsto u_t(\varphi)$ is a zero mean space white noise with covariance $\|\varphi\|_{L_0^2}^2/2$.

iii. For any $T > 0$ the reverse process $\hat{u}_t = u_{T-t}$ has again properties i, ii with martingale $\hat{M}$ and finite variation part $\hat{\mathcal{A}}$ such that $\hat{\mathcal{A}}_t(\varphi) = -\mathcal{A}_t(\varphi)$.

Any pair of processes $(u, \mathcal{A})$ satisfying these condition will be called controlled by the OU process and we will denote the set of all these processes with $\mathcal{Q}_{\mathrm{ou}}$.

Theorem 4. *Assume that $(u, \mathcal{A}) \in \mathcal{Q}_{\mathrm{ou}}$ and for any $N \geqslant 1$, $t \geqslant 0$, $\varphi \in \mathscr{S}$ let*

$$\mathcal{N}_t^N(\varphi) = \int_0^t \partial_x(\Pi_N u_s)^2(\varphi)\mathrm{d}s$$

Then for any $p \geqslant 1$ $(\mathcal{N}^N)_{N\geqslant 1}$ converges in probability to a space–time distribution $\mathcal{N} \in C^{1/2-}H^{-1-}$.

We are now at a point were we can give a meaning to our original equation.

Definition 5. *A pair of random distribution $(u, \mathcal{A}) \in \mathcal{Q}_{\mathrm{ou}}$ is an* energy solution *to the stochastic Burgers equation if it satisfies*

$$u_t(\varphi) = u_0(\varphi) + \int_0^t u_s(\Delta\varphi)\mathrm{d}s + \mathcal{N}_t(\varphi) + M_t(\varphi)$$

for all $t \geqslant 0$ and $\varphi \in \mathscr{S}$. That is if $\mathcal{A} = \mathcal{N}$.

Now we are in a relatively standard setting of needing to prove existence and uniqueness of such energy solutions. Note that in general the solutions are *pairs* of processes $(u, \mathcal{A})$.

Remark 6. The notion of energy solution has been introduced in the work of Gonçalves and Jara on macroscopic fluctuations of weakly asymmetric interacting particle systems.

2.5 Existence of solutions

For the existence the way to proceed is quite standard. We approximate the equation, construct approximate solutions and then try to have enough compactness to have limiting points which then naturally will satisfy the requirements for energy solutions. For any $N \geqslant 1$ consider solutions u^N to

$$\partial_t u^N = \Delta u^N + \partial_x \Pi_N (\Pi_N u^N)^2 + \partial_x \xi$$

These are generalized functions such that

$$\mathrm{d} u_t^N(e_k) = -k^2 u_t^N(e_k)\mathrm{d}t + [\partial_x \Pi_N(\Pi_N u^N)^2](e_k)\mathrm{d}t + i\,k\,\mathrm{d}\beta_t(k)$$

for $k \in \mathbb{Z}$ and $t \geqslant 0$. We take u_0 to be the white noise with covariance $u_t(\varphi) \sim \mathcal{N}(0, \|\varphi\|^2/2)$. The point of our choice of the non–linearity is that this (infinite–dimensional) system of equations decomposes into a finite dimensional system for $(v^N(k) = \Pi_N u^N(e_k))_{k:|k| \leqslant N}$ and an infinite number of one–dimensional equations for each $u^N(e_k)$ with $|k| > N$. Indeed if $|k| > N$ we have $[\partial_x \Pi_N(\Pi_N u^N)^2](e_k) = 0$ so $u_t(e_k) = X_t(e_k)$ the OU process with initial condition $X_0 = u_0$ which renders it stationary in time (check it). The equation for $(v^N(k))_{|k| \leqslant N}$ reads

$$\mathrm{d} v_t^N(k) = -k^2 v_t^N(k)\mathrm{d}t + b_k(v_t^N)\mathrm{d}t + i\,k\,\mathrm{d}\beta_t(k), \qquad |k| \leqslant N, t \geqslant 0$$

where

$$b_k(v_t^N) = i\,k \sum_{\ell + m = k} \mathbb{I}_{|\ell|, |k|, |m| \leqslant N} v_t^N(\ell) v_t^N(m).$$

This is a standard finite–dimensional ODE having global solutions for all initial conditions and which give rise to a nice Markov process. The fact that solutions do not blow up even if the interaction is quadratic it can be seen by computing the evolution of the norm

$$A_t = \sum_{|k| \leqslant N} |v_t^N(k)|^2$$

and showing that

$$\mathrm{d}A_t = 2 \sum_{|k| \leqslant N} v_t^N(-k) \mathrm{d} v_t^N(-k) = -2k^2 A_t \mathrm{d}t + 2 \sum_{|k| \leqslant N} v_t^N(-k) b_k(v_t^N)\mathrm{d}t + 2\iota\,k \sum_{|k| \leqslant N} v_t^N(-k)\mathrm{d}\beta_t(k)$$

but now

$$\sum_{|k| \leqslant N} v_t^N(-k) b_k(v_t^N) = 2\,\iota \sum_{k,\ell,m:\ell+m=k} \mathbb{I}_{|\ell|,|k|,|m| \leqslant N}\, k\, v_t^N(\ell) v_t^N(m) v_t^N(-k)$$

$$= -2\,\iota \sum_{k,\ell,m:\ell+m+k=0} \mathbb{I}_{|\ell|,|k|,|m| \leqslant N}\,(k)\, v_t^N(\ell) v_t^N(m) v_t^N(k)$$

but by symmetry of this expression it equals to

$$= -\frac{2}{3}\iota \sum_{k,\ell,m:\ell+m+k=0} \mathbb{I}_{|\ell|,|k|,|m| \leqslant N}\,(k + \ell + m)\, v_t^N(\ell) v_t^N(m) v_t^N(k) = 0$$

so $A_t = A_0 + M_t$ where $\mathrm{d}M_t = 2\sum_{|k| \leqslant N} \mathbb{I}_{|k| \leqslant N}(\iota k) v_t^N(-k)\mathrm{d}\beta_t(k)$. Now

$$\mathbb{E}[M_T^2] \lesssim \int_0^T \sum_{|k| \leqslant N} k^2 |v_t^N(k)|^2 \mathrm{d}t \lesssim N^2 \int_0^T A_t \mathrm{d}t$$

and then by martingales inequalities

$$\mathbb{E}[\sup_{t \in [0,T]} (A_t)^2] \leqslant 2\mathbb{E}[A_0^2] + 2\mathbb{E}[\sup_{t \in [0,T]} (M_t)^2] \leqslant 2\mathbb{E}[A_0^2] + 2\mathbb{E}[M_T^2]$$

$$\leqslant 2\mathbb{E}[A_0^2] + C N^2 \int_0^T \mathbb{E}(A_t)\mathrm{d}t$$

and by Gronwall inequality

$$\mathbb{E}[\sup_{t\in[0,T]}(A_t)^2]\lesssim e^{CN^2T}\mathbb{E}[A_0^2].$$

From which we can deduce (by a continuation argument) that almost surely there is no blowup at finite time for the dynamics. From the Galerkin approximations the Itô trick will provide enough compactness in order to pass to the limit and build an energy solution to the Stochastic Burgers equation.

3 Distributions and Besov spaces

Here we collect some classical results from harmonic analysis which we will need in the following. Fix $d\in\mathbb{N}$ and denote by $\mathbb{T}^d=(\mathbb{R}/(2\pi\mathbb{Z}))^d$ the d-dimensional torus. We concentrate here on distributions and SPDEs on the torus, but everything in this Section applies mutatis mutandis on the full space $\mathbb{R}^d$, see [GIP13]. The only problem is that then the stochastic terms will no longer be in the Besov spaces $\mathscr{C}^\alpha$ which we encounter below but rather in weighted Besov spaces. Since we did not develop paracontrolled distributions on weighted Besov spaces yet, we are currently unable to solve SPDEs on $\mathbb{R}^d$.

The space of distributions $\mathscr{D}'=\mathscr{D}'(\mathbb{T}^d)$ is defined as the set of linear maps f from $C^\infty=C^\infty(\mathbb{T}^d,\mathbb{C})$ to $\mathbb{C}$, such that there exist $k\in\mathbb{N}$ and $C>0$ with

$$|\langle f,\varphi\rangle|:=|f(\varphi)|\leqslant C\sup_{|\mu|\leqslant k}\|\partial^\mu\varphi\|_{L^\infty(\mathbb{T}^d)}$$

for all $\varphi\in C^\infty$. In particular, the Fourier transform $\mathscr{F}f:\mathbb{Z}^d\to\mathbb{C}$, $\mathscr{F}f(k)=\langle f,e^{-ik\cdot}\rangle$, is defined for all $f\in\mathscr{D}'$, and it satisfies $|\mathscr{F}f(k)|\leqslant|P(k)|$ for a suitable polynomial P. We will also write $\hat{f}(k)=\mathscr{F}f(k)$. Conversely, if $(g(k))_{k\in\mathbb{Z}^d}$ is at most of polynomial growth, then its inverse Fourier transform

$$\mathscr{F}^{-1}g=(2\pi)^{-d}\sum_{k\in\mathbb{Z}^d}e^{i\langle k,\cdot\rangle}g(k)$$

defines a distribution, and we have $\mathscr{F}^{-1}\mathscr{F}f=f$ as well as $\mathscr{F}\mathscr{F}^{-1}g=g$. To see this, it suffices to note that the Fourier transform of $\varphi\in C^\infty$ decays faster than any rational function (we say that it is of *rapid decay*). Indeed, for $\mu\in\mathbb{N}_0^d$ we have $|k^\mu\hat{g}(k)|=|\mathscr{F}(\partial^\mu g)(k)|\leqslant\|\partial^\mu g\|_{L^1(\mathbb{T}^d)}$ for all $k\in\mathbb{Z}^d$. As a consequence we get the Parseval formula $\langle f,\overline{\varphi}\rangle=(2\pi)^{-d}\sum_k\hat{f}(k)\overline{\hat{\varphi}(k)}$ for $f\in\mathscr{D}'$ and $\varphi\in C^\infty$.

Linear maps on $\mathscr{D}'$ can be defined by duality: if A: $C^\infty\to C^\infty$ is such that for all $k\in\mathbb{N}$ there exists $n\in\mathbb{N}$ and $C>0$ with $\sup_{|\mu|\leqslant k}\|\partial^\mu(A\varphi)\|_{L^\infty}\leqslant C\sup_{|\mu|\leqslant n}\|\partial^\mu\varphi\|$, then we set $\langle {}^tAf,\varphi\rangle=\langle f,A\varphi\rangle$. Differential operators are defined by $\langle\partial^\mu f,\varphi\rangle=(-1)^{|\mu|}\langle f,\partial^\mu\varphi\rangle$. If $\varphi:\mathbb{Z}^d\to\mathbb{C}$ grows at most polynomially, then it defines a *Fourier multiplier*

$$\varphi(\mathrm{D})f=\mathscr{F}^{-1}(\varphi\mathscr{F}f),$$

which gives us a distribution $\varphi(\mathrm{D})f\in\mathscr{D}'$ for every $f\in\mathscr{D}'$.

Example 7. Clearly $L^p=L^p(\mathbb{T}^d)\subset\mathscr{D}'$ for all $p\geqslant 1$, and also the space of finite measures on $(\mathbb{T}^d,\mathscr{B}(\mathbb{T}^d))$ is contained in $\mathscr{D}'$. Another example of a distribution is $\varphi\mapsto\partial^\mu\varphi(x)$ for $\mu\in\mathbb{N}_0^d$ and $x\in\mathbb{T}$.

Exercise 1. Show that for $f\in\mathscr{D}'$, $g\in C^\infty$ and for $u,v:\mathbb{Z}^d\to\mathbb{C}$ with u of polynomial growth and v of rapid decay

$$\mathscr{F}(fg)(k)=(2\pi)^{-d}\sum_\ell\hat{f}(k-\ell)\hat{g}(\ell)\quad\text{and}\quad\mathscr{F}^{-1}(uv)(x)=\int_{\mathbb{T}^d}\mathscr{F}^{-1}u(x-y)\mathscr{F}^{-1}v(y)\mathrm{d}y.$$

Littlewood-Paley blocks give a decomposition of any distribution on $\mathscr{D}'$ into an infinite series of smooth functions. Of course, we have already such a decomposition at our disposal, namely $f=\sum_k(2\pi)^{-d}\hat{f}(k)e^{i\langle k,\cdot\rangle}$. But it turns out to be convenient not to consider each Fourier coefficient separately, but to work with projections on dyadic Fourier blocks.

Definition 8. *A dyadic partition of unity consists of two nonnegative radial functions* χ, $\rho \in C^\infty(\mathbb{R}^d, \mathbb{R})$, *where* ρ *is supported in a ball* $\mathscr{B} = \{|x| \leqslant c\}$ *and* ρ *is supported in an annulus* $\mathscr{A} = \{a \leqslant |x| \leqslant b\}$ *for suitable* $a, b, c > 0$, *such that*

1. $\chi + \sum_{j \geqslant 0} \rho(2^{-j} \cdot) \equiv 1$ *and*
2. $\chi \rho(2^{-j} \cdot) \equiv 0$ *for* $j \geqslant 1$ *and* $\rho(2^{-i} \cdot) \rho(2^{-j} \cdot) \equiv 0$ *for all* $i, j \geqslant 0$ *with* $|i - j| \geqslant 1$.

We will often write $\rho_{-1} = \chi$ *and* $\rho_j = \rho(2^{-j} \cdot)$ *for* $j \geqslant 0$.

Dyadic partitions of unity exist, see [BCD11]. The reason for considering smooth partitions rather than indicator functions is that indicator functions do not have good Fourier multiplier properties. From now on we fix a dyadic partition of unity (χ, ρ) and define the dyadic blocks

$$\Delta_j f = \rho_j(\mathrm{D}) f = \mathscr{F}^{-1}(\rho_j \hat{f}), \quad j \geqslant -1.$$

We also use the notation

$$S_j f = \sum_{i \leqslant j-1} \Delta_i f.$$

Every dyadic block has a compactly supported Fourier transform and is therefore in C^∞. It is easy to see that $f = \sum_{j \geqslant -1} \Delta_j f = \lim_{j \to \infty} S_j f$ for all $f \in \mathscr{D}'$.

For $\alpha \in \mathbb{R}$, the Hölder-Besov space $\mathscr{C}^\alpha$ is given by $\mathscr{C}^\alpha = B^\alpha_{\infty,\infty}(\mathbb{T}^d, \mathbb{R})$, where for $p, q \in [1, \infty]$ we define

$$B^\alpha_{p,q} = B^\alpha_{p,q}(\mathbb{T}^d, \mathbb{R}) = \left\{ f \in \mathscr{D}' : \|f\|_{B^\alpha_{p,q}} = \left(\sum_{j \geqslant -1} (2^{j\alpha} \|\Delta_j f\|_{L^p})^q \right)^{1/q} < \infty \right\},$$

with the usual interpretation as ℓ^∞ norm in case $q = \infty$. Then $B^\alpha_{p,q}$ is a Banach space and while the norm $\|\cdot\|_{B^\alpha_{p,q}}$ depends on (χ, ρ), the space $B^\alpha_{p,q}$ does not, and any other dyadic partition of unity corresponds to an equivalent norm. We write $\|\cdot\|_\alpha$ instead of $\|\cdot\|_{B^\alpha_{\infty,\infty}}$.

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 (see page 99 of [BCD11]). Note however, that for $k \in \mathbb{N}$ the space $\mathscr{C}^k$ is strictly larger than C^k, the space of k times continuously differentiable functions.

Exercise 2. Show that $\|\cdot\|_\alpha \leqslant \|\cdot\|_\beta$ for $\alpha \leqslant \beta$, that $\|\cdot\|_{L^\infty} \lesssim \|\cdot\|_\alpha$ for $\alpha > 0$, that $\|\cdot\|_\alpha \lesssim \|\cdot\|_{L^\infty}$ for $\alpha \leqslant 0$, and that $\|S_j \cdot\|_{L^\infty} \lesssim 2^{j\alpha} \|\cdot\|_\alpha$ for $\alpha < 0$.

Hint: When proving $\|\cdot\|_\alpha \lesssim \|\cdot\|_{L^\infty}$ for $\alpha \leqslant 0$, you might need to bound $\|\mathscr{F}^{-1} \rho_j\|_{L^1(\mathbb{T}^d)}$. Here it may be helpful to use Poisson's summation formula

$$(2\pi)^{-d} \sum_{k \in \mathbb{Z}^d} \rho(2^{-j} k) e^{i \langle k, x \rangle} = (2^j)^d \sum_{k \in \mathbb{Z}^d} (\mathscr{F}^{-1}_{\mathbb{R}^d} \rho)(2^j (2\pi k + x)),$$

where $\mathscr{F}^{-1}_{\mathbb{R}^d} \rho(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i \langle z, x \rangle} \rho(z) \mathrm{d}z$, which holds for all Schwartz functions ρ: $\mathbb{R}^d \to \mathbb{R}$. Alternatively, you can periodically extend $f \in L^\infty(\mathbb{T}^d)$ to $\bar{f} \in L^\infty(\mathbb{R}^d)$ and note that

$$\Delta_j \bar{f}(x) := \mathscr{F}^{-1}_{\mathbb{R}^d}(\rho_j \mathscr{F}_{\mathbb{R}^d} \bar{f})(x) = \int_{\mathbb{R}^d} \mathscr{F}^{-1}_{\mathbb{R}^d} \rho_j(x - y) \bar{f}(y) \mathrm{d}y.$$

Exercise 3. Let δ_0 denote the Dirac delta in 0. Show that $\delta_0 \in \mathscr{C}^{-d}$.

We will often rely on the following characterization of Besov regularity for functions that can be decomposed into pieces which are localized in Fourier space.

Lemma 9.

1. *Let* $\mathscr{A}$ *be an annulus, let* $\alpha \in \mathbb{R}$, *and let* (u_j) *be a sequence of smooth functions such that* $\mathscr{F} u_j$ *has its support in* $2^j \mathscr{A}$, *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 $\mathscr{B}$ be a ball, let $\alpha > 0$, and let (u_j) be a sequence of smooth functions such that $\mathscr{F}u_j$ has its support in $2^j\mathscr{B}$, 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 for $i \sim j$. Hence, we obtain

$$\|\Delta_i u\|_{L^\infty} \leqslant \sum_{j:j\sim i} \|\Delta_i u_j\|_{L^\infty} \leqslant \sup_{k \geqslant -1} \{2^{k\alpha}\|u_k\|_{L^\infty}\} \sum_{j:j\sim i} 2^{-j\alpha} \simeq \sup_{k \geqslant -1} \{2^{k\alpha}\|u_k\|_{L^\infty}\} 2^{-i\alpha}.$$

If $\mathscr{F}u_j$ is supported in $2^j\mathscr{B}$, then $\Delta_i u_j \neq 0$ only for $i \lesssim j$. Therefore,

$$\|\Delta_i u\|_{L^\infty} \leqslant \sum_{j:j\gtrsim i} \|\Delta_i u_j\|_{L^\infty} \leqslant \sup_{k \geqslant -1} \{2^{k\alpha}\|u_k\|_{L^\infty}\} \sum_{j:j\gtrsim i} 2^{-j\alpha} \lesssim \sup_{k \geqslant -1} \{2^{k\alpha}\|u_k\|_{L^\infty}\} 2^{-i\alpha},$$

using $\alpha > 0$ in the last step. □

The following Bernstein inequalities are extremely useful when dealing with functions with compactly supported Fourier transform.

Lemma 10. *Let $\mathscr{A}$ be an annulus and let $\mathscr{B}$ be a ball. For any $k \in \mathbb{N}_0$, $\lambda > 0$, and $1 \leqslant p \leqslant q \leqslant \infty$ we have that*

1. *if $u \in L^p$ is such that $\operatorname{supp}(\mathscr{F}u) \subseteq \lambda\mathscr{B}$, then*

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

2. *if $u \in L^p$ is such that $\operatorname{supp}(\mathscr{F}u) \subseteq \lambda\mathscr{A}$, then*

$$\lambda^k \|u\|_{L^p} \lesssim_k \max_{\mu \in \mathbb{N}^d : |\mu| = k} \|\partial^\mu u\|_{L^p}.$$

It then follows immediately that for $\alpha \in \mathbb{R}$, $f \in \mathscr{C}^\alpha$, $\mu \in \mathbb{N}_0^d$, we have $\partial^\mu f \in \mathscr{C}^{\alpha - |\mu|}$. Another simple application of the Bernstein inequalities is the Besov embedding theorem, the proof of which we leave as an exercise.

Lemma 11. *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 $B^\alpha_{p_1,q_1}$ is continuously embedded into $B^{\alpha - d(1/p_1 - 1/p_2)}_{p_2,q_2}$.*

Exercise 4. Let ξ be a spatial white noise on $\mathbb{T}^d$, i.e. ξ is a centered Gaussian process indexed by $L^2(\mathbb{T}^d)$, with covariance

$$\mathbb{E}[\xi(f)\xi(g)] = \int_{\mathbb{T}^d} f(x) g(x) \mathrm{d}x.$$

Show that there exists $\tilde{\xi}$ with $\mathbb{P}(\tilde{\xi}(f) = \xi(f)) = 1$ for all $f \in L^2$, such that $\mathbb{E}\big[\|\tilde{\xi}\|^p_{-d/\varepsilon - 2}\big] < \infty$ for all $p \geqslant 1$ and $\varepsilon > 0$ (so that $\tilde{\xi} \in \mathscr{C}^{-d/2-}$ almost surely).

Hint: Define $\tilde{\xi} = \mathscr{F}^{-1}\hat{\xi} = (2\pi)^{-d}\sum_k e^{i\langle k,\cdot\rangle}\xi\big(e^{-i\langle k,\cdot\rangle}\big)$ and estimate $\mathbb{E}\Big[\|\tilde{\xi}\|^{2p}_{B^\alpha_{2p,2p}}\Big]$ using Gaussian hypercontractivity (equivalence of moments). Then apply Besov embedding.

When solving SPDEs, we will need the smoothing properties of the heat semigroup. For that purpose we study functions of time with values in distribution spaces. If $\mathbb{X}$ is a Banach space with norm $\|\cdot\|_\mathbb{X}$ and if $T > 0$, then we define $C\mathbb{X}$ and $C_T\mathbb{X}$ as the spaces of continuous functions from $[0, \infty)$ respectively $[0, T]$ to $\mathbb{X}$, and $C_T\mathbb{X}$ is equipped with the supremum norm $\|\cdot\|_{C_T\mathbb{X}}$. If $\alpha \in (0,1)$ then we write $C^\alpha\mathbb{X}$ for the functions in $C\mathbb{X}$ that are α–Hölder continuous on every interval $[0, T]$, and we write

$$\|f\|_{C^\alpha_T\mathbb{X}} = \sup_{0 \leqslant s < t \leqslant T} \frac{\|f(t) - f(s)\|}{|t - s|^\alpha}.$$

We then define $\mathscr{L}^\alpha = C\mathscr{C}^\alpha \cap C^{\alpha/2}L^\infty$ for $\alpha \in (0,2)$. For $T > 0$ we set $\mathscr{L}^\alpha_T = C_T\mathscr{C}^\alpha \cap C^{\alpha/2}_T L^\infty$ and we equip $\mathscr{L}^\alpha_T$ with the norm

$$\|\cdot\|_{\mathscr{L}^\alpha_T} = \max\{\|\cdot\|_{C_T\mathscr{C}^\alpha}, \|\cdot\|_{C^{\alpha/2}_T L^\infty}\}.$$

The notation $\mathscr{L}^\alpha$ is chosen to be reminiscent of the operator $\mathscr{L}$ and indeed the parabolic spaces $\mathscr{L}^\alpha$ are adapted to $\mathscr{L}$ in the sense that the temporal regularity “counts twice”, which is due to the fact that $\mathscr{L} = \partial_t - \Delta$ contains a first order temporal but a second order spatial derivative. If we would replace Δ by a fractional Laplacian $-(-\Delta)^\sigma$, then we would have to consider the space $C\mathscr{C}^\alpha \cap C^{\alpha/(2\sigma)}L^\infty$ instead of $\mathscr{L}^\alpha$.

We have the following Schauder estimate on the scale of $(\mathscr{L}^\alpha)_\alpha$ spaces:

Lemma 12. *Let $\alpha \in (0,2)$ and let $(P_t)_{t \geqslant 0}$ be the semigroup generated by the periodic Laplacian, $\mathscr{F}(P_t f)(k) = e^{-t|k|^2}\mathscr{F}f(k)$. For $f \in C\mathscr{C}^{\alpha-2}$ define $Jf(t) = \int_0^t P_{t-s} f_s \mathrm{d}s$. Then*

$$\|Jf\|_{\mathscr{L}_T^\alpha} \lesssim (1+T)\|f\|_{C_T\mathscr{C}^{\alpha-2}}$$

for all $T > 0$. If $u \in \mathscr{C}^\alpha$, then $t \mapsto P_t u \in \mathscr{L}^\alpha$ and

$$\|t \mapsto P_t u\|_{\mathscr{L}_T^\alpha} \lesssim \|u\|_\alpha.$$

Bibliographic notes For a gentle introduction to Littlewood-Paley theory and Besov spaces see the recent monograph [BCD11], where most of our results are taken from. There the case of tempered distributions on $\mathbb{R}^d$ is considered. The theory on the torus is developed in [ST87]. The Schauder estimates for the heat semigroup are classical and can be found in [GIP13, GP14].

4 Diffusion in a random environment

Let us consider the following d-dimensional homogeneisation problem. Fix $\varepsilon > 0$ and let $u^\varepsilon\colon \mathbb{R}_+ \times \mathbb{T}_\varepsilon \to \mathbb{R}$ be the solution to the Cauchy problem

$$\partial_t u^\varepsilon(t,x) = \Delta u^\varepsilon(t,x) + \varepsilon^{-\alpha} V(x/\varepsilon) u^\varepsilon(t,x), \qquad t \geqslant 0, x \in \mathbb{T}^d$$

with $u^\varepsilon(0,\cdot) = u_0(\cdot)$, where $\mathbb{T} = [0,2\pi]$ is the one dimensional torus and where $V\colon \mathbb{T}_\varepsilon^d \to \mathbb{R}$ is a random field defined on the rescaled torus $\mathbb{T}_\varepsilon^d = \mathbb{T}^d/\varepsilon$. This model describes the diffusion of particles in a random medium (replacing ∂_t by $i\partial_t$ gives the Schrödinger equation of a quantum particle evolving in a random potential). For a review of related results the reader can look at the recent paper of Bal and Gu [BG13]. The limit $\varepsilon \to 0$ corresponds to looking at the long scale behaviour of the model since it can be understood as the equation for the *macroscopic* density $u^\varepsilon(t,x) = u(t/\varepsilon^2, x/\varepsilon)$ which corresponds to a *microscopic* density $u\colon \mathbb{R}_+ \times \mathbb{T}_\varepsilon^d \to \mathbb{R}$ evolving according to the parabolic equation

$$\partial_t u(t,x) = \Delta u(t,x) + \varepsilon^{2-\alpha} V(x) u(t,x), \qquad t \geqslant 0, x \in \mathbb{T}_\varepsilon^d.$$

Let $V_\varepsilon(x) = \varepsilon^{-\alpha} V(x/\varepsilon)$ and assume that $V\colon \mathbb{T}_\varepsilon^d \to \mathbb{R}$ is Gaussian and has mean zero and homogeneous correlation function C_ε given by

$$C_\varepsilon(x-y) = \mathbb{E}[V(x)V(y)] = (\varepsilon/2\pi)^d \sum_{k \in \varepsilon\mathbb{Z}_0^d} e^{i\langle x-y,k\rangle} R(k)$$

where $\mathbb{Z}_0^d = \mathbb{Z}^d \backslash \{0\}$. On $R\colon \mathbb{R}^d \to \mathbb{R}_+$ we make the following hypothesis: for some $\beta \in (0,d]$ we have $R(k) = |k|^{\beta-d}\tilde{R}(k)$ where $\tilde{R} \in \mathscr{S}(\mathbb{R}^d)$ is a smooth bounded function of rapid decrease. When $\beta < d$ this is equivalent to require that spatial correlations (in the limit $\varepsilon \to 0$) decay as $|x|^{-\beta}$. When $\beta = d$ this hypothesis means that spatial correlations decay in an integrable way. Indeed by dominated convergence

$$\lim_{\varepsilon \to 0} C_\varepsilon(x) = \int \frac{\mathrm{d}k}{(2\pi)^d} e^{i\langle x,k\rangle} R(k) = \int \frac{\mathrm{d}k}{(2\pi)^d} e^{i\langle x,k\rangle} |k|^{\beta-d}\tilde{R}(k) = \mathscr{F}^{-1}(|\cdot|^{\beta-d}) * \mathscr{F}^{-1}(\tilde{R})(x)$$

Now $\mathscr{F}^{-1}(\tilde{R}) \in \mathscr{S}(\mathbb{R}^d)$ and $\mathscr{F}^{-1}(|\cdot|^{\beta-d})(x) = |x|^{-\beta}$ if $0 < \beta < d$ so $\lim_{\varepsilon \to 0} C_\varepsilon(x) \lesssim |x|^{-\beta}$ if $|x| \to +\infty$.

Let us now compute the variance of the LP blocks of V_ε.

Lemma 13. *Assume $\beta - 2\alpha \geqslant 0$. We have that for any $\varepsilon > 0$ and $i \geqslant 0$ and any $0 \leqslant \kappa \leqslant \beta - 2\alpha$:*

$$\mathbb{E}[|\Delta_i V_\varepsilon(x)|^2] \lesssim 2^{(2\alpha-\kappa)i}\varepsilon^\kappa$$

This estimate implies that if $\beta > 2\alpha$ we have $V_\varepsilon \to 0$ in $L^2(\Omega; B_{2,2}^{-\alpha-\kappa/2}(\mathbb{T}^d))$ as $\varepsilon \to 0$ for some $\kappa > 0$.

Proof. A spectral computation gives

$$\begin{aligned}
\mathbb{E}[|\Delta_i V_\varepsilon(x)|^2] &= \varepsilon^{-2\alpha}\int_{\mathbb{R}^d\times\mathbb{R}^d} K_i(x-y)K_i(x-z)C_\varepsilon((y-z)/\varepsilon)\mathrm{d}y\mathrm{d}z \\
&= (\varepsilon/2\pi)^d\varepsilon^{-2\alpha}\int_{\mathbb{R}^d\times\mathbb{R}^d} K_i(x-y)K_i(x-z)\sum_{k\in\varepsilon\mathbb{Z}_0^d} e^{i\langle y-z,k/\varepsilon\rangle}R(k)\mathrm{d}y\mathrm{d}z \\
&= (\varepsilon/2\pi)^d\varepsilon^{-2\alpha}\sum_{k\in\varepsilon\mathbb{Z}^d}\rho_i(k/\varepsilon)^2\, e^{i\langle x,k/\varepsilon\rangle}R(k) \qquad (2)\\
&= (2\pi)^d\varepsilon^{d-2\alpha}\sum_{k\in\varepsilon\mathbb{Z}_0^d}\rho(k/(\varepsilon 2^i))^2\, R(k) \\
&\lesssim \varepsilon^{d-2\alpha}2^{id}\sup_{k\in\varepsilon 2^i\mathscr{A}} R(k).
\end{aligned}$$

Now if $\varepsilon 2^i \leqslant 1$ we have $\mathbb{E}[|\Delta_i V_\varepsilon(x)|^2] \lesssim 2^{id}\varepsilon^{d-2\alpha}(\varepsilon 2^i)^{\beta-d} \lesssim \varepsilon^{\beta-2\alpha}2^{i\beta}$. The assumption $\beta-2\alpha\geqslant 0$ implies that $\mathbb{E}[|\Delta_i V_\varepsilon(x)|^2]\lesssim 2^{(2\alpha+\kappa)i}\varepsilon^\kappa$ for any $0\leqslant\kappa\leqslant\beta-2\alpha$. In the case $\varepsilon 2^i>1$, due to the remark that $\int_{B(0,1)^c} R(k)\mathrm{d}k<+\infty$, we can estimate

$$\varepsilon^d\sum_{k\in\varepsilon\mathbb{Z}^d}\rho(k/(\varepsilon 2^i))^2\, R(k)\lesssim\varepsilon^d\sum_{k\in\mathbb{Z}^d}R(\varepsilon k)\lesssim\int_{\mathbb{R}^d}R(k)\mathrm{d}k<+\infty$$

and then $\mathbb{E}[|\Delta_i V_\varepsilon(x)|^2]\lesssim\varepsilon^{-2\alpha}\lesssim 2^{2\alpha i}(\varepsilon 2^i)^\kappa$ for any small $\kappa>0$. □

Note that the computation carried on in eq. (2) implies also that if $\beta-2\alpha<0$ then essentially V_ε does not converge in any reasonable sense since the variance of the LP blocks explodes.

The previous analysis shows that it is reasonable to take $\alpha\leqslant\beta/2$ in order to hope in any well defined limit as $\varepsilon\to 0$. In this case V_ε stay bounded (at least) in spaces of distribution of regularity $\alpha-$. This brings us to the problem of obtaining estimates for the parabolic PDE ($\mathscr{L}=\partial_t-\Delta$)

$$\mathscr{L}u^\varepsilon(t,s)=V_\varepsilon(x)u^\varepsilon(t,x),\qquad t\geqslant 0, x\in\mathbb{T}^d,$$

depending only on negative regularity norms of V_ε. On one side the regularity of u^ε is then limited by the regularity of the r.h.s. which cannot be better than that of V_ε. On the other side the product of V_ε with u_ε can cause problems since we try to multiply an (a–priori) irregular object with one of limited regularity.

Assume that $V_\varepsilon\in\mathscr{C}^{\gamma-2}$ with $\gamma>0$. It is reasonable then to assume also that $V_\varepsilon u^\varepsilon\in C_T\mathscr{C}^{\gamma-2}$ and that $u^\varepsilon\in C_T\mathscr{C}^\gamma$ as a consequence of the regularising effect of the heat operator. We see then that the product $V_\varepsilon u^\varepsilon$ is under control only if $\gamma+\gamma-2>0$, that is if $\gamma>1$. So if $V_\varepsilon\to 0$ in $\mathscr{C}^{-1+}$ then it is not difficult to show that u^ε converges as $\varepsilon\to 0$ to the solution u of the linear equation $\partial_t u(t,x)-\Delta u(t,x)=0$. In this case the random potential will not have any effect in the limit.

The interesting situation then is when $\gamma\leqslant 1$. To understand what could happen in this case let us use a simple transformation of the solution. Write $u^\varepsilon=\exp(X^\varepsilon)v^\varepsilon$ where X^ε satisfy the equation $\mathscr{L}X^\varepsilon=V^\varepsilon$ with initial condition $X^\varepsilon(0,\cdot)=0$. Then

$$\mathscr{L}u^\varepsilon=\exp(X^\varepsilon)(v^\varepsilon\mathscr{L}X^\varepsilon+\mathscr{L}v^\varepsilon-v^\varepsilon(\partial_x X^\varepsilon)^2-\partial_x X^\varepsilon\partial_x v^\varepsilon)=\exp(X^\varepsilon)v^\varepsilon V_\varepsilon$$

which implies that v^ε satisfies

$$\mathscr{L}v^\varepsilon-v^\varepsilon(\partial_x X^\varepsilon)^2-\partial_x X^\varepsilon\partial_x v^\varepsilon=0,\qquad (t,x)\in[0,T]\times\mathbb{T}^d$$

since $\exp(X^\varepsilon)>0$ on $[0,T]\times\mathbb{T}^d$. Parabolic estimates imply that $X^\varepsilon\in C_T\mathscr{C}^\gamma$ with bounds uniform in $\varepsilon>0$ so that here the problematic term is $(\partial_x X^\varepsilon)^2$ for which this estimate does not guarantee existence.

Lemma 14. *Assume that*

$$\sigma^2=\int\frac{R(k)}{k^2}\frac{\mathrm{d}k}{(2\pi)^d}<+\infty.$$

Then if $\alpha=1$ we have

$$\lim_{\varepsilon\to 0}\mathbb{E}[(\partial_x X^\varepsilon)^2(t,x)]=\sigma^2$$

and if $\alpha<1$

$$\lim_{\varepsilon\to 0}\mathbb{E}[(\partial_x X^\varepsilon)^2(t,x)]=0.$$

Moreover if $d>2$ *we have*

$$\mathrm{Var}[\Delta_q(\partial_x X^{\varepsilon})^2(t,x)] \lesssim \varepsilon^{d-2}\varepsilon^{4-4\alpha}\sigma^2.$$

Proof. A computation similar to that leading to eq. (2) gives

$$\mathbb{E}[(\partial_x X^{\varepsilon})^2(t,x)] = \left[\frac{\varepsilon}{2\pi}\right]^d \varepsilon^{-2\alpha} \sum_{k\in\varepsilon\mathbb{Z}^d} (k/\varepsilon)^2 \left[\int_0^t e^{-(t-s)(k/\varepsilon)^2} \mathrm{d}s\right] R(k)$$

$$= \varepsilon^{2-2\alpha}\left[\frac{\varepsilon}{2\pi}\right]^d \sum_{k\in\varepsilon\mathbb{Z}^d} \frac{\left[1-e^{-t(k/\varepsilon)^2}\right]^2}{k^2} R(k)$$

which as $\varepsilon\to 0$ tends to

$$\lim_{\varepsilon\to 0} \mathbb{E}[(\partial_x X^{\varepsilon})^2(t,x)] = \varepsilon^{2-2\alpha}\int \frac{R(k)}{k^2}\frac{\mathrm{d}k}{(2\pi)^d} = \sigma^2.$$

Let us now study the variance of the LP blocks of $(\partial_x X^{\varepsilon})^2(t,x)$. Let us observe that

$$(\partial_x X^{\varepsilon})^2 = 2\,\partial_x X^{\varepsilon} \prec \partial_x X^{\varepsilon} + \partial_x X^{\varepsilon} \circ \partial_x X^{\varepsilon}$$

with $\partial_x X^{\varepsilon}\circ\partial_x X^{\varepsilon} = \sum_{|i-j|\leqslant 1} (\partial_x\Delta_i X^{\varepsilon})(\partial_x \Delta_j X^{\varepsilon})$. Calling $\partial_x\Delta_i X^{\varepsilon}(t,x) = Y_i$ we have

$$\mathrm{Var}[\Delta_q(\partial_x X^{\varepsilon}\circ\partial_x X^{\varepsilon})(t,x)] \lesssim \sum_{|i-j|\leqslant 1, i\gtrsim q}\ \sum_{|i'-j'|\leqslant 1, i'\gtrsim q} |\mathrm{Cov}(Y_i Y_j; Y_{i'} Y_{j'})|$$

$$\lesssim \sum_{i\sim i'\sim j\sim j'\gtrsim q} |\mathbb{E}[Y_i Y_{i'}]\mathbb{E}[Y_j Y_{j'}]|$$

Now as above if $i\sim i'$ we have

$$\mathbb{E}[Y_i Y_{i'}] = \left[\frac{\varepsilon}{2\pi}\right]^d \varepsilon^{-2\alpha}\sum_{k\in\varepsilon\mathbb{Z}^d} k^2 \rho_i(k/\varepsilon)\rho_{i'}(k/\varepsilon)\left[\int_0^t e^{-(t-s)(k/\varepsilon)^2}\mathrm{d}s\right] R(k)$$

$$= \left[\frac{\varepsilon}{2\pi}\right]^d \varepsilon^{-2\alpha}\sum_{k\in\varepsilon\mathbb{Z}^d} \rho_i(k/\varepsilon)\rho_{i'}(k/\varepsilon)\frac{\left[1-e^{-t(k/\varepsilon)^2}\right]^2}{(k/\varepsilon)^2} R(k)$$

so

$$\mathrm{Var}[\Delta_q(\partial_x X^{\varepsilon}\circ\partial_x X^{\varepsilon})(t,x)] \lesssim \sum_{i\gtrsim q}\left\{\left[\frac{\varepsilon}{2\pi}\right]^d \varepsilon^{-2\alpha}\sum_{k\in\varepsilon\mathbb{Z}^d}\rho_i(k/\varepsilon)\rho_{i'}(k/\varepsilon)\frac{\left[1-e^{-t(k/\varepsilon)^2}\right]^2}{(k/\varepsilon)^2}R(k)\right\}^2$$

$$\lesssim \left(\left[\frac{\varepsilon}{2\pi}\right]^d\varepsilon^{-2\alpha}\right)^2 \sum_{k\in\varepsilon\mathbb{Z}^d} \mathbb{I}_{|k|\gtrsim\varepsilon 2^q}\left(\frac{\left[1-e^{-t(k/\varepsilon)^2}\right]^2}{(k/\varepsilon)^2}R(k)\right)^2$$

$$\lesssim \varepsilon^{d+2-4\alpha}\left[\frac{\varepsilon}{2\pi}\right]^d \sum_{k\in\varepsilon\mathbb{Z}^d}\mathbb{I}_{|k|\gtrsim\varepsilon 2^q}\frac{R(k)}{k^2} \lesssim \varepsilon^{d-2}\varepsilon^{4-4\alpha}\sigma^2$$

Which goes to zero as $\varepsilon\to 0$ if $d>2$. □

This lemma shows that the interesting situation is when $\alpha=1$. Then provided $\sigma^2<+\infty$ and $d>2$ we have $(\partial_x X^{\varepsilon})^2\to\sigma^2$ in $C_T\mathscr{C}^{0-}$ (an additional argument is needed to provide the uniformity in time of the convergence but this can be done along the lines of the above computations). An easy consequence of this is that v^{ε} converges to the solution of the PDE

$$\mathscr{L}v = \sigma^2 v \tag{3}$$

and since $X^{\varepsilon}\to 0$ in $C_T\mathscr{C}^{\gamma}$ we finally obtain the convergence of $(u^{\varepsilon})_{\varepsilon>0}$ to the same v.

4.1 The 2d generalized parabolic Anderson model

The case $\alpha=1$ and $d=2$ remains open in the previous analysis. When $d=2$ we cannot expect σ^2 to be finite and moreover from the above computations we see that the variance of $(\partial_x X^{\varepsilon})^2$ remains finite and does not go to zero so the limiting object should satisfy a stochastic PDE. If we let $\sigma_{\varepsilon}^2(t) = \mathbb{E}[(\partial_x X^{\varepsilon})^2(t,x)]$ (which depends on time but which is easily shown independent of $x\in\mathbb{T}^2$) then we expect that solving the *renormalized* equation

$$\mathscr{L}\tilde{u}^{\varepsilon} = V_{\varepsilon}\tilde{u}^{\varepsilon} - \sigma_{\varepsilon}^2\tilde{u}^{\varepsilon}$$

should give rise in the limit to a well defined random field $\tilde{u}$ satisfying $\tilde{u} = e^X \tilde{v}$ where

$$\mathscr{L}\tilde{v} = \tilde{v}\,\eta + \partial_x X \partial_x \tilde{v}$$

and X is the limit of X^ε as $\varepsilon \to 0$ while η is the limit of $(\partial_x X^\varepsilon)^2 - \sigma_\varepsilon^2(t)$. The relation of u^ε with $\tilde{u}^\varepsilon$ is easily seen to be $\tilde{u}^\varepsilon(t,x) = e^{-\int_0^t \sigma_\varepsilon^2(s)\mathrm{d}s} u^\varepsilon(t,x)$. The renormalization procedure is equivalent then to a time–dependent rescaling of the solution to the initial problem.

We will study the renormalization and convergence problem for a more general equation of the form

$$\mathscr{L}u^\varepsilon = F(u^\varepsilon) V_\varepsilon \tag{4}$$

where $F: \mathbb{R} \to \mathbb{R}$ is a general, sufficiently smooth, non–linearity. One possible motivation is that, if z^ε solves the linear PDE and we set $u^\varepsilon = \varphi(z^\varepsilon)$ for some invertible $\varphi: \mathbb{R} \to \mathbb{R}$ such that $\varphi' > 0$ then

$$\mathscr{L}u^\varepsilon = \varphi'(z^\varepsilon)\mathscr{L}z^\varepsilon - \varphi''(z^\varepsilon)(\partial_x z^\varepsilon)^2 = \varphi'(z^\varepsilon) z^\varepsilon V_\varepsilon - \varphi''(z^\varepsilon)(\varphi'(z^\varepsilon))^{-2}(\partial_x u^\varepsilon)^2$$

and u^ε satisfies in turn the PDE

$$\mathscr{L}u^\varepsilon = F_1(u^\varepsilon) V_\varepsilon + F_2(u^\varepsilon)(\partial_x u^\varepsilon)^2$$

where $F_1(x) = \varphi'(\varphi^{-1}(x))\varphi^{-1}(x)$ and $F_2(x) = -\varphi''(\varphi^{-1}(x))(\varphi'(\varphi^{-1}(x)))^{-2}$. In the situation we are interested in the second term in the r.h.s. it is simpler to treat than the first term so, for the time being, we will drop it and we will concentrate on the equation (4) in $d = 2$ with $\alpha = 1$ and short ranged ($\beta = d$) potential V which we refer to as *generalized parabolic Anderson model* (GPAM).

Under these conditions V_ε converges to the white noise in space which we usually denote with ξ and our aim will be to set up a theory in which the non–linear operations involved in the definition of the dynamics of the GPAM are well understood, including the possibility of the renormalization which already appears in the linear case as hinted above.

While the reader should have in mind always a limit procedure from a well defined model like the ones we were considering so far, in the following we will mostly discuss the limit equation. The specific phenomena appearing when trying to track the oscillation of the term $F(u^\varepsilon)V_\varepsilon$ as $\varepsilon \to 0$ will be described by a *renormalized product* $F(u) \diamond \xi$ and so we write the GPAM as

$$\mathscr{L}u(t,x) = F(u(t,x)) \diamond \xi(x), \qquad u(0) = u_0. \tag{5}$$

In the linear case $F(u) = u$, the problem of the renormalization can be solved along the lines suggested above. Another possible line of attack comes from the theory of Gaussian spaces and in particular from Wick products, see for example [Hu02]. However, the definition of the Wick product relies on the concrete chaos expansion of its factors, and since nonlinear functions change the chaos expansion in a complicated way, there is little hope of directly extending the Wick product approach to the nonlinear case and moreover the kind of solution which can be obtained using these non–local (in the probability space) objects can deliver solutions which are not physically acceptable.

Eq. (5) is structurally very similar to the stochastic differential equation

$$\partial_t v(t) = F(v(t)) \partial_t B^H(t), \qquad v(0) = v_0, \tag{6}$$

where B^H denotes a fractional Brownian motion with Hurst index $H \in (0,1)$. There are many ways to solve (6) in the Brownian case. Since we are interested in a way that might extend to (5), we should exclude all approaches based on information and filtrations; in particular, any approach that works for $H \neq 1/2$ might seem promising. Lyons' theory of rough paths [Lyo98] equips us exactly with the techniques we need to solve (6) for general H. More precisely, if for $H > 1/3$ we are given $\int_0^\cdot B_s^H \mathrm{d}B_s^H$, then we can use controlled rough path integral [Gub04] to make sense of $\int_0^\cdot f_s \mathrm{d}B_s^H$ for any f which "looks like" B^H. The product $f\partial_t B^H$ can then again be defined by differentiation. So the main ingredients required for controlled rough paths are the integral $\int_0^\cdot B_s^H \mathrm{d}B_s^H$ for the reference path B^H, and the fact that we can describe paths which look like B^H. It is worthwhile to note that while we need probability theory to construct $\int_0^\cdot B_s^H \mathrm{d}B_s^H$, the construction of $\int_0^\cdot f_s \mathrm{d}B_s^H$ is achieved using pathwise arguments and it is given as a continuous map of f and $\left(B^H, \int_0^\cdot B_s^H \mathrm{d}B_s^H\right)$. As a consequence, the solution to the SDE (6) depends pathwise continuously on $\left(B^H, \int_0^\cdot B_s^H \mathrm{d}B^H\right)$.

By the structural similarity of (5) and (6), we might hope to extend the rough path approach to (5). The equivalent of B^H is given by the solution ϑ to $\mathscr{L}\vartheta=\xi$, $\vartheta(0)=0$, so that we should assume the renormalized product $\vartheta\diamond\xi$ as given. Then we might hope to define $f\diamond\xi$ for all f that "look like ϑ", however this is to be interpreted. Of course, rough paths can only be applied to functions of a one-dimensional index variable, while for (5) the problem lies in the irregularity of ξ in the spatial variable $x\in\mathbb{T}^2$.

In the following we combine the ideas from controlled rough paths with Bony's paraproduct, a tool from functional analysis that allows us to extend them to functions of a multidimensional parameter. Using the paraproduct, we are able to make precise in a simple way what we mean by "distributions looking like a reference distribution". We can then define products of suitable distributions and solve (5) as well as many other interesting singular SPDEs.

4.2 More singular problems

Keeping the homogeneisation problem as leit–motiv for these lectures we could consider also space–time varying environments $V_\varepsilon(t,x)=\varepsilon^{-\alpha}V(t/\varepsilon^2,x/\varepsilon)$. The scaling of the temporal variable is chosen so that it is compatible with the diffusive scaling from a microscopic description where $V(t,x)$ has typical variation in space and time in scales of order 1. Assume that $d=1$, then when the random field V is Gaussian, zero mean and with short–range space–time correlations the natural choice for the magnitude of the macroscopic fluctuations is $\alpha=3/2$. In this case V_ε converges as $\varepsilon\to0$ to a space–time white noise ξ. Understanding the limit dynamics of u^ε, solution to the linear equation $\mathscr{L}u^\varepsilon=V_\varepsilon u^\varepsilon$ as $\varepsilon\to0$ represents now a more difficult problem than in the time independent situation. A Gaussian computation shows that the random field X^ε, solution to $\mathscr{L}X^\varepsilon=V_\varepsilon$ (e.g. with zero initial condition) stay bounded in $C_T\mathscr{C}^{1/2-}$ as $\varepsilon\to0$. Since $\mathscr{L}$ is a second order operator (if we use an appropriate parabolic weighting of the time and space regularities) then ξ is expected to live in a space of distributions of regularity $-3/2-$. This is to be compared with the $-1-$ of the space white noise which had to be dealt with in the gpam. Renormalization effects are then expected to be stronger in this setting and the limiting object, which we denote with w should satisfy a (suitably renormalized) linear stochastic heat equation with multiplicative noise (she)

$$\mathscr{L}w(t,x)=w(t,x)\diamond\xi(t,x),\qquad w(0)=w_0. \tag{7}$$

As hinted by the computations in the more regular case, it is useful to consider the change of variables $w=e^h$ which is called Cole–Hopf transformation. Here $h\colon[0,\infty)\times\mathbb{T}\to\mathbb{R}$ is a new unknown which satisfy now the Kardar–Parisi–Zhang (kpz) equation:

$$\mathscr{L}h(t,x)=(\partial_x h(t,x))^{\diamond2}+\xi(t,x),\qquad h(0)=h_0 \tag{8}$$

where the difficulty comes now from the squaring of the derivative but which has the nice feature to be additively perturbed by the space–time white noise, feature which simplifies many considerations. Another relevant model in applications is obtained by taking the space derivative of kpz and letting $u(t,x)=\partial_x h(t,x)$ in order to obtain the stochastic conservation law

$$\mathscr{L}u(t,x)=\partial_x(u(t,x))^{\diamond2}+\partial_x\xi(t,x),\qquad u(0)=u_0, \tag{9}$$

which we will refer to as the stochastic Burgers equation sbe. In all these cases, $\diamond$ denotes a suitably renormalized product.

The kpz equation was derived by Kardar–Parisi–Zhang in 1986 as a universal model for the random growth of an interface [KPZ86]. For a long time it could not be solved due to the fact that there was no way to make sense of the nonlinearity $(\partial_x h)^{\diamond2}$ in (8). The only way to make sense of (8) was to apply the Cole-Hopf transform [BG97]: start by solving the stochastic heat equation (7) (which is accessible to Itô integration) and then set $h=\log w$. But there was no was no interpretation for the nonlinearity $(\partial_x h)^{\diamond2}$, and in particular there was no intrinsic definition of what it means to solve (8). Finally in 2012, when Hairer [Hai13] used rough paths to solve give a meaning to the equation and obtain directly solutions at the kpz level. Here we recover his solution in the paracontrolled setting. Application of the techniques used by Hairer to solve the kpz problem to a more general homogeneisation problem with general ergodic potentials (not necessarily Gaussian) have been studied in [HPP13].

4.3 Hairer's regularity structures

In [Hai14], Hairer introduces a theory of regularity structures which can also be considered a generalization of the theory of controlled rough paths to functions of a multidimensional index variable. Hairer fundamentally rethinks the notion of regularity. Usually a function is called smooth if it can be approximated around every point by a polynomial of a given degree (the Taylor polynomial). Naturally, the solution to an SPDE driven by –say– Gaussian space-time white noise is not smooth. So in Hairer's theory, a function is called smooth if locally it can be approximated by the noise (and higher order terms constructed from the noise). This induces a natural topology in which the solutions to semilinear SPDEs depend continuously on the driving signal. Hairer's approach is very general and allows to handle more general problems than the ones that can be currently treated with the paracontrolled techniques.

Compared with the theory of regularity structures, we can hardly call paracontrolled distributions a *theory*. It is just a set of tools which allows us to understand better the multiplication of distributions and thus to solve some SPDEs. We use classical notions of regularity and only observe that there exist interesting settings (i.e. beyond the Young integral conditions) where the pointwise multiplication extends to a bounded operator in a suitable topology. If there is a merit in this approach, then its relative simplicity, the fact that it seems to be very adaptable so that it can be easily modified to treat problems with a different structure, and that we make the connection between rough paths and harmonic analysis.

5 The paracontrolled PAM

As we have tried to motivate in the previous sections we are looking for a theory for the PAM which describes the possible limits of the equation

$$\mathscr{L}u = F(u)\eta \tag{10}$$

driven by sufficiently regular η but as η is converging to the space white noise ξ. From this point of view we are looking at a–priori estimates on the solutions u to the above equation which depends only on distributional norms of η so in the following we will assume that we have at hand only a uniform control of η in $C_T\mathscr{C}^{\gamma-2}$ for some $\gamma > 0$, for the application to the 2d space white noise we could take $\gamma = 1-$ but we will not use this specific information in order to probe the range of applicability of our approach and we will assume only that the exponent γ is such that $3\gamma - 2 > 0$, this includes the case $\gamma = 1-$.

Assume for a moment that we are in the simpler situation $\gamma > 1$ and $u_0 \in \mathscr{C}^\gamma$. Trying to solve eq. (10) via Picard iterations $(u^n)_{n \geqslant 0}$ starting from $u^0 \equiv u_0$. Since F preserves $C\mathscr{C}^\gamma$-regularity (which can be seen by identifying $C\mathscr{C}^\gamma$ with the classical space of bounded Hölder-continuous functions of space), the product $F(u^0(t))\eta$ is well defined as an element of $\mathscr{C}^{\gamma-2}$ for all $t \geqslant 0$ since $2\gamma - 2 > 0$ and we are in condition to apply Corollary 16 below on the product of elements in Holder–Besov spaces. Now by Lemma 12, the Laplacian gains two degrees of regularity so that the solution u^1 to $\mathscr{L}u^1 = F(u^0)\eta$, $u^1(0) = u_0$, is in $C\mathscr{C}^\gamma$. From here we obtain a contraction on $C_T\mathscr{C}^\gamma$ for some small $T > 0$ whose value does not depend on u_0, which gives us global in time existence and uniqueness of solutions. Note that if we are in one dimension the space–white noise has regularity $\mathscr{C}^{-1/2-}$ (see e.g. Exercise 4) so taking $\gamma = 3/2-$ we have determined that the 1d PAM can be solved globally in time with standard techniques.

When the condition $2\gamma - 2 > 0$ is not satisfied we still have that if $\eta \in C_T\mathscr{C}^{\gamma-2}$ then $u \in \mathscr{L}^\gamma = C_T\mathscr{C}^{\gamma-2} \cap C_T^{\gamma/2}L^\infty$ by the standard parabolic estimates of Lemma 12. However with the regularities at hand we cannot guarantee anymore the continuity of the operator $(u, \eta) \mapsto F(u)\eta$ using Corollary 16. Moreover, as already seen in the simpler homogeneisation problems above this is not a technical difficulty but a real issue of the regime $\gamma \leqslant 1$. We expect that controlling the model in this regime can be quite tricky since limits exists when $\eta \to 0$ but the limit solution still feels residual order one effects from the vanishing driving signal η. This situation cannot be improved from the point of view of the standard analytic considerations. What is needed is a finer control of the solution u which allows to analyse more in detail the possible resonances between fluctuations of u and that of η.

Before going on however we will revise the problem of multiplication of distributions in the scale of Hölder–Besov spaces introducing the basic tool of our general analysis: Bony's paraproduct.

5.1 The paraproduct and the resonant term

Paraproducts are bilinear operations introduced by Bony [Bon81] in order to linearize a class of nonlinear hyperbolic PDE problems in order to analyse the regularity of their solutions. In terms of Littlewood–Paley blocks, a general product $f g$ of two distributions can be (at least formally) decomposed as

$$f g = \sum_{j \geqslant -1} \sum_{i \geqslant -1} \Delta_i f \, \Delta_j g = f \prec g + f \succ g + f \circ g.$$

Here $f \prec g$ is the part of the double sum with $i < j - 1$, $f \succ g$ is the part with $i > j + 1$, and $f \circ g$ is the "diagonal" part, where $|i - j| \leqslant 1$. More precisely, we define

$$f \prec g = g \succ f = \sum_{j \geqslant -1} \sum_{i = -1}^{j-2} \Delta_i f \, \Delta_j g \qquad \text{and} \qquad f \circ g = \sum_{|i - j| \leqslant 1} \Delta_i f \, \Delta_j g.$$

Of course, the decomposition depends on the dyadic partition of unity used to define the blocks Δ_j, and also on the particular choice of the pairs (i, j) in the diagonal part. The choice of taking all (i, j) with $|i - j| \leqslant 1$ into the diagonal part corresponds to the fact that the partition of unity can be chosen such that $\operatorname{supp} \mathscr{F}(\Delta_i f \Delta_j g) \subseteq 2^j \mathscr{A}$ if $i < j - 1$. If $|i - j| \leqslant 1$ the only apriori information on the spectral support of the various term in the double sum is $\operatorname{supp} \mathscr{F}(\Delta_i f \Delta_j g) \subseteq 2^j \mathscr{B}$, that is they are supported in balls and in particular they can have non–zero contributions to very low wave vectors. We call $f \prec g$ and $f \succ g$ *paraproducts*, and $f \circ g$ the *resonant* term.

Bony's crucial observation is that $f \prec g$ (and thus $f \succ g$) is always a well-defined distribution. Heuristically, $f \prec g$ behaves at large frequencies like g (and thus retains the same regularity), and f provides only a frequency modulation of g. The only difficulty in constructing $f g$ for arbitrary distributions lies in handling the diagonal term $f \circ g$. The basic result about these bilinear operations is given by the following estimates.

Theorem 15. *(Paraproduct estimates) For any $\beta \in \mathbb{R}$ and $f, g \in \mathscr{D}'$ we have*

$$\| f \prec g \|_{\beta} \lesssim_{\beta} \| f \|_{L^{\infty}} \| g \|_{\beta}, \tag{11}$$

and for $\alpha < 0$ furthermore

$$\| f \prec g \|_{\alpha + \beta} \lesssim_{\alpha, \beta} \| f \|_{\alpha} \| g \|_{\beta}. \tag{12}$$

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

$$\| f \circ g \|_{\alpha + \beta} \lesssim_{\alpha, \beta} \| f \|_{\alpha} \| g \|_{\beta}. \tag{13}$$

Proof. There exists an annulus $\mathscr{A}$ such that $S_{j-1} f \, \Delta_j g$ has Fourier transform supported in $2^j \mathscr{A}$, and for $f \in L^{\infty}$ we have

$$\| S_{j-1} f \, \Delta_j g \|_{L^{\infty}} \leqslant \| S_{j-1} f \|_{L^{\infty}} \| \Delta_j g \|_{L^{\infty}} \lesssim \| f \|_{L^{\infty}} 2^{-j\beta} \| g \|_{\beta}.$$

By Lemma 9, we thus obtain (11). The proof of of (12) and (13) works in the same way, where for estimating $f \circ g$ we need $\alpha + \beta > 0$ because the terms of the series are supported in a ball $2^j \mathscr{B}$ and not in an annulus. □

A simple corollary is then the following :

Corollary 16. *Let $f \in \mathscr{C}^{\alpha}$ and $g \in \mathscr{C}^{\beta}$ with $\alpha + \beta > 0$, then the product $(f, g) \mapsto f g$ is a continuous bounded bilinear map $\mathscr{C}^{\alpha} \times \mathscr{C}^{\beta} \mapsto \mathscr{C}^{\alpha \wedge \beta}$.*

The ill-posedness of $f \circ g$ for $\alpha + \beta \leqslant 0$ can be interpreted as a resonance effect since $f \circ g$ contains exactly those part of the double series where f and g are in the same frequency range. The paraproduct $f \prec g$ can be interpreted as frequency modulation of g, which should become more clear in the following example.

Example 17. In Figure 1 we see a slowly oscillating positive function u, while Figure 2 depicts a fast sine curve v. The product uv, which here equals the paraproduct $u \prec v$ since u has no rapidly oscillating components, is shown in Figure 3. We see that the local fluctuations of uv are due to v, and that uv is essentially oscillating with the same speed as v.

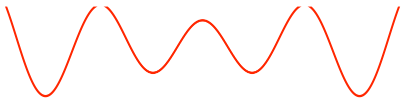

Figure 1.

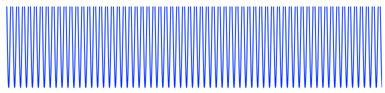

Figure 2.

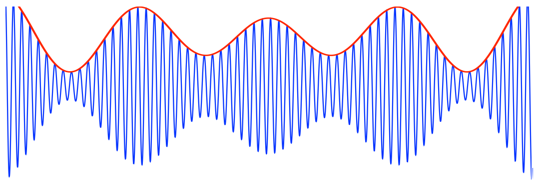

Figure 3.

As a corollary of Theorem 15, the product fg of $f \in \mathscr{C}^\alpha$ and $g \in \mathscr{C}^\beta$ is well defined as soon as $\alpha + \beta > 0$, and it belongs to $\mathscr{C}^\gamma$ where $\gamma = \min\{\alpha, \beta, \alpha + \beta\}$. Using smooth approximations, we see that while $f \prec g$, $f \succ g$, and $f \circ g$ depend on the specific dyadic partition of unity, the product fg does not.

Example 18. Let B^H be a fractional Brownian bridge on $\mathbb{T}$ (or simply a fractional Brownian motion on $[0, \pi]$, reflected on $[\pi, 2\pi]$) and assume that $H > 1/2$. We have $B^H \in \mathscr{C}^{H-}_{\mathrm{loc}}$, and in particular $B^H \partial_t B^H$ is well-defined. This can be used to solve SDEs driven by B^H in a pathwise sense. In that way we recover exactly the Young integral, and in fact the condition $\alpha + \beta > 0$ corresponds to the Young condition $\gamma + \delta > 1$: if $f \in \mathscr{C}^\gamma$ and $g \in \mathscr{C}^\delta$ and we want to construct $\int f \mathrm{d}g$, then this is equivalent to constructing $f \partial_t g$, and since $\partial_t g \in \mathscr{C}^{\delta - 1}$ we recover the Bony condition $\gamma + (\delta - 1) > 0$.

The condition $\alpha + \beta > 0$ is essentially sharp, at least at this level of generality, see [You36] for counterexamples. It excludes of course the problem of Brownian case: if B is a Brownian motion, then almost surely $B \in \mathscr{C}^\alpha_{\mathrm{loc}}$ for all $\alpha < 1/2$, so that $\partial_t B \in \mathscr{C}^{\alpha - 1}_{\mathrm{loc}}$ and thus $B \circ \partial_t B$ fails to be well defined. See also [LCL07], Proposition 1.29 for an instructive example which shows that this is not a shortcoming of our description of regularity, but that it is indeed impossible to define the product $B \partial_t B$ as a continuous bilinear operation on distribution spaces.

Other counterexamples are given by our discussion of the homogeneisation above. More simply one can consider the following situation.

Example 19. Consider the sequence of functions $f_n \colon \mathbb{T} \to \mathbb{C}$ given by $f_n(t) = e^{in^2 t}/n$. Then it is easy to show that $\|f_n\|_\gamma \to 0$ for all $\gamma < 1/2$. However let $g_n(t) = \mathrm{Re}\, f_n(t)\, \mathrm{Im}\, \partial_t f_n(t) = \left(1 - e^{i2n^2 t} - e^{-i2n^2 t}\right)/(2i)$. Then $g_n \to 1/2$ in $\mathscr{C}^{0-}$ which shows that the map $f \mapsto (\mathrm{Re}\, f)\, (\partial_t \mathrm{Im}\, f)$ cannot be continuous in $\mathscr{C}^\gamma$ if $\gamma < 1/2$. Pictorially the situation is resumed in Fig. 4 where we sketched the three dimensional curve given by $t \mapsto \left(\mathrm{Re}\, f_n(t), \mathrm{Im}\, f_n(t), \int_0^t g_n(s) \mathrm{d}s\right)$ for various values of n and in the limit.

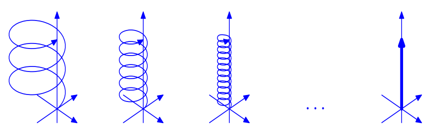

Figure 4. Resonances give macroscopic effects

5.2 Commutator estimates and paralinearization

The product $F(u)\eta$ appearing in the r.h.s. of PAM can be decomposed via the paraproduct $\prec$ as a sum of three terms

$$F(u)\eta = F(u) \prec \eta + F(u) \circ \eta + F(u) \succ \eta.$$

The first and the last of these terms are continuous in any topology we choose for $F(u)$ and η. The *resonant term* $F(u) \circ \eta$ however is problematic. It gathers the products of the oscillations of $F(u)$ and η on comparable dyadic scales and these products can contribute to the results at any other higher scale in such a way that microscopic oscillations can build up to a macroscopic effect which does not disappear in the limit (as we already have seen). If the function F is smooth enough these resonances corresponds to the resonances of u and η.

We use the paraproduct decomposition to write the right hand side of (10) as a sum of the three terms

$$\underbrace{F(u) \prec \eta}_{\gamma-2} + \underbrace{F(u) \circ \eta}_{2\gamma-2} + \underbrace{F(u) \succ \eta}_{2\gamma-2} \tag{14}$$

(where the quantity indicated by the underbrace corresponds to the expected regularity of each term). Unless $2\gamma - 2 > 0$, the resonant term $F(u) \circ \eta$ cannot be controlled using only the $C\mathscr{C}^{\gamma}$–norm of u and the $C\mathscr{C}^{\gamma-2}$–norm of η.

If F is at least C^2, we can use a *paralinearization* result (stated precisely in Lemma 23 below) to rewrite this term as

$$F(u) \circ \eta = F'(u)\,(u \circ \xi) + \Pi_F(u, \xi), \tag{15}$$

where the remainder $\Pi_F(u, \xi) \in \mathscr{C}^{3\gamma-2}$ provided $3\gamma - 2 > 0$. The difficulty is now localized in the linearized resonant product $u \circ \eta$. In order to control this term, we would like to exploit the fact that the function u is not a generic element of $C\mathscr{C}^{\gamma}$ but that it has a specific structure, since $\mathscr{L}u$ has to match the paraproduct decomposition given in (14) where the less regular term is expected to be $F(u) \prec \eta \in C\mathscr{C}^{\gamma-2}$.

In order to do so, we postulate that the solution u is given by the following *paracontrolled ansatz:*

$$u = u^{\vartheta} \prec \vartheta + u^{\sharp},$$

for functions u^{ϑ}, ϑ, $u^{\sharp}$ such that u^{ϑ}, $\vartheta \in \mathscr{C}^{\gamma}$ and the remainder $u^{\sharp} \in C\mathscr{C}^{2\gamma}$. This decomposition allows for a finer analysis of the resonant term $u \circ \xi$: indeed, we have

$$u \circ \eta = (u^{\vartheta} \prec \vartheta) \circ \eta + u^{\sharp} \circ \xi = u^{\vartheta}\,(\vartheta \circ \eta) + C(u^{\vartheta}, \vartheta, \eta) + u^{\sharp} \circ \eta, \tag{16}$$

where the *commutator* is defined by $C(u^{\vartheta}, \vartheta, \eta) = (u^{\vartheta} \prec \vartheta) \circ \eta - u^{\vartheta}\,(\vartheta \circ \eta)$. Observe now that the term $u^{\sharp} \circ \eta$ does not pose any further problem, as it is bounded in $C\mathscr{C}^{3\gamma-2}$. The key point is now that the commutator is a bounded multilinear function of its arguments as long as the sum of their regularities is strictly positive, see Lemma 20 below. By assumption, we have $3\gamma - 2 > 0$, and therefore $C(u^{\vartheta}, \vartheta, \eta) \in C\mathscr{C}^{3\gamma-2}$.

The only problematic term which remains to be handled is thus the bilinear functional of the noise given by $\vartheta \circ \eta$. Here we need to make the assumption that $\vartheta \circ \eta \in C\mathscr{C}^{2\gamma-2}$ in order for the product $u^{\vartheta}\,(\vartheta \circ \eta)$ to be well defined. That assumption is not guaranteed by the analytical estimates at hand, and it has to be added as a further requirement to our construction.

Granting this last step, we have obtained that the right hand side of equation (10) is well defined and a continuous function of $(u, u^{\vartheta}, u^{\sharp}, \vartheta, \eta, \vartheta \circ \eta)$.

It remains to check that the paracontrolled ansatz is coherent with the equation satisfied by solutions to PAM. The ansatz and the Leibniz rule for the paraproduct imply that (10) can be rewritten as

$$\mathscr{L}u = \mathscr{L}(u^{\vartheta} \prec \vartheta + u^{\sharp}) = u^{\vartheta} \prec \mathscr{L}\vartheta + [\mathscr{L}, u^{\vartheta} \prec]\vartheta + \mathscr{L}u^{\sharp} = F(u) \prec \eta + F(u) \circ \eta + F(u) \succ \eta.$$

If we choose ϑ such that $\mathscr{L}\vartheta = \eta$ and we set $u^\vartheta = F(u)$, then we can use (15) and (16) to obtain the following equation for the remainder $u^\sharp$:

$$\begin{aligned} \mathscr{L} u^\sharp \; = & F'(u)\, F(u)\, (\vartheta \circ \eta) + F(u) \succ \eta - [\mathscr{L}, F(u) \prec]\vartheta \\ & + F'(u)\, C(F(u), \vartheta, \eta) + F'(u)\, (u^\sharp \circ \eta) + \Pi_F(u, \eta). \end{aligned} \tag{17}$$

Lemma 25 below ensure that $\mathscr{L}^{-1}[\mathscr{L}, F(u) \prec]\vartheta \in C\mathscr{C}^{2\gamma}$ if $u \in \mathscr{L}^\gamma$ and we have already seen that all the other terms in the r.h.s. are in $C\mathscr{C}^{2\gamma-2}$ so this equation implies estimates for $u^\sharp \in C\mathscr{C}^{2\gamma}$ in terms of the r.h.s. of eq. (17) . Together with the equation $u = F(u) \prec \vartheta + u^\sharp$, this equation gives an equivalent description of the solution and allows us to obtain an a priori estimate on u and $u^\sharp$ in terms of $(u_0, \|\eta\|_{\gamma-2}, \|\vartheta \circ \eta\|_{2\gamma-2})$. It is now easy to show that if $F \in C_b^3$, then u depends continuously on the data $(u_0, \eta, \vartheta \circ \eta)$, so that we have a robust strategy to pass to the limit in (4) and make sense of the solution to (10) also for irregular $\eta \in C\mathscr{C}^{\gamma-2}$ as long as $\gamma > 2/3$.

In the rest of this section we will prove the results (paralinearization and various key commutators) which we used in this discussion before going on to gather the consequences of this analysis in the next section. When the time dependence does not play any role we state the results for distributions depending only on the space variable, the extension to time varying functions will not add further difficulty.

Lemma 20. *Assume that $\alpha, \beta, \gamma \in \mathbb{R}$ are such that $\alpha + \beta + \gamma > 0$ and $\beta + \gamma < 0$. Then for $f, g, h \in C^\infty$ the trilinear operator*

$$C(f, g, h) = ((f \prec g) \circ h) - f\,(g \circ h)$$

allows for the bound

$$\|C(f, g, h)\|_{\beta+\gamma} \lesssim \|f\|_\alpha \|g\|_\beta \|h\|_\gamma, \tag{18}$$

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

Proof. By definition

$$\begin{aligned} C(f, g, h) &= \sum_{i,j,k,\ell} \Delta_i(\Delta_j f \Delta_k g) \Delta_\ell h (1_{j<k-1} 1_{|i-\ell| \leqslant 1} - 1_{|k-\ell| \leqslant 1}) \\ &= \sum_{i,j,k,\ell} \Delta_i(\Delta_j f \Delta_k g) \Delta_\ell h (1_{j<k-1} 1_{|i-\ell| \leqslant 1} 1_{|k-\ell| \leqslant N} - 1_{|k-\ell| \leqslant 1}), \end{aligned}$$

where we used that $S_{k-1} f \Delta_k g$ has support in an annulus $2^k \mathscr{A}$, so that $\Delta_i(S_{k-1} f \Delta_k g) \neq 0$ only if $|i - k| \leqslant N - 1$ for some fixed $N \in \mathbb{N}$, which in combination with $|i - \ell| \leqslant 1$ yields $|k - \ell| \leqslant N$. Now for fixed k, the term $\sum_\ell 1_{2 \leqslant |k-\ell| \leqslant N} \Delta_k g \Delta_\ell h$ is spectrally supported in an annulus $2^k \mathscr{A}$, so that $\sum_{k,\ell} 1_{2 \leqslant |k-\ell| \leqslant N} \Delta_k g \Delta_\ell h \in \mathscr{C}^{\beta+\gamma}$ and we may add and subtract $f \sum_{k,\ell} 1_{2 \leqslant |k-\ell| \leqslant N} \Delta_k g \Delta_\ell h$ to $C(f, g, h)$ while maintaining the bound (18). It remains to treat

$$\begin{aligned} & \sum_{i,j,k,\ell} \Delta_i(\Delta_j f \Delta_k g) \Delta_\ell h 1_{|k-\ell| \leqslant N} (1_{j<k-1} 1_{|i-\ell| \leqslant 1} - 1) \\ = & - \sum_{i,j,k,\ell} \Delta_i(\Delta_j f \Delta_k g) \Delta_\ell h 1_{|k-\ell| \leqslant N} (1_{j \geqslant k-1} + 1_{j<k-1} 1_{|i-\ell| > 1}). \end{aligned} \tag{19}$$

We estimate both terms on the right hand side separately. For $m \geqslant -1$ we have

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

using $\beta+\gamma<0$. It remains to estimate the second term in (19). For $|i-\ell|>1$ and $i\sim k\sim\ell$, any term of the form $\Delta_i()\Delta_\ell()$ is spectrally supported in an annulus $2^\ell\mathscr{A}$, and therefore

$$\begin{aligned}\left\|\Delta_m\left(\sum_{i,j,k,\ell}\Delta_i(\Delta_j f\Delta_k g)\Delta_\ell h 1_{|k-\ell|\leqslant N}1_{j<k-1}1_{|i-\ell|>1}\right)\right\|_{L^\infty}\\ \lesssim\sum_{i,j,k,\ell}\mathbb{I}_{j<k-1}\mathbb{I}_{i\sim k\sim\ell\sim m}\|\Delta_i(\Delta_j f\Delta_k g)\Delta_\ell h\|_{L^\infty}\\ \lesssim\sum_{j\lesssim m}2^{-j\alpha}\|f\|_\alpha 2^{-m\beta}\|g\|_\beta 2^{-m\gamma}\|h\|_\gamma\lesssim 2^{-m(\beta+\gamma)}\|f\|_\alpha\|g\|_\beta\|h\|_\gamma.\end{aligned}$$

□

Remark 21. The restriction $\beta+\gamma<0$ is not problematic. If $\beta+\gamma>0$, then $(f\prec g)\circ h$ can be treated with the usual paraproduct estimates, without the need of introducing the commutator. If $\beta+\gamma=0$, then we can apply the commutator estimate with $\gamma'<\gamma$ sufficiently close to γ such that $\alpha+\beta+\gamma'>0$.

Our next result is a simple paralinearization lemma for non-linear operators.

Lemma 22. (see also [BCD11], Theorem 2.92) *Let $\alpha\in(0,1)$, $\beta\in(0,\alpha]$, and let $F\in C_b^{1+\beta/\alpha}$. There exists a locally bounded map $R_F\colon\mathscr{C}^\alpha\to\mathscr{C}^{\alpha+\beta}$ such that*

$$F(f)=F'(f)\prec f+R_F(f) \tag{20}$$

for all $f\in\mathscr{C}^\alpha$. More precisely, we have

$$\|R_F(f)\|_{\alpha+\beta}\lesssim\|F\|_{C_b^{1+\beta/\alpha}}(1+\|f\|_\alpha^{1+\beta/\alpha}).$$

If $F\in C_b^{2+\beta/\alpha}$, then R_F is locally Lipschitz continuous:

$$\|R_F(f)-R_F(g)\|_{\alpha+\beta}\lesssim\|F\|_{C_b^{2+\beta/\alpha}}(1+\|f\|_\alpha+\|g\|_\alpha)^{1+\beta/\alpha}\|f-g\|_\alpha.$$

Proof. The difference $F(f)-F'(f)\prec f$ is given by

$$R_F(f)=F(f)-F'(f)\prec f=\sum_{i\geqslant-1}[\Delta_i F(f)-S_{i-1}F'(f)\,\Delta_i f]=\sum_{i\geqslant-1}u_i,$$

and every u_i is spectrally supported in a ball $2^i\mathscr{B}$. For $i<1$, we simply estimate $\|u_i\|_{L^\infty}\lesssim\|F\|_{C_b^1}(1+\|f\|_\alpha)$. For $i\geqslant1$ we use the fact that f is a bounded function to write the Littlewood-Paley projections as convolutions and obtain

$$\begin{aligned}u_i(x)&=\int K_i(x-y)\,K_{<i-1}(x-z)\,[F(f(y))-F'(f(z))\,f(y)]\,\mathrm{d}y\,\mathrm{d}z\\&=\int K_i(x-y)\,K_{<i-1}(x-z)\,[F(f(y))-F(f(z))-F'(f(z))\,(f(y)-f(z))]\,\mathrm{d}y\,\mathrm{d}z,\end{aligned}$$

where $K_i=\mathscr{F}^{-1}\rho_i$, $K_{<i-1}=\sum_{j<i-1}K_j$, and where we used that $\int K_i(y)\,\mathrm{d}y=\rho_i(0)=0$ for $i\geqslant0$ and $\int K_{<i-1}(z)\,\mathrm{d}z=1$ for $i\geqslant1$. Now we can apply a first order Taylor expansion to F and use the β/α–Hölder continuity of F' in combination with the α–Hölder continuity of f, to deduce

$$\begin{aligned}|u_i(x)|&\lesssim\|F\|_{C_b^{1+\beta/\alpha}}\|f\|_\alpha^{1+\beta/\alpha}\int|K_i(x-y)\,K_{<i-1}(x-z)|\times|z-y|^{\alpha+\beta}\,\mathrm{d}y\,\mathrm{d}z\\&\lesssim\|F\|_{C_b^{1+\beta/\alpha}}\|f\|_\alpha^{1+\beta/\alpha}2^{-i(\alpha+\beta)}.\end{aligned}$$

Therefore, the estimate for $R_F(f)$ follows from Lemma 9. The estimate for $R_F(f)-R_F(g)$ is shown in the same way. □

Let g be a distribution belonging to $\mathscr{C}^\beta$ for some $\beta<0$. Then the map $f\mapsto f\circ g$ behaves, modulo smoother correction terms, like a derivative operator:

Lemma 23. *Let $\alpha \in (0,1)$, $\beta \in (0,\alpha]$, $\gamma \in \mathbb{R}$ be such that $\alpha+\beta+\gamma>0$ but $\alpha+\gamma<0$. Let $F \in C_b^{1+\beta/\alpha}$. Then there exists a locally bounded map $\Pi_F : \mathscr{C}^\alpha \times \mathscr{C}^\gamma \to \mathscr{C}^{\alpha+\beta+\gamma}$ such that*

$$F(f) \circ g = F'(f)\,(f \circ g) + \Pi_F(f, g) \tag{21}$$

for all $f \in \mathscr{C}^\alpha$ and all smooth g. More precisely, we have

$$\|\Pi_F(f, g)\|_{\alpha+\beta+\gamma} \lesssim \|F\|_{C_b^{1+\beta/\alpha}} (1 + \|f\|_\alpha^{1+\beta/\alpha}) \|g\|_\gamma.$$

If $F \in C_b^{2+\beta/\alpha}$, then Π_F is locally Lipschitz continuous:

$$\|\Pi_F(f, g) - \Pi_F(u, v)\|_{\alpha+\beta+\gamma} \lesssim \|F\|_{C_b^{2+\beta/\alpha}} (1 + (\|f\|_\alpha + \|u\|_\alpha)^{1+\beta/\alpha} + \|v\|_\gamma)\,(\|f-u\|_\alpha + \|g-v\|_\gamma).$$

Proof. Use the paralinearization and commutator lemmas above to deduce that

$$\begin{aligned}\Pi(f, g) &= F(f) \circ g - F'(f)\,(f \circ g) = R_F(f) \circ g + (F'(f) \prec f) \circ g - F'(f)\,(f \circ g)\\ &= R_F(f) \circ g + C(F'(f), f, g),\end{aligned}$$

so that the claimed bounds easily follow from Lemma 20 and Lemma 22. □

Besides this sort of chain rule, we also have a Leibniz rule for $f \mapsto f \circ g$:

Lemma 24. *Let $\alpha \in (0,1)$ and $\gamma < 0$ be such that $2\alpha + \gamma > 0$ but $\alpha + \gamma < 0$ Then there exists a bounded trilinear operator $\Pi_\times : \mathscr{C}^\alpha \times \mathscr{C}^\alpha \times \mathscr{C}^\gamma \to \mathscr{C}^{2\alpha+\gamma}$, such that*

$$(f u) \circ g = f\,(u \circ g) + u\,(f \circ g) + \Pi_\times(f, u, g)$$

for all $f, u \in \mathscr{C}^\alpha(\mathbb{R})$ and all smooth g.

Proof. It suffices to note that $f u = f \prec u + f \succ u + f \circ u$, which leads to

$$\Pi_\times(f, u, g) = (f u) \circ g - f\,(u \circ g) + u\,(f \circ g) = C(f, u, g) + C(u, f, g) + (f \circ u) \circ g. \qquad \square$$

Lemma 25. *Let $f \in \mathscr{L}^\beta$, $G \in C\mathscr{C}^\alpha$ such that $\mathscr{L}G \in C\mathscr{C}^{\alpha-2}$. There exists $H = H(f, G)$ such that $\mathscr{L}H = [\mathscr{L}, f \prec]G$. Moreover $H \in C_T\mathscr{C}^{\alpha+\beta} \cap C_T^{(\alpha\wedge\beta)/2} L^\infty$ and*

$$\|H\|_{C_T^{(\alpha\wedge\beta)/2} L^\infty} + \|H\|_{\alpha+\beta} \lesssim \|f\|_{\mathscr{L}^\beta} (\|G\|_\alpha + \|\mathscr{L}G\|_{\alpha-2})$$

Proof. Let f_ε be a time mollification of f such that $\|\partial_t f_\varepsilon\|_{L^\infty} \lesssim \varepsilon^{\beta/2-1}$ and $\|f_\varepsilon - f\|_{L^\infty} \lesssim \|f\|_{\mathscr{L}^\beta} \varepsilon^{\beta/2}$ for all $\varepsilon > 0$, for example we can take $f_\varepsilon = \rho_\varepsilon * f$ with $\rho_\varepsilon(t) = \rho(t/\varepsilon)/\varepsilon$ and $\rho : \mathbb{R} \to \mathbb{R}$ compactly supported, positive, smooth and of unit integral. Consider the decomposition for $\mathscr{L}H$ given by

$$\mathscr{L}\Delta_i H = \Delta_i[(f - f_\varepsilon) \prec \mathscr{L}G - \mathscr{L}((f - f_\varepsilon) \prec G)] + \Delta_i[f_\varepsilon \prec \mathscr{L}H - \mathscr{L}(f_\varepsilon \prec H)]$$

Then

$$\begin{aligned}\mathscr{L}\Delta_i(H + (f - f_\varepsilon) \prec G) &= \Delta_i[(f - f_\varepsilon) \prec \mathscr{L}G] + \Delta_i[f_\varepsilon \prec \mathscr{L}G - \mathscr{L}(f_\varepsilon \prec G)]\\ &= \Delta_i[(f - f_\varepsilon) \prec \mathscr{L}G] + \Delta_i[\mathscr{L}f_\varepsilon \prec G + \partial_x f_\varepsilon \prec \partial_x G].\end{aligned}$$

Schauder estimates for $\mathscr{L}$ give

$$\|\Delta_i(H + (f - f_\varepsilon) \prec G)\|_{\mathscr{L}^{\alpha+\beta}} \lesssim \|\Delta_i[(f - f_\varepsilon) \prec \mathscr{L}G] + \Delta_i[\mathscr{L}f_\varepsilon \prec G + \partial_x f_\varepsilon \prec \partial_x G]\|_{C_T\mathscr{C}^{\alpha+\beta-2}}$$

Now choosing $\varepsilon = 2^{-2i}$ we have

$$\|\Delta_i((f - f_\varepsilon) \prec G)\|_{\alpha+\beta} \lesssim 2^{\beta i} \|\Delta_i((f - f_\varepsilon) \prec G)\|_\alpha \lesssim 2^{\beta i} \|f - f_\varepsilon\|_{L^\infty} \|G\|_\alpha \lesssim \|f\|_{\mathscr{L}^\beta} \|G\|_\alpha$$

and

$$\begin{aligned}\|\Delta_i[(f - f_\varepsilon) \prec \mathscr{L}G]\|_{\alpha+\beta-2} &\lesssim 2^{\beta i} \|\Delta_i[(f - f_\varepsilon) \prec \mathscr{L}G]\|_{\alpha-2} \lesssim 2^{\beta i} \|f - f_\varepsilon\|_{L^\infty} \|\mathscr{L}G\|_{\alpha-2}\\ &\lesssim \|f\|_{\mathscr{L}^\beta} \|\mathscr{L}G\|_{\alpha-2}\end{aligned}$$

and since $\beta < 1$:

$$\|\Delta_i[\mathscr{L} f_\varepsilon \prec G + \partial_x f_\varepsilon \prec \partial_x G]\|_{\alpha+\beta-2} \lesssim 2^{i(\beta-2)}\|\partial_t f_\varepsilon\|_{L^\infty}\|G\|_\alpha + \|f_\varepsilon\|_\beta \|G\|_\alpha$$
$$\lesssim \|f\|_{\mathscr{L}^\beta}\|G\|_\alpha + \|f\|_\beta\|G\|_\alpha$$

Finally

$$\|\Delta_i H\|_{\alpha+\beta} \lesssim \|f\|_{\mathscr{L}^\beta}(\|G\|_\alpha + \|\mathscr{L} G\|_{\alpha-2})$$

which implies the thesis since $\|\Delta_i H\|_{L^\infty} \lesssim 2^{-(\alpha+\beta)i}\|\Delta_i H\|_{\alpha+\beta}$. The estimate on the time regularity of H can be deduced similarly by noting that $(f - f_\varepsilon) \prec G \in C_T^{(\alpha\wedge\beta)/2} L^\infty$ uniformly in ε. □

5.3 Paracontrolled distributions

Here we build a calculus of distributions satisfying a paracontrolled ansatz. We start by a suitable space of such objects.

Definition 26. *Let $\alpha \in \mathbb{R}$, $\beta > 0$, and let $u \in \mathscr{C}^\alpha$. A pair of distributions $(f, f') \in \mathscr{C}^\alpha \times \mathscr{C}^\beta$ is called paracontrolled by g if we have a smoother remainder*

$$f^\sharp = f - f' \prec u \in \mathscr{C}^{\alpha+\beta}.$$

In that case we abuse notation and write $f \in \mathscr{D}^\beta = \mathscr{D}^\beta(u)$, and we define the norm

$$\|f\|_{\mathscr{D}^\beta} = \|f'\|_\beta + \|f^\sharp\|_{\alpha+\beta}.$$

If $\tilde{u} \in \mathscr{C}^\alpha$ and $(\tilde{f}, \tilde{f}') \in \mathscr{D}^\beta(\tilde{u})$, then we also write

$$d_{\mathscr{D}^\beta}(f, \tilde{f}) = \|f' - f'\|_\beta + \|f^\sharp - \tilde{f}^\sharp\|_{\alpha+\beta}.$$

Of course, we should really write $(f, f') \in \mathscr{D}^\beta$ since given f and g, the derivative f' is usually not uniquely determined. But in the applications there will always be an obvious candidate for the derivative, and no confusion will arise.

Remark 27. The space $\mathscr{D}^\beta$ does not depend on the specific dyadic partition of unity. Indeed, Bony [Bon81] has shown that if $\tilde{\prec}$ is the paraproduct constructed from another partition of unity, then $\|f' \prec u - f' \tilde{\prec} u\|_{\alpha+\beta} \lesssim \|f'\|_\beta \|u\|_\alpha$.

Nonlinear operations As an immediate consequence of Lemma 3, we can multiply paracontrolled distributions provided that we know how to multiply the reference distributions.

Theorem 28. *Let $\alpha, \beta \in \mathbb{R}$, $\gamma < 0$, with $\alpha + \beta + \gamma > 0$. Let $u \in \mathscr{C}^\alpha$, $v \in \mathscr{C}^\gamma$, and let $\eta \in \mathscr{C}^{\alpha+\gamma}$ be such that there exist sequences of smooth functions (u_n) and (v_n) converging to u and v respectively for which $(u_n \circ v_n)$ converges to η. Then*

$$\mathscr{D}^\beta(u) \ni f \mapsto f \cdot v := f \prec v + f \succ v + f^\sharp \circ v + C(f', u, v) + f'\eta \in \mathscr{C}^\gamma$$

defines a linear operator which admits the bound

$$\|(fv)^\sharp\|_{\alpha+\gamma} := \|f \cdot v - f \prec v\|_{\alpha+\gamma} \lesssim \|f\|_{\mathscr{D}^\beta}(\|v\|_\gamma + \|u\|_\alpha \|v\|_\gamma + \|\eta\|_{\alpha+\gamma}).$$

The operator does not depend on the partition of unity used to construct it.

Furthermore, there exists a quadratic polynomial P so that if $\tilde{u}, \tilde{v}, \tilde{\eta}$ satisfy the same assumptions as u, v, η respectively, if $\tilde{f} \in \mathscr{D}^\beta(\tilde{u})$, and if

$$M := \max\left\{\|u\|_\alpha, \|v\|_\gamma, \|\eta\|_{\alpha+\gamma}, \|\tilde{u}\|_\alpha, \|\tilde{v}\|_\gamma, \|\tilde{\eta}\|_{\alpha+\gamma}, \|f\|_{\mathscr{D}^\beta(u)}, \|\tilde{f}\|_{\mathscr{D}^\beta(\tilde{u})}\right\},$$

then

$$\|(fv)^\sharp - (\tilde{f}\tilde{v})^\sharp\|_{\alpha+\gamma} \leqslant P(M)\big(d_{\mathscr{D}^\beta}(f, \tilde{f}) + \|u - \tilde{u}\|_\alpha + \|v - \tilde{v}\|_\gamma + \|\eta - \tilde{\eta}\|_{\alpha+\gamma}\big).$$

Given Lemma 3 (and the paraproduct estimates Theorem 15), the proof is a simple exercise and we omit it. From now on we will usually write fv rather than $f \cdot v$.

To solve equations involving general nonlinear functions, we need to examine the stability of paracontrolled distributions under smooth functions.

Theorem 29. *Let* $\alpha \in (0, 1)$ *and* $\beta \in (0, \alpha]$. *Let* $u \in \mathscr{C}^\alpha$, $f \in \mathscr{D}^\beta(u)$, *and* $F \in C_b^{1+\beta/\alpha}$. *Then* $F(f) \in \mathscr{D}^\beta$ *with derivative* $(F(f))' = F'(f) f'$, *and*

$$\|F(f)\|_{\mathscr{D}^\beta} \lesssim \|F\|_{C_b^{1+\beta/\alpha}} (1 + \|f\|_{\mathscr{D}^\beta}^2)(1 + \|u\|_\alpha^2).$$

Moreover, there exists a polynomial P *which satisfies for all* $F \in C_b^{2+\beta/\alpha}$, $\tilde{u} \in \mathscr{C}^\alpha$, $\tilde{f} \in \mathscr{D}^\beta(\tilde{u})$, *and*

$$M := \max\left\{ \|u\|_\alpha, \|\tilde{u}\|_\alpha, \|f\|_{\mathscr{D}^\beta(u)}, \|\tilde{f}\|_{\mathscr{D}^\beta(\tilde{u})} \right\}$$

the bound

$$d_{\mathscr{D}^\beta}(F(f), F(\tilde{f})) \leqslant P(M) \|F\|_{C_T^{2+\beta/\alpha}} \big(d_{\mathscr{D}^\beta}(f, \tilde{f}) + \|u - \tilde{u}\|_\alpha \big).$$

The proof is not very complicated but rather lengthy, and we do not present it here.

Example 30. Let B^1 and B^2 be two independent Brownian bridges on $[0, 2\pi]$. From the identification of $\mathscr{C}^\alpha$ with the space of α–Hölder continuous functions we get $B^1, B^2 \in \mathscr{C}^{1/2-}$. If $\int_0^\cdot B_s^1 \circ \mathrm{d} B_s^2$ denotes the Stratonovich integral, then we can set

$$B^1 \circ \partial_t B^2 = \partial_t \int_0^\cdot B_s^1 \circ \mathrm{d} B^2 - B^1 \prec \partial_t B - B^1 \succ \partial_t B.$$

In that case the existing results for the Stratonovich integral easily give us $B^1 \circ \partial_t B^2 \in \mathscr{C}^{0-}$ and that $(B^{1,n} \circ \partial_t B^{2,n})_n$ converges to $B^1 \circ \partial_t B^2$ for a wide range of smooth approximations.

Note that while we cannot control $B^{1,n} \circ \partial_t B^{2,n}$ using analytical arguments, there are stochastic cancellations that appear due to the correlation structure of the Brownian bridge. Theorem 28 equips us with a tool to take these cancellations into account when studying objects which "look like" the Brownian bridge, which is most useful when solving SDEs.

Of course, the theory of rough paths does exactly the same thing. And in the case of SDEs it actually does a much better job. But the advantage of Theorem 28 is that here the dimensionality of the index parameter plays no role and therefore it is applicable to SPDEs driven by more complicated objects, say space-time white noise.

Schauder estimate for paracontrolled distributions The Schauder estimate Lemma 12 is not quite enough: since we are working on spaces of paracontrolled distributions, we need to understand how the integration against the heat kernel acts on the paracontrolled structure. For that purpose let us adapt the notion of paracontrolled distributions to take into account that solutions to SPDEs will be functions of time with values in distributional spaces. We also restrict the parameter range for α and β. While one could imagine extending the result to more general settings, for the equations under consideration here the following definition will be sufficient.

Definition 31. *Let* $\alpha, \beta > 0$ *be such that* $\alpha + \beta \in (0, 2)$, *and let* $u \in \mathscr{L}^\alpha$. *A pair of distributions* $(f, f') \in \mathscr{L}^\alpha \times \mathscr{L}^\beta$ *is called paracontrolled by* g *if* $f^\sharp = f - f' \prec u \in C\mathscr{C}^{\alpha+\beta}$. *In that case we write* $f \in \mathscr{D}^\beta = \mathscr{D}^\beta(u)$, *and for all* $T > 0$ *we define the norm*

$$\|f\|_{\mathscr{D}_T^\beta} = \|f'\|_{\mathscr{L}_T^\beta} + \|f^\sharp\|_{C_T \mathscr{C}^{\alpha+\beta}} + \|f^\sharp\|_{C_T^{(\alpha \wedge \beta)/2} L^\infty}.$$

If $\tilde{u} \in \mathscr{L}^\alpha$ *and* $(\tilde{f}, \tilde{f}') \in \mathscr{D}^\beta(\tilde{u})$, *then we also write*

$$d_{\mathscr{D}_T^\beta}(f, \tilde{f}) = \|f' - \tilde{f}'\|_{\mathscr{L}_T^\beta} + \|f^\sharp - \tilde{f}^\sharp\|_{C_T \mathscr{C}^{\alpha+\beta}} + \|f^\sharp - \tilde{f}^\sharp\|_{C_T^{(\alpha \wedge \beta)/2} L^\infty}.$$

Theorem 32. *Let* $\beta \in (0, 1)$, $\alpha \in (0, 2 - \beta)$, $u \in C\mathscr{C}^{\alpha-2}$, *and let* $\mathscr{L} U = u$ *with* $U(0) = 0$. *Let* $f' \in \mathscr{L}^\beta$, $f^\sharp \in C\mathscr{C}^{\alpha+\beta-2}$, *and* $g_0 \in \mathscr{C}^{\alpha+\beta}$. *Then* $(g, f') \in \mathscr{D}^\beta(U)$, *where* g *solves*

$$\mathscr{L} g = f' \prec u + f^\sharp, \qquad g(0) = g_0,$$

and we have the bound

$$\|g\|_{\mathscr{D}_T^\beta(U)} \lesssim \|g_0\|_{\alpha+\beta} + (1+T)(\|f'\|_{\mathscr{L}_T^\beta}(1+\|u\|_{C_T\mathscr{C}^{\alpha-2}}) + \|f^\sharp\|_{C_T\mathscr{C}^{\alpha+\beta-2}})$$

for all $T>0$. If furthermore $\tilde{u}, \tilde{U}, \tilde{f}', \tilde{f}^\sharp, \tilde{g}_0, \tilde{g}$ satisfy the same assumptions as $u, U, f', f^\sharp, g_0, g$ respectively, and if $M = \max\{\|f'\|_{\mathscr{L}_T^\beta}, \|\tilde{u}\|_{C_T\mathscr{C}^{\alpha-2}}, 1\}$, then

$$d_{\mathscr{D}_T^\beta}(g,\tilde{g}) \lesssim \|g_0 - \tilde{g}_0\|_{\alpha+\beta} + (1+T)M(\|f' - \tilde{f}'\|_{\mathscr{L}_T^\beta} + \|u - \tilde{u}\|_{C_T\mathscr{C}^{\alpha-2}} + \|f^\sharp - \tilde{f}^\sharp\|_{C_T\mathscr{C}^{\alpha+\beta-2}}).$$

Proof. Let us derive an equation for the remainder $g^\sharp$. We have

$$\begin{aligned}\mathscr{L}g^\sharp = \mathscr{L}(f' \prec U) - \mathscr{L}g &= [\mathscr{L}(f' \prec U) - f' \prec \mathscr{L}U] + f' \prec \mathscr{L}U - [f' \prec u + f^\sharp] \\ &= [\mathscr{L}(f' \prec U) - f' \prec \mathscr{L}U] - f^\sharp.\end{aligned}$$

We know also that there exists $H \in C\mathscr{C}^{\alpha+\beta}$ such that $\mathscr{L}H = [\mathscr{L}, f' \prec]U$ so we can apply standard Schauder estimates of Lemma 12 to $\mathscr{L}(g^\sharp - H) = -f^\sharp$ to get

$$\|g^\sharp\|_{C_T\mathscr{C}^{\alpha+\beta}} + \|g^\sharp\|_{C_T^{(\alpha\wedge\beta)/2}L^\infty} \lesssim \|f'\|_{\mathscr{L}_T^\beta}(\|U\|_{C_T\mathscr{C}^\alpha} + \|\mathscr{L}U\|_{C_T\mathscr{C}^{\alpha-2}}) + \|f^\sharp\|_{C_T\mathscr{C}^{\alpha+\beta-2}}.$$

The estimate for $g^\sharp - \tilde{g}^\sharp$ can be derived in the same way. □

Bibliographic notes Paraproducts were introduced in [Bon81], for a nice introduction see [BCD11]. The commutator estimate Lemma 3 is from [GIP13], but the proof here is new and the statement is slightly different. In [GIP13], we require the additional assumption $\alpha \in (0,1)$ under which C maps $\mathscr{C}^\alpha \times \mathscr{C}^\beta \times \mathscr{C}^\gamma$ to $\mathscr{C}^{\alpha+\beta+\gamma}$ and not only to $\mathscr{C}^{\beta+\gamma}$. Theorem 29 is from [GIP13] and relies on a paralinearization result due to Bony.

Theorem 32 is new, but it is implicitly used in [GIP13, GP14]. The estimates presented here will only allow us to consider regular initial conditions. More general situations can be covered by working on "explosive spaces" of the type

$$\{f \in C((0,\infty), \mathscr{C}^\alpha) \colon \sup_{t\in(0,T]} \|t^{-\gamma} f(t)\|_{\mathscr{C}^\alpha} < \infty \text{ for all } T>0\}$$

and similar for the temporal regularity. This is also done in [GIP13, GP14].

Of course it is easily possible to replace the Laplacian by more general pseudo-differential operators. We only used two properties of Δ: the fact that $\Delta(f' \prec U) - f' \prec (\Delta U)$ is smooth, and that the semigroup generated by Δ has a sufficiently strong regularization effect. This is also true for fractional Laplacians and more generally for a wide range of pseudo-differential operators.

5.4 Fixpoint

Let us now give the details of the solution to PAM in the space of paracontrolled distributions. Assume that $F \colon \mathbb{R} \to \mathbb{R}$ is in $C_b^{2+\varepsilon}$ for some $\varepsilon > 0$ such that $(2+\varepsilon)\gamma > 2$.

We know from Theorem 29 that if $X \in C\mathscr{C}^\gamma$ and $u \in \mathscr{D}^\gamma(X)$, then $F(u) \in \mathscr{D}^\gamma(X)$. So if $X \circ \eta \in C\mathscr{C}^{2\gamma-2}$ is given, then Theorem 28 shows that $F(u)\eta = (F(u)\eta)^\sharp + F(u) \prec \eta$ with $(F(u)\eta)^\sharp \in C\mathscr{C}^{2\gamma-2}$. Integrating against the heat kernel and assuming that $u_0 \in \mathscr{C}^{2\gamma}$, we obtain from Theorem 32 that the solution v to $\mathscr{L}v = F(u)\eta$, $v(0) = u_0$, is in $\mathscr{D}^\gamma(\vartheta)$, where ϑ solves $\mathscr{L}\vartheta = \eta$ and $\vartheta(0) = 0$. So for all $T > 0$ we can define a map

$$\Gamma_T \colon \mathscr{D}_T^\gamma(X) \to \mathscr{D}_T^\gamma(\vartheta), \qquad \Gamma(u|_{[0,T]}) = v|_{[0,T]}.$$

We want to set up a Picard iteration using Γ_T, so domain and image space should coincide which means we should take $X = \vartheta$. Refining the analysis, we obtain a scaling factor T^δ when estimating the $\mathscr{D}_T^\gamma(\vartheta)$–norm of v. This allows us to show that for small $T > 0$, Γ_T leaves suitable balls in $\mathscr{D}_T^\gamma(\vartheta)$ invariant, and therefore we obtain the (local in time) *existence* of solutions to the equation under the assumption $\vartheta \circ \eta \in C\mathscr{C}^{2\gamma-2}$.

To obtain *uniqueness*, it suffices to note that by Theorem 29 the map $u \mapsto F(u)$ is locally Lipschitz continuous from $\mathscr{D}_T^\gamma(\vartheta)$ to $\mathscr{D}_T^{\varepsilon\gamma}(\vartheta)$ (recalling that $F \in C_b^{2+\varepsilon}$), while Theorem 28 and Theorem 32 show that multiplication with ξ in concatenation with integration against the heat kernel defines a Lipschitz continuous map from $\mathscr{D}_T^{\varepsilon\gamma}(\vartheta)$ to $\mathscr{D}_T^\gamma(\vartheta)$ (it is here that we use $\varepsilon\gamma + 2\gamma - 2 > 0$). Again we can obtain a scaling factor T^δ, so that Γ_T defines a contraction from $\mathscr{D}_T^\gamma(\vartheta)$ to $\mathscr{D}_T^\gamma(\vartheta)$ for some small $T > 0$.

Even better, Γ_T not only depends locally Lipschitz continuously on u, but also on the extended data $(u_0, \eta, \vartheta \circ \eta)$, and therefore the solution to (10) depends locally Lipschitz continuously on $(u_0, \eta, \vartheta \circ \eta)$.

5.5 Renormalization

So far we argued under the assumption that $\vartheta \circ \eta$ exists and has a sufficient regularity. This should be understood via approximations as the existence of a sequence of smooth functions (η_n) that converges to η, such that $(\vartheta_n \circ \eta_n)$ converges to $\vartheta \circ \eta$. However this hypothesis is questionable and, recalling our homogeneisation setting, actually not satisfied at all. More concretely this can be checked on some approximation sequence when for example $\eta = \xi$ is the 2d space white noise. Indeed, if φ is a Schwartz function on $\mathbb{R}^2$ and if $\varphi_n = n\varphi(n\cdot)$ and

$$\eta_n(x) = \varphi_n * \xi(x) = \int_{\mathbb{R}^2} \varphi_n(x - y)\xi(y)\mathrm{d}y = \sum_{k \in \mathbb{Z}^2} \langle \xi, \varphi_n(x + 2\pi k - \cdot)\rangle,$$

then we will see below that there exist constants (c_n) with $\lim_n c_n = \infty$, such that $(\vartheta_n \circ \eta_n - c_n)$ converges in $C_T\mathscr{C}^{2\gamma-2}$ for all $T > 0$.

This is not a problem with this specific approximation. The homogeneisation setting shows that even when $\eta \to 0$ there are cases where the limiting equation is nontrivial. In the paracontrolled setting we have continuous dependence on data so this non–triviality of the limit can only mean that is $\vartheta \circ \eta$ which does not converge to zero.

Another way to see that there is indeed a problem consider the following representation of the resonant term: use $\mathscr{L}\vartheta = \eta$ to write

$$\vartheta \circ \eta = \vartheta \circ \mathscr{L}\vartheta = \frac{1}{2}\mathscr{L}(\vartheta \circ \vartheta) + \frac{1}{2}\partial_x\vartheta \circ \partial_x\vartheta = \frac{1}{2}|\partial_x\vartheta|^2 + \frac{1}{2}\mathscr{L}(\vartheta \circ \vartheta) - \partial_x\vartheta \prec \partial_x\vartheta.$$

Integrating this equation over the torus and over $t \in [0, T]$ we get

$$\int_0^T \int_{\mathbb{T}} \vartheta \circ \eta = \frac{1}{2}\int_0^T \int_{\mathbb{T}} |\partial_x\vartheta|^2 + \frac{1}{2}\int_{\mathbb{T}} (\vartheta(T) \circ \vartheta(T)) - \int_0^T \int_{\mathbb{T}} (\partial_x\vartheta \prec \partial_x\vartheta)$$

but now if $\vartheta \circ \eta \in C_T\mathscr{C}^{2\gamma-2}$ and $\vartheta \in C_T\mathscr{C}^\gamma$ then all the terms should be well defined and finite but this is cannot be since would mean that $\int_0^T \int_{\mathbb{T}} |\partial_x\vartheta|^2 < +\infty$. By direct computation however we can check that

$$\int_{\mathbb{T}} |\partial_x\vartheta(t, \cdot)|^2 = +\infty$$

for any $t > 0$ almost surely if η is the space white noise. Note also that the problematic term $|\partial_x\vartheta|^2$ is exactly the correction term $(\partial_x X(t, x))^2$ appearing in the analysis of the linear homogeneisation problem above.

In order to obtain convergence of the smooth solutions in general, we should introduce corrections to the equation to remove the divergent constant c_n. Let us see where the resonant product $\vartheta \circ \eta$ appears. We have

$$(F(u)\eta)^\sharp = F(u) \succ \eta + (F(u))^\sharp \circ \eta + C((F(u))', \vartheta, \eta) + (F(u))'(\vartheta \circ \eta).$$

Now $(F(u))' = F'(u)u'$ by Theorem 29, and if u solves the equation, then $u' = F(u)$ by Theorem 32. So we should really consider the renormalized equation

$$\mathscr{L}u_n = F(u_n) \diamond \eta_n = F(u_n)\xi_n - F'(u_n)F(u_n)c_n$$

and in the limit

$$\mathscr{L}u = F(u)\diamond\eta := (F(u)\diamond\eta)^\sharp + F(u) \prec \eta, \tag{22}$$

where the paracontrolled product $(F(u)\diamond\eta)^\sharp$ is calculated using $\vartheta\diamond\eta = \lim_n (\vartheta_n \circ \eta_n - c_n)$. Formally, we also denote this product by $F(u)\diamond\eta = F(u)\eta - F'(u)u'\cdot\infty$, so that the solution u will satisfy

$$\mathscr{L}u = F(u) - F'(u)F(u)\cdot\infty.$$

Note that the correction term has exactly the same form as the Itô/Stratonovich corrector for SDEs and indeed we should consider (22) as the "Itô form" of the equation: one can show that for $F(u) = u$ and for the right choice of the constants (c_n), our solution agrees with the Wick product solution of [Hu02].

Remark 33. The convergence properties of $(\vartheta_n \circ \eta_n)$ are in stark contrast to the case of ODEs: If we replace η_n by $\partial_t B_n$, then we should replace ϑ_n by B_n. In one dimension we then have $B_n \circ \partial_t B_n = 1/2\partial_t(B_n \circ B_n)$ for smooth data, so that convergence of the term $B_n \circ \partial_t B_n$ comes for free with the convergence of B_n (which is of course false in higher dimensions). This comes from the Leibniz rule for the first order differential operator ∂_t. For $\mathscr{L}$ we have different rules and as we have discussed

$$(\vartheta_n \circ \eta_n) = (\vartheta_n \circ \mathscr{L}\vartheta_n) = \frac{1}{2}\mathscr{L}(\vartheta_n \circ \vartheta_n) + (\partial_x\vartheta_n \circ \partial_x\vartheta_n).$$

Since convergence of $(\partial_x\vartheta_n \circ \partial_x\vartheta_n)$ is equivalent to the convergence of the positive term $|\partial_x\vartheta_n|^2$, we cannot hope to have simple cancellation properties and this is the reason why we need to introduce an additive renormalization when considering the parabolic Anderson model while this renormalization is not needed in the 1d ODE setting.

These considerations lead us to give the following definition.

Definition 34. *(Rough distribution) Let $\gamma \in (2/3, 1)$ and let*

$$\mathcal{Y}_{\text{pam}} \subseteq \mathscr{C}^{\gamma-2} \times C\mathscr{C}^{2\gamma-2}$$

be the closure of the image of the map

$$\Theta_{\text{pam}}: C^\infty \times C([0,\infty), \mathbb{R}) \to \mathcal{Y}_{\text{pam}},$$

given by

$$\Theta_{\text{pam}}(\theta, f) = (\theta, \Phi\diamond\theta) := (\theta, \Phi \circ \theta - f), \tag{23}$$

where $\Phi = J\theta$, that is $\mathscr{L}\Phi = \theta$ and $\Phi(0) = 0$. We will call $\Theta_{\text{pam}}(\theta, f)$ the renormalized PAM-enhancement of the driving distribution θ. For $T > 0$ we define $\mathcal{Y}_{\text{pam}}(T) = \mathcal{Y}_{\text{pam}}|_{[0,T]}$ and we write $\|\mathbb{Y}\|_{\mathcal{Y}_{\text{pam}}(T)}$ for the norm of $\mathbb{Y} \in \mathcal{Y}_{\text{pam}}(T)$ in the Banach space $\mathscr{C}^{\gamma-2} \times C_T\mathscr{C}^{2\gamma-2}$. Moreover, we define the distance $d_{\mathcal{Y}_{\text{pam}}(T)}(\mathbb{Y}, \tilde{\mathbb{Y}}) = \|\mathbb{Y} - \tilde{\mathbb{Y}}\|_{\mathcal{Y}_{\text{pam}}(T)}$.

Remark 35. It would be more elegant to renormalize $\Phi \circ \theta$ with a constant and not with a time-dependent function (and indeed this is what we need to do to recover the Wick product solution in the linear case). But since we chose $\Phi(0) = 0$, we have $\Phi(0) \circ \theta = 0$ and therefore $(\Phi_n(0) \circ \theta_n - c_n)$ diverges for any diverging sequence of constants (c_n). A simple way of avoiding this problem is to consider the stationary version $\tilde{\Phi}$ given by

$$\tilde{\Phi}(x) = \int_0^\infty P_t \Pi_{\neq 0}\theta(x)\mathrm{d}t,$$

where $\Pi_{\neq 0}$ denotes the projection on the non-zero Fourier modes, $\Pi_{\neq 0}u = u - \hat{u}(0)$. But then $\tilde{\Phi}$ does not depend on time and in particular $\tilde{\Phi}(0) \neq 0$, so that we have to consider irregular initial conditions in the paracontrolled approach, which require "explosive" Besov spaces and complicate the presentation. Alternatively, we could observe that in the white noise case there exist constants (c_n) so that $(\vartheta_n(t) \circ \xi_n - c_n)$ converges for all $t > 0$, and while the limit $(\vartheta(t)\diamond\xi)$ diverges as $t \to 0$, it can be integrated against the heat kernel. Again, this would complicate the presentation and here we choose the simple (and cheap) solution of taking a time-dependent renormalization.

Theorem 36. *Let $\gamma \in (2/3, 1)$ and $\varepsilon > 0$ be such that $(2+\varepsilon)\gamma > 2$. Let $\mathbb{Y} \in \mathcal{Y}_{\mathrm{pam}}$, $F \in C_b^{2+\varepsilon}$, and $u_0 \in \mathscr{C}^{2\alpha}$. Then there exists a unique solution $u \in \mathscr{D}^\gamma(\Phi)$ to the equation*

$$\mathscr{L}u = F(u) \diamond \theta, \qquad u(0) = u_0,$$

up to the (possibly finite) explosion time $\tau = \tau(u) = \inf\{t \geqslant 0 \colon \|u\|_{\mathscr{D}_t^\alpha} = \infty\} > 0$.

Moreover, u depends on $(u_0, \mathbb{Y}) \in \mathscr{C}^{2\gamma} \times \mathcal{Y}_{\mathrm{pam}}$ in a locally Lipschitz continuous way: if $M, T > 0$ are such that for all $(u_0, \mathbb{Y})$ with $\|u_0\|_{2\gamma} \vee \|\mathbb{Y}\|_{\mathcal{Y}_{\mathrm{pam}}(T)} \leqslant M$, the solution u to the equation driven by $(u_0, \mathbb{Y})$ satisfies $\tau(u) > T$, and if $(\tilde{u}_0, \tilde{\mathbb{Y}})$ is another set of data bounded in the above sense by M, then there exists $C(F, M) > 0$ for which

$$d_{\mathscr{D}_T^\alpha}(u, \tilde{u}) \leqslant C(F, M)(\|u_0 - \tilde{u}_0\|_{2\gamma} + d_{\mathcal{Y}_{\mathrm{pam}}(T)}(\mathbb{Y}, \tilde{\mathbb{Y}})).$$

Remark 37. It is not necessary to assume $F \in C_b^{2+\varepsilon}$, for the local in time existence it suffices if $F \in C^{2+\varepsilon}$. This can be seen by considering a ball containing $u_0(x)$ for all $x \in \mathbb{T}^d$, by considering $\tilde{F} \in C_b^{2+\varepsilon}$ which coincides with F on this ball, and by stopping u upon exiting the ball.

In the linear case $F(u) = u$ we actually have global in time solutions. In the general case, we only have local in time solutions because we pick up a quadratic estimate when applying the paralinearization result Theorem 29. This step is not necessary if F is linear, and all the other estimates are linear in u. When $F \in C_b^3$ a more refined argument can get rid of the quadratic estimate and estabilish global existence in full generality. For the details see [GIP13].

Proof. It is an easy exercise to turn the formal discussion above into rigorous mathematics to obtain existence and uniqueness on a small time interval. It remains to show three things: how to pick up the scaling factor T^δ from the map Γ_T, how to iterate the construction so that we get existence up to the explosion time, and how to obtain the local Lipschitz continuity.

As for the scaling factor, note that in the estimate of Theorem 28 we have considerable space in the parameter β, so that we can estimate

$$\|\Gamma_T(u)\|_{\mathscr{D}_T^\gamma} \leqslant C_{F,\mathbb{Y}}(1 + \|u_0\|_{2\gamma} + \|u\|_{\mathscr{D}_T^\beta}^2)$$

for all β with $2\gamma + \beta > 0$, with a constant $C_{F,\mathbb{Y}}$ depending only on $\|F\|_{C_b^2}$ and $\|\mathbb{Y}\|_{\mathcal{Y}_{\mathrm{pam}}(T)}$. But now any fixed point u for Γ_T will satisfy $u'(0) = F(u_0)$ and $u^\sharp(0) = u_0$ (since $\vartheta(0) = 0$), so that we may restrict our attention to those u. Then we can use the estimate

$$\|u'\|_{\mathscr{L}_T^\beta} \lesssim \|u'(0)\|_\beta + T^{(\gamma-\beta)/2}\|u'\|_{\mathscr{L}_T^\gamma}$$

for all $\beta \in (0, \gamma)$, and similarly for $\|u^\sharp\|_{C_T\mathscr{C}^{\gamma+\beta}} + \|u^\sharp\|_{C_T^{(\alpha\wedge\beta)/2}L^\infty}$, to show that for small $T = T(M) > 0$ and suitable $C(M) > 0$, the map Γ_T leaves the set $\{u \in \mathscr{D}_T^\gamma \colon (u', u^\sharp)(0) = (F(u_0), u_0), \|u\|_{\mathscr{D}_T^\gamma} \leqslant C(M)\}$ invariant. In the same way we can show that Γ_T defines a contraction on a comparable set, possibly after further decreasing T.

Now let us assume that we constructed the solution u on $[0, T_0]$ for some $T_0 > 0$. Then we slightly have to adapt our arguments to extend u to $[T_0, T_0 + T]$: Now we no longer have $\vartheta(T_0) = 0$, and also the initial condition $u(T_0)$ is no longer in $\mathscr{C}^{2\gamma}$. But we only used that $\vartheta(0) = 0$ to derive the initial conditions for u' and $u^\sharp$. Since we already know u on $[0, T_0]$, we do not need this anymore. And we used $u_0 \in \mathscr{C}^{2\gamma}$ only to obtain smooth initial conditions for $u^\sharp$. But again, $u^\sharp(T_0) \in \mathscr{C}^{2\gamma}$ is already fixed.

Let us come to the continuity in $(u_0, \mathbb{Y})$. If $(\tilde{u}_0, \tilde{\mathbb{Y}})$ is another set of data also bounded by M, then we know that the solutions u and $\tilde{u}$ both are bounded in $\mathscr{D}_T^\gamma$ by some constant $C = C(F, M) > 0$. So by the continuity properties of the paracontrolled product (and the other operations involved), we can estimate

$$d_{\mathscr{D}_T^\alpha}(u, \tilde{u}) \leqslant P(C)\big(\|u_0 - \tilde{u}_0\|_{2\gamma} + d_{\mathcal{Y}_{\mathrm{pam}}(T)}(\mathbb{Y}, \tilde{\mathbb{Y}}) + T^{(\gamma-\beta)/2} d_{\mathscr{D}_T^\gamma}(u, \tilde{u})\big)$$

for suitable $\beta < \gamma$ and for a polynomial P. The local Lipschitz continuity on $[0, T]$ immediately follows if we choose $T > 0$ small enough. This can be iterated to obtain the local Lipschitz continuity on "macroscopic" intervals. □

5.6 Construction of the extended data

In order to apply Theorem 36 to eq. (10) with white noise perturbation, it remains to show that if ξ is a spatial white noise on $\mathbb{T}^2$, then ξ defines an element of $\mathcal{Y}_{\text{pam}}$. In other words, we need to construct $\vartheta \diamond \xi$.

Since $P_t \xi$ is a smooth function for every $t > 0$, the resonant term $P_t \xi \circ \xi$ is a smooth function, and therefore we could formally set $\vartheta \circ \xi = \int_0^\infty (P_t \xi \circ \xi)\, \mathrm{d}\, t$. But we will see that this expression does not make sense.

We will need that $(\hat{\xi}(k))_{k \in \mathbb{Z}^2}$ is a complex valued, centered Gaussian process with covariance

$$\mathbb{E}[\hat{\xi}(k) \hat{\xi}(k')] = (2\pi)^2 1_{k=-k'}, \tag{24}$$

and such that $\overline{\hat{\xi}(k)} = \hat{\xi}(-k)$.

Lemma 38. *For any $x \in \mathbb{T}^2$ and $t > 0$ we have*

$$g_t = \mathbb{E}[(P_t \xi \circ \xi)(x)] = \mathbb{E}[\Delta_{-1} (P_t \xi \circ \xi)(x)] = (2\pi)^{-2} \sum_{k \in \mathbb{Z}^2} e^{-t|k|^2}.$$

In particular, g_t does not depend on the partition of unity used to define the $\circ$ operator, and $\int_0^t g_s \,\mathrm{d}\, s = \infty$ for all $t > 0$.

Proof. Let $x \in \mathbb{T}^2$, $t > 0$, and $\ell \geq -1$. Then

$$\mathbb{E}[\Delta_\ell (P_t \xi \circ \xi)(x)] = \sum_{|i-j| \leqslant 1} \mathbb{E}[\Delta_\ell (\Delta_i (P_t \xi) \Delta_j \xi)(x)],$$

where exchanging summation and expectation is justified because it can be easily verified that the partial sums of $\Delta_\ell (P_t \xi \circ \xi)(x)$ are uniformly L^p–bounded for any $p \geq 1$. Now $P_t = e^{-t|\cdot|^2}(\mathrm{D})$, and therefore we get from (24)

$$\begin{aligned}
&\mathbb{E}[\Delta_\ell (\Delta_i (P_t \xi) \Delta_j \xi)(x)] \\
&\quad = (2\pi)^{-4} \sum_{k,k' \in \mathbb{Z}^2} e^{i\langle k+k', x\rangle} \rho_\ell (k+k')\, \rho_i(k)\, e^{-t|k|^2} \rho_j(k') \mathbb{E}[\hat{\xi}(k) \hat{\xi}(k')] \\
&\quad = (2\pi)^{-2} \sum_{k \in \mathbb{Z}^2} \rho_\ell(0)\, \rho_i(k)\, e^{-t|k|^2} \rho_j(k) = (2\pi)^{-2} 1_{\ell=-1} \sum_{k \in \mathbb{Z}^2} \rho_i(k)\, \rho_j(k)\, e^{-t|k|^2}.
\end{aligned}$$

For $|i-j| > 1$ we have $\rho_i(k)\, \rho_j(k) = 0$. This implies, independently of $x \in \mathbb{T}^2$, that

$$g_t = \mathbb{E}[(P_t \xi \circ \xi)(x)] = (2\pi)^{-2} \sum_{k \in \mathbb{Z}^2} \sum_{i,j} \rho_i(k)\, \rho_j(k)\, e^{-t|k|^2} = (2\pi)^{-2} \sum_{k \in \mathbb{Z}^2} e^{-t|k|^2},$$

while $\mathbb{E}[(P_t \xi \circ \xi)(x) - \Delta_{-1} (P_t \xi \circ \xi))(x)] = 0$. □

Exercise 5. Let φ be a Schwartz function on $\mathbb{R}^2$ and set

$$\xi_n(x) = ((n^2 \varphi(n\,\cdot\,)) * \xi)(x) = \int_{\mathbb{R}^2} n^2 \varphi(n(x-y)) \xi(y) \mathrm{d}y = \sum_{k \in \mathbb{Z}^2} \langle \xi, n^2 \varphi(n(x + 2\pi k - \cdot)) \rangle$$

for $x \in \mathbb{T}^2$. Write $\mathscr{F}_{\mathbb{R}^2} \varphi(z) = \int_{\mathbb{R}^2} e^{-i\langle z, x\rangle} \varphi(x) \mathrm{d}x$. Show that

$$\mathbb{E}[(P_t \xi_n \circ \xi_n)(x)] = \mathbb{E}[\Delta_{-1} (P_t \xi_n \circ \xi_n)(x)] = (2\pi)^{-2} \sum_{k \in \mathbb{Z}^2} e^{-t|k|^2} |\mathscr{F}_{\mathbb{R}^2} \varphi(k/n)|^2.$$

Hint: Poisson summation yields $\sum_k \psi(x - y + 2\pi k) = (2\pi)^{-2} \sum_k \mathscr{F}_{\mathbb{R}^2} \psi(k) e^{i\langle k, x-y\rangle}$ for any Schwartz function ψ.

The diverging time integral motivates us to study the renormalized product $\vartheta \circ \xi - \int_0^\cdot g_s \mathrm{d}s$, where $\int_0^\cdot g_s \mathrm{d}s$ is an "infinite function":

Lemma 39. *Set*

$$(\vartheta \diamond \xi)(t) = \int_0^t (P_s \xi \circ \xi - g_s)\, \mathrm{d}s.$$

Then $\mathbb{E}[\|\vartheta\diamond\xi\|_{C_T\mathscr{C}^{2\gamma-2}(\mathbb{T}^2)}^p]<\infty$ *for all* $\gamma<1$, $p\geq 1$, $T>0$. *Moreover, if* φ *is a Schwartz function on* $\mathbb{R}^2$ *which satisfies* $\int\varphi(x)\mathrm{d}x=1$, *and if* $\xi_n=\varphi_n*\xi$ *with* $\varphi_n=n^2\varphi(n\,\cdot\,)$ *for* $n\in\mathbb{N}$, *and* $\vartheta_n(t)=\int_0^\infty P_t\,\xi_n\mathrm{d}t$, *then*

$$\lim_{n\to\infty}\mathbb{E}[\|\vartheta\diamond\xi-(\vartheta_n\circ\xi_n-f_n)\|_{C_T\mathscr{C}^{2\gamma-2}(\mathbb{T}^2)}^p]=0$$

for all $p\geq 1$, *where for all* $x\in\mathbb{T}^2$

$$\begin{aligned}f_n(t)&=\mathbb{E}[\vartheta_n(t,x)\xi_n(x)]=\mathbb{E}[(\vartheta_n(t)\circ\xi_n)(x)]\\&=(2\pi)^{-2}\sum_{k\in\mathbb{Z}^2\setminus\{0\}}\frac{|\mathscr{F}_{\mathbb{R}^2}\varphi(k/n)|^2}{|k|^2}\big(1-e^{-t|k|^2}\big)+(2\pi)^{-2}t.\end{aligned}$$

Proof. To lighten the notation, we will only show that $\mathbb{E}[\|\vartheta\diamond\xi\|_{C_T\mathscr{C}^{2\gamma-2}}^p]<\infty$. The convergence of $(\vartheta_n\circ\xi_n-f_n)$ to $\vartheta\diamond\xi$ is shown using similar arguments, and we leave it as an exercise. Let $t>0$ and define $\Xi_t=P_t\,\xi\circ\xi-g_t$. Using Gaussian hypercontractivity, we will be able to reduce everything to estimating $\mathbb{E}[|\Delta_\ell\Xi_t(x)|^2]$ for $\ell\geqslant-1$ and $x\in\mathbb{T}^2$. Lemma 38 yields $\Delta_\ell\,g_t=0=\mathbb{E}[\Delta_\ell\,(P_t\,\xi\circ\xi)(x)]$ for $\ell\geq 0$ and $x\in\mathbb{T}^2$, and $\Delta_{-1}\,g_t=g_t=\mathbb{E}[\Delta_{-1}(P_t\xi\circ\xi)(x)]$, so that $\mathbb{E}[|\Delta_\ell\Xi_t(x)|^2]=\mathrm{Var}(\Delta_\ell(P_t\xi\circ\xi)(x))$, where Var denotes the variance. We have

$$\begin{aligned}\Delta_\ell(P_t\xi\circ\xi)(x)&=(2\pi)^{-2}\sum_{k\in\mathbb{Z}^2}e^{i\langle k,x\rangle}\rho_\ell(k)\mathscr{F}(P_t\,\xi\circ\xi))(k)\\&=(2\pi)^{-4}\sum_{k_1,k_2\in\mathbb{Z}^2}\sum_{|i-j|\leqslant 1}e^{i\langle k_1+k_2,x\rangle}\rho_\ell(k_1+k_2)\rho_i(k_1)e^{-t|k_1|^2}\hat{\xi}(k_1)\,\rho_j(k_2)\hat{\xi}(k_2).\end{aligned}$$

Hence

$$\begin{aligned}&\mathrm{Var}(\Delta_\ell(P_t\xi\circ\xi)(x))\\&\quad=(2\pi)^{-8}\sum_{k_1,k_2}\sum_{k_1',k_2'}\sum_{|i-j|\leqslant 1}\sum_{|i'-j'|\leqslant 1}e^{i\langle k_1+k_2,x\rangle}\rho_\ell(k_1+k_2)\rho_i(k_1)e^{-t|k_1|^2}\rho_j(k_2)\\&\qquad\times e^{i\langle k_1'+k_2',x\rangle}\rho_\ell(k_1'+k_2')\rho_{i'}(k_1')e^{-t|k_1'|^2}\rho_{j'}(k_2')\mathrm{cov}(\hat{\xi}(k_1)\hat{\xi}(k_2),\hat{\xi}(k_1')\hat{\xi}(k_2')),\end{aligned}$$

where exchanging summation and expectation can be justified a posteriori by the uniform L^p–boundedness of the partial sums, and where cov denotes the covariance. Now Wick's theorem ([Jan97], Theorem 1.28) gives us

$$\begin{aligned}\mathrm{cov}(\hat{\xi}(k_1)\hat{\xi}(k_2),\hat{\xi}(k_1')\hat{\xi}(k_2'))&=\mathbb{E}\big[\hat{\xi}(k_1)\hat{\xi}(k_2)\hat{\xi}(k_1')\hat{\xi}(k_2')\big]-\mathbb{E}\big[\hat{\xi}(k_1)\hat{\xi}(k_2)\big]\mathbb{E}\big[\hat{\xi}(k_1')\hat{\xi}(k_2')\big]\\&=\mathbb{E}\big[\hat{\xi}(k_1)\hat{\xi}(k_2)\big]\mathbb{E}\big[\hat{\xi}(k_1')\hat{\xi}(k_2')\big]+\mathbb{E}\big[\hat{\xi}(k_1)\hat{\xi}(k_1')\big]\mathbb{E}\big[\hat{\xi}(k_2)\hat{\xi}(k_2')\big]\\&\quad+\mathbb{E}\big[\hat{\xi}(k_1)\hat{\xi}(k_2')\big]\mathbb{E}\big[\hat{\xi}(k_2)\hat{\xi}(k_1')\big]-\mathbb{E}\big[\hat{\xi}(k_1)\hat{\xi}(k_2)\big]\mathbb{E}\big[\hat{\xi}(k_1')\hat{\xi}(k_2')\big]\\&=(2\pi)^4(1_{k_1=-k_1'}1_{k_2=-k_2'}+1_{k_1=-k_2'}1_{k_2=-k_1'}),\end{aligned}$$

and therefore

$$\begin{aligned}\mathrm{Var}(\Delta_\ell(P_t\xi\circ\xi)(x))&=\sum_{k_1,k_2}\sum_{|i-j|\leqslant 1}\sum_{|i'-j'|\leqslant 1}1_{\ell\lesssim i}1_{\ell\lesssim i'}\rho_\ell^2(k_1+k_2)\rho_i(k_1)\rho_j(k_2)\\&\quad\times\big[\rho_{i'}(k_1)\rho_{j'}(k_2)e^{-2t|k_1|^2}+\rho_{i'}(k_2)\rho_{j'}(k_1)e^{-t|k_1|^2-t|k_2|^2}\big].\end{aligned}$$

There exists $c>0$ such that $e^{-2t|k|^2}\lesssim e^{-tc2^{2i}}$ for all $k\in\mathrm{supp}(\rho_i)$ and for all $i\geq-1$. In the remainder of the proof the value of this strictly positive c may change from line to line. If $|i-j|\leqslant 1$, then we also have $e^{-t|k|^2}\lesssim e^{-tc2^{2i}}$ for all $k\in\mathrm{supp}(\rho_j)$. Thus

$$\begin{aligned}&\mathrm{Var}(\Delta_\ell\,(P_t\xi\circ\xi)(x))\\&\quad\lesssim\sum_{i,j,i',j'}1_{\ell\lesssim i}1_{i\sim j\sim i'\sim j'}\sum_{k_1,k_2}1_{\mathrm{supp}(\rho_\ell)}(k_1+k_2)1_{\mathrm{supp}(\rho_i)}(k_1)1_{\mathrm{supp}(\rho_j)}(k_2)e^{-2tc2^{2i}}\\&\quad\lesssim\sum_{i:i\gtrsim\ell}2^{2i}2^{2\ell}e^{-tc2^{2i}}\lesssim\frac{2^{2\ell}}{t}\sum_{i:i\gtrsim\ell}e^{-tc2^{2i}}\lesssim\frac{2^{2\ell}}{t}e^{-tc2^{2\ell}},\end{aligned}$$

where we used that $t 2^{2i} \lesssim e^{t(c-c')2^{2i}}$ for any $c' < c$.

Consider now $\vartheta \diamond \xi(t) = \int_0^t \Xi_s \mathrm{d}s$. We have for all $0 \leqslant s < t$

$$\mathbb{E}\Big[\|\vartheta \diamond \xi(t) - \vartheta \diamond \xi(s)\|_{B^{2\gamma-2}_{2p,2p}}^{2p}\Big] = \sum_{\ell} 2^{2p\ell(2\gamma-2)} \int_{\mathbb{T}^2} \mathbb{E}[|\Delta_\ell(\vartheta \diamond \xi(t) - \vartheta \diamond \xi(s))(x)|^{2p}] \mathrm{d}x.$$

Since the random variable $\Delta_\ell(\vartheta \diamond \xi(t) - \vartheta \diamond \xi(s))(x)$ lives in the second non-homogeneous chaos generated by the Gaussian white noise ξ, we may use Gaussian hypercontractivity (see [FV10], Appendix D.4) to obtain

$$\mathbb{E}[|\Delta_\ell(\vartheta \diamond \xi(t) - \vartheta \diamond \xi(s))(x)|^{2p}] \lesssim \mathbb{E}[|\Delta_\ell(\vartheta \diamond \xi(t) - \vartheta \diamond \xi(s))(x)|]^{2p} \leqslant \left(\int_s^t \mathbb{E}[|\Delta_\ell \Xi_r(x)|] \mathrm{d}r\right)^{2p}.$$

But we just showed that

$$\mathbb{E}[|\Delta_\ell \Xi_r(x)|] \leqslant \mathbb{E}[|\Delta_\ell \Xi_r(x)|^2]^{1/2} \lesssim r^{-1/2} 2^\ell e^{-\frac{1}{2} r c 2^{2\ell}} \leqslant r^{-1/2} 2^\ell e^{-r c 2^{2\ell}}$$

(changing again the value of c), and therefore

$$\begin{aligned}
\mathbb{E}\Big[\|\vartheta \diamond \xi(t) - \vartheta \diamond \xi(s)\|_{B^{2\gamma-2}_{2p,2p}}^{2p}\Big] &\lesssim \sum_{\ell} 2^{2p\ell(2\gamma-2)} \left(\int_s^t r^{-1/2} 2^\ell e^{-rc2^{2\ell}} \mathrm{d}r\right)^{2p} \\
&\leqslant \left(\sum_{\ell} \int_s^t r^{-1/2} 2^{\ell(2\gamma-1)} e^{-rc2^{2\ell}} \mathrm{d}r\right)^{2p} \\
&\lesssim \left(\int_s^t r^{-1/2} \int_{-1}^\infty (2^x)^{2\gamma-1} e^{-rc2^{2x}} \mathrm{d}x \,\mathrm{d}r\right)^{2p}.
\end{aligned}$$

The change of variable $y = \sqrt{r} 2^x$ leads to

$$\lesssim \left(\int_s^t r^{-1/2} r^{-(2\gamma-1)/2} \int_0^\infty y^{2\gamma-2} e^{-cy^2} \mathrm{d}y \,\mathrm{d}r\right)^{2p}.$$

For $\alpha > 1/2$, the integral in y is finite and we end up with

$$\lesssim \left(\int_s^t r^{-\gamma} \mathrm{d}r\right)^{2p} \lesssim |t-s|^{2p(1-\gamma)}$$

provided that $\gamma \in (1/2, 1)$. So for large enough p we can use Kolmogorov's continuity criterion to deduce that $\mathbb{E}\Big[\|\vartheta \diamond \xi\|_{C_T B^{2\gamma-2}_{2p,2p}}^{2p}\Big] < \infty$ for all $T > 0$. The claim now follows from the Besov embedding theorem, Lemma 11. □

Combining Theorem 36 and Lemma 39, we are finally able to solve (10).

Corollary 40. *Let $\varepsilon > 0$ and let $F \in C_b^{2+\varepsilon}$ and assume that u_0 is a random variable that almost surely takes its values in $\mathscr{C}^{2\gamma}$ for some α with $(2+\varepsilon)\gamma > 2$. Let ξ be a spatial white noise on $\mathbb{T}^2$. Then there exists a unique solution u to*

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

up to the (possibly finite) explosion time $\tau = \tau(u) = \inf\{t \geqslant 0 \colon \|u\|_{\mathscr{D}_t^\alpha} = \infty\}$ which is almost surely strictly positive.

If (φ_n) and (ξ_n) are as described in Lemma 39, and if $(u_{n,0})$ converges in probability in $\mathscr{C}^{2\gamma}$ to u_0, then u is the limit in probability of the solutions u_n to

$$\mathscr{L} u_n = F(u_n) \diamond \xi_n, \qquad u_n(0) = u_{n,0}.$$

Remark 41. We even have a stronger result: We can fix a null set outside of which $\vartheta \diamond \xi$ is regular enough, and once we dispose of that null set we can solve all equations for any regular enough u_0 and F simultaneously, without ever having to worry about null sets again. This is for example interesting when studying stochastic flows.

The pathwise continuous dependence on the signal is also powerful in several other applications, for example support theorems and large deviations. For examples in the theory of rough paths see [FV10].

6 The stochastic Burgers equation

Let us now get to our main example, the "KPZ family" of equations. We concentrate here on the Burgers equation SBE, but essentially the same analysis works for the KPZ equation. We can also treat the heat equation in the same way, although in that case we need to set up the equation in the right way. We will indicate how to do this after we treat Burgers equation.

Recall that Burgers equation is

$$\mathscr{L}u = \partial_x u^2 + \partial_x \xi, \qquad u(0) = u_0, \tag{25}$$

where $u: [0, \infty) \times \mathbb{T} \to \mathbb{R}$, ξ is a space-time white noise, and ∂_x denotes the spatial derivative. To begin with let us analyze the difficulty for this equation. As we argued before, the solution u cannot be expected to behave better than the solution X of the linear equation

$$\mathscr{L}X = \partial_x \xi$$

(for example with zero initial condition at time 0). Arguing similarly as in Exercise 4, one can show that almost surely $X \in C\mathscr{C}^{-1/2-}$. In particular, u^2 is the square of a distribution, and therefore a priori it does not make sense.

But what raises some hope though is that if X_n denotes the solution to the linear equation with regularized noise ξ_n, then $(\partial_x X_n^2)_n$ converges to a space–time distribution $\partial_x X^2$. So as in the previous examples there are stochastic cancellations going into $\partial_x X^2$, and Theorem 28 will allow us to take these cancellations into account in the full solution u.

6.1 Structure of the solution

In this discussion we consider the case of zero initial condition and smooth noise ξ, and we analyze the structure of the solution. Let us expand u around the solution X to the linear equation $\mathscr{L}X = \partial_x \xi$, $X(0) = 0$. Setting $u = X + u^{\geqslant 1}$, we have

$$\mathscr{L}u^{\geqslant 1} = \partial_x(u^2) = \partial_x(X^2) + 2\partial_x(X u^{\geqslant 1}) + \partial_x((u^{\geqslant 1})^2).$$

Let us define the bilinear map

$$B(f, g) = J\partial_x(fg) = \int_0^{\cdot} P_{\cdot - s}\partial_x(f(s)g(s))\mathrm{d}s.$$

Then we can proceed by performing a further change of variables in order to remove the term $\partial_x(X^2)$ from the equation by setting

$$u = X + B(X, X) + u^{\geqslant 2}. \tag{26}$$

Now $u^{\geqslant 2}$ satisfies

$$\begin{aligned} \mathscr{L}u^{\geqslant 2} &= 2\partial_x(X B(X, X)) + \partial_x(B(X, X)B(X, X)) \\ &\quad + 2\partial_x(X u^{\geqslant 2}) + 2\partial_x(B(X, X) u^{\geqslant 2}) + \partial_x((u^{\geqslant 2})^2). \end{aligned} \tag{27}$$

We can imagine to make a similar change of variables to get rid of the term

$$2\partial_x(X B(X, X)) = \mathscr{L}B(X, B(X, X)).$$

As we proceed in this inductive expansion, we generate a number of explicit terms, obtained by various combinations of X and B. Since we will have to deal explicitly with at least some of these terms, it is convenient to represent them with a compact notation involving binary trees. A binary tree $\tau \in \mathcal{T}$ is either the root $\bullet$ or the combination of two smaller binary trees $\tau = (\tau_1 \tau_2)$, where the two edges of the root of τ are attached to τ_1 and τ_2 respectively.

Then we define recursively

$$X^{\bullet} = X, \qquad X^{(\tau_1 \tau_2)} = B(X^{\tau_1}, X^{\tau_2}),$$

giving

$$X^{\vee} = B(X, X), \quad X^{(\vee\bullet)} = B(X, X^{\vee}), \quad X^{(\bullet(\vee\bullet))} = B(X, X^{(\vee\bullet)}), \quad X^{(\vee\vee)} = B(X^{\vee}, X^{\vee})$$

and so on, where

$$(\bullet\bullet) = \vee, \quad (\vee\bullet) = (\vee\bullet), \quad (\bullet(\vee\bullet)) = (\bullet(\vee\bullet)), \quad (\vee\vee) = (\vee\vee), \quad \ldots$$

In this notation, the expansion (26)-(27) reads

$$u = X + X^{\vee} + u^{\geqslant 2}, \tag{28}$$

$$u^{\geqslant 2} = 2X^{(\vee\bullet)} + X^{(\vee\vee)} + 2B(X, u^{\geqslant 2}) + 2B(X^{\vee}, u^{\geqslant 2}) + B(u^{\geqslant 2}, u^{\geqslant 2}). \tag{29}$$

Remark 42. We observe that formally the solution u of equation (25) can be expanded as an infinite sum of terms labelled by binary trees:

$$u = \sum_{\tau \in \mathcal{T}} c(\tau) X^{\tau},$$

where $c(\tau)$ is a combinatorial factor counting the number of planar trees which are isomorphic (as graphs) to τ. For example $c(\bullet) = 1$, $c(\vee) = 1$, $c((\vee\bullet)) = 2$, $c((\bullet(\vee\bullet))) = 4$, $c((\vee\vee)) = 1$ and in general $c(\tau) = \sum_{\tau_1, \tau_2 \in \mathcal{T}} 1_{(\tau_1 \tau_2) = \tau}\, c(\tau_1) c(\tau_2)$. Alternatively, we may truncate the summation at trees of degree at most n and set

$$u = \sum_{\tau \in \mathcal{T}, d(\tau) < n} c(\tau) X^{\tau} + u^{\geqslant n},$$

where we denote by $d(\tau) \in \mathbb{N}_0$ the degree of the tree τ, given by $d(\bullet) = 0$ and then inductively $d((\tau_1 \tau_2)) = 1 + d(\tau_1) + d(\tau_2)$. For example $d(\vee) = 1$, $d((\vee\bullet)) = 2$, $d((\bullet(\vee\bullet))) = 3$, $d((\vee\vee)) = 3$. We then obtain for the remainder

$$u^{\geqslant n} = \sum_{\substack{\tau_1, \tau_2 : d(\tau_1) < n, d(\tau_2) < n \\ d((\tau_1 \tau_2)) \geqslant n}} c(\tau_1) c(\tau_2) X^{(\tau_1 \tau_2)} + \sum_{\tau : d(\tau) < n} c(\tau) B(X^{\tau}, u^{\geqslant n}) + B(u^{\geqslant n}, u^{\geqslant n}). \tag{30}$$

Our aim is to control the truncated expansion under the natural regularity assumptions in the white noise case, where we have $X \in C\mathscr{C}^{-1/2-}$. Since (30) contains the term $B(X, u^{\geqslant n})$ which in turn contains the paraproduct $J\partial_x(u^{\geqslant n} \prec X)$, the remainder $u^{\geqslant n}$ will be at best in $C\mathscr{C}^{1/2-}$. But then the sum of the regularities of X and $u^{\geqslant n}$ is negative, and the term $B(X, u^{\geqslant n})$ is a priori not well defined. We therefore continue the expansion up to the point (turning out to be $u^{\geqslant 3}$) where we can set up a paracontrolled Ansatz for the remainder, which will allow us to make sense of $X \circ u^{\geqslant n}$ and thus of $B(X, u^{\geqslant n})$.

6.2 Paracontrolled solution

Inspired by the partial tree series expansion of u, we set up a paracontrolled Ansatz of the form

$$u = X + X^{\vee} + 2X^{(\vee\bullet)} + u^{Q}, \qquad u^{Q} = u' \prec Q + u^{\sharp}, \tag{31}$$

where the functions u', Q and $u^{\sharp}$ are for the moment arbitrary, but we assume $u', Q \in \mathscr{L}^{\alpha}$ and $u^{\sharp} \in \mathscr{L}^{2\alpha}$, where from now on we fix $\alpha \in (1/3, 1/2)$. For such u, the nonlinear term takes the form

$$\begin{aligned} \partial_x u^2 = {} & \partial_x(X^2 + 2X^{\vee}X + (X^{\vee})^2 + 4X^{(\vee\bullet)}X) + 2\partial_x(u^{Q}X) \\ & + 2\partial_x(X^{\vee}(u^{Q} + 2X^{(\vee\bullet)})) + \partial_x((u^{Q} + 2X^{(\vee\bullet)})^2). \end{aligned} \tag{32}$$

As a consequence, we derive the following equation for u^{Q}:

$$\begin{aligned} \mathscr{L}u^{Q} = {} & \partial_x((X^{\vee})^2 + 4X^{(\vee\bullet)}X) + 2\partial_x(u^{Q}X) + 2\partial_x(X^{\vee}(u^{Q} + 2X^{(\vee\bullet)})) + \partial_x((u^{Q} + 2X^{(\vee\bullet)})^2) \\ = {} & \mathscr{L}X^{(\vee\vee)} + 4\mathscr{L}X^{(\bullet(\vee\bullet))} + 2\partial_x(u^{Q}X) + 2\partial_x(X^{\vee}(u^{Q} + 2X^{(\vee\bullet)})) + \partial_x((u^{Q} + 2X^{(\vee\bullet)})^2). \end{aligned} \tag{33}$$

If we formally apply the paraproduct estimates Theorem 15 (which is of course not possible since the regularity requirements for the resonant term are not satisfied), we derive the following natural regularities for the driving terms: $X \in C\mathscr{C}^{-1/2-}$, $X^{\vee} \in C\mathscr{C}^{0-}$, $X^{\vee_3} \in \mathscr{L}^{1/2-}$, $X^{\vee_4} \in \mathscr{L}^{1/2-}$, and $X^{\vee\vee} \in \mathscr{L}^{1-}$. In terms of α, we can encode this as

$$X \in C\mathscr{C}^{\alpha-1}, \quad X^{\vee} \in C\mathscr{C}^{2\alpha-1}, \quad X^{\vee_3} \in \mathscr{L}^{\alpha}, \quad X^{\vee_4} \in \mathscr{L}^{\alpha}, \quad X^{\vee\vee} \in \mathscr{L}^{2\alpha}.$$

Under these regularity assumptions, the term $2\partial_x(X^{\vee}(u^Q + X^{\vee_3})) + \partial_x((u^Q + X^{\vee_3})^2)$ is well defined and the only problematic term in (33) is $\partial_x(u^Q X)$. Using the paracontrolled structure of u^Q, we can make sense of $\partial_x(u^Q X)$ as a bounded operator provided that $Q \circ X \in C\mathscr{C}^{2\alpha-1}$ is given. In other words, the right hand side of (33) is well defined for paracontrolled distributions.

Next, we should specify how to choose Q and which form u' will take for the solution u^Q. We have formally

$$\begin{aligned}\mathscr{L}u^Q &= \mathscr{L}X^{\vee\vee} + 4\mathscr{L}X^{\vee_4} + 2\partial_x(u^Q X) + 2\partial_x(X^{\vee}(u^Q + 2X^{\vee_3})) + \partial_x((u^Q + 2X^{\vee_3})^2)\\ &= 4\partial_x(X^{\vee_3} X) + 2\partial_x(u^Q X) + C\mathscr{C}^{2\alpha-2} = 4X^{\vee_3} \prec \partial_x X + 2u^Q \prec \partial_x X + C\mathscr{C}^{2\alpha-2},\end{aligned}$$

where we assumed that not only $\mathscr{L}X^{\vee_4} \in C\mathscr{C}^{\alpha-2}$, but that $\partial_x(X^{\vee_3} \circ X) \in C\mathscr{C}^{2\alpha-1}$ (which implies $\mathscr{L}X^{\vee_4} \in C\mathscr{C}^{\alpha-2}$, but also the stronger statement $\mathscr{L}X^{\vee_4} - X^{\vee_3} \prec \partial_x X \in C\mathscr{C}^{2\alpha-2}$). By Theorem 32, this shows that u^Q is paracontrolled by $J(\partial_x X)$, in other words we should set $Q = J(\partial_x X)$. The derivative u' of the solution u^Q will then be given by $u' = 4X^{\vee_3} + 2u^Q$.

Unlike for the parabolic Anderson model, here we do not need to introduce a renormalization. This is due to the fact that we differentiate after taking the square: to construct u^2, we would have to subtract an infinite constant and formally consider $u^{\diamond 2} = u^2 - \infty$. But then

$$\partial_x u^{\diamond 2} = \partial_x(u^2 - \infty) = \partial_x u^2.$$

So we obtain the following set of driving data for Burgers equation:

Definition 43. *(Rough distribution) Let $\alpha \in (1/3, 1/2)$ and let*

$$\mathcal{X}_{\text{rbe}} \subseteq C\mathscr{C}^{\alpha-1} \times C\mathscr{C}^{2\alpha-1} \times \mathscr{L}^{\alpha} \times \mathscr{L}^{2\alpha} \times C\mathscr{C}^{2\alpha-1} \times C\mathscr{C}^{2\alpha-1}$$

be the closure of the image of the map $\Theta_{\text{rbe}}: C(\mathbb{R}_+, C^{\infty}(\mathbb{T})) \to \mathcal{X}_{\text{rbe}}$ given by

$$\Theta_{\text{rbe}}(\theta) = (X(\theta), X^{\vee}(\theta), X^{\vee_3}(\theta), X^{\vee\vee}(\theta), (X^{\vee_3} \circ X)(\theta), (Q \circ X)(\theta)), \tag{34}$$

where

$$\begin{array}{rcl} X(\theta) &=& J(\partial_x \theta),\\ X^{\vee}(\theta) &=& B(X(\theta), X(\theta)),\\ X^{\vee_3}(\theta) &=& B(X^{\vee}(\theta), X(\theta)),\\ X^{\vee\vee}(\theta) &=& B(X^{\vee}(\theta), X^{\vee}(\theta)),\\ Q(\theta) &=& J(\partial_x X(\theta)). \end{array} \tag{35}$$

We will call $\Theta_{\text{rbe}}(\theta)$ the RBE-enhancement of the driving distribution θ. For $T > 0$ we define $\mathcal{X}_{\text{rbe}}(T) = \mathcal{X}_{\text{rbe}}|_{[0,T]}$ and we write $\|\mathbb{X}\|_{\mathcal{X}_{\text{rbe}}(T)}$ for the norm of $\mathbb{X}$ in the Banach space $C_T\mathscr{C}^{\alpha-1} \times C_T\mathscr{C}^{2\alpha-1} \times \mathscr{L}_T^{\alpha} \times \mathscr{L}_T^{2\alpha} \times C_T\mathscr{C}^{2\alpha-1} \times C_T\mathscr{C}^{2\alpha-1}$. Moreover, we define the distance $d_{\mathcal{X}_{\text{rbe}}(T)}(\mathbb{X}, \tilde{\mathbb{X}}) = \|\mathbb{X} - \tilde{\mathbb{X}}\|_{\mathcal{X}_{\text{rbe}}(T)}$.

Naturally, the acronym rbe stands for *rough Burgers equation*. For every $\mathbb{X} \in \mathcal{X}_{\text{rbe}}$, there is an associated space of paracontrolled distributions:

Definition 44. *Let $\mathbb{X} \in \mathcal{X}_{\text{rbe}}$ and let $\beta \in (0, \alpha]$ be such that $2\alpha + \beta > 1$. Then the space of paracontrolled distributions $\mathscr{D}^{\beta}(\mathbb{X})$ is defined as the set of all $(u, u') \in C\mathscr{C}^{\alpha-1} \times \mathscr{L}^{\beta}$ with*

$$u = X + X^{\vee} + 2X^{\vee_3} + u' \prec Q + u^{\sharp},$$

where $u^\sharp \in \mathscr{L}^{\alpha+\beta}$. *For* $T > 0$ *we define*

$$\|u\|_{\mathscr{D}_T^\beta} = \|u'\|_{\mathscr{L}_T^\beta} + \|u^\sharp\|_{C_T\mathscr{C}^{\alpha+\beta}}.$$

If $\tilde{\mathbb{X}} \in \mathcal{X}_{\mathrm{rbe}}$ *and* $(\tilde{u}, \tilde{u}') \in \mathscr{D}^\beta(\tilde{\mathbb{X}})$, *then we also write*

$$d_{\mathscr{D}_T^\beta}(u, \tilde{u}) = \|u' - \tilde{u}'\|_{\mathscr{L}_T^\beta} + \|u^\sharp - \tilde{u}^\sharp\|_{C_T\mathscr{C}_T^{\alpha+\beta}}.$$

We now have everything in place to solve the rough Burgers equation driven by some $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$.

Theorem 45. *Let* $\alpha \in (1/3, 1/2)$. *Let* $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$, *write* $\partial_x \theta = \mathscr{L} X$, *and let* $u_0 \in \mathscr{C}^{2\alpha}$. *Then there exists a unique solution* $u \in \mathscr{D}^\alpha(\mathbb{X})$ *to the equation*

$$\mathscr{L} u = \partial_x u^2 + \partial_x \theta, \qquad u(0) = u_0, \tag{36}$$

up to the (possibly finite) explosion time $\tau = \tau(u) = \inf\{t \geqslant 0 \colon \|u\|_{\mathscr{D}_t^\alpha} = \infty\} > 0$.

Moreover, u *depends on* $(u_0, \mathbb{X}) \in \mathscr{C}^{2\alpha} \times \mathcal{X}_{\mathrm{rbe}}$ *in a locally Lipschitz continuous way: if* $M, T > 0$ *are such that for all* $(u_0, \mathbb{X})$ *with* $\|u_0\|_{2\alpha} \vee \|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)} \leqslant M$, *the solution* u *to the equation driven by* $(u_0, \mathbb{X})$ *satisfies* $\tau(u) > T$, *and if* $(\tilde{u}_0, \tilde{\mathbb{X}})$ *is another set of data bounded in the above sense by* M, *then there exists* $C(M) > 0$ *for which*

$$d_{\mathscr{D}_T^\alpha}(u, \tilde{u}) \leqslant C(M)\big(\|u_0 - \tilde{u}_0\|_{2\alpha} + d_{\mathcal{X}_{\mathrm{rbe}}(T)}(\mathbb{X}, \tilde{\mathbb{X}})\big).$$

Proof. By definition of the term $\partial_x u^2$, the distribution $u \in \mathscr{D}^\alpha(\mathbb{X})$ solves (36) if and only if $u^Q = u - X - X^{\mathsf{V}} - 2X^{\mathsf{V}\!\!\!\mathsf{V}}$ solves

$$\mathscr{L} u^Q = \mathscr{L} X^{\mathsf{VV}} + 4\partial_x(X^{\mathsf{V}\!\!\!\mathsf{V}} X) + 2\partial_x(u^Q X) + 2\partial_x(X^{\mathsf{V}}(u^Q + 2X^{\mathsf{V}\!\!\!\mathsf{V}})) + \partial_x((u^Q + 2X^{\mathsf{V}\!\!\!\mathsf{V}})^2)$$

with initial condition $u^Q(0) = u_0$. This equation is structurally very similar to the parabolic Anderson model (10) and can be solved using the same arguments, which we do not reproduce here. □

Of course, to be of any use we still have to show that if ξ is the space-time white noise, then there is almost surely an element of $\mathbb{X}$ associated to $\partial_x \xi$. While for the parabolic Anderson model we needed to construct only one term, here we have to construct five terms: X^{V}, $X^{\mathsf{V}\!\!\!\mathsf{V}}$, X^{VV}, $X^{\mathsf{V}\!\!\!\mathsf{V}} \circ X$, $Q \circ X$. This construction will be included in [GP14], alternatively we may simply differentiate the extended data which Hairer constructed for the KPZ equation in Chapter 5 of [Hai13].

The same approach allows us to solve the KPZ equation $\mathscr{L} h = (\partial_x h)^{\diamond 2} + \xi$, and if we are careful how to interpret the product $w \diamond \xi$, then also the linear heat equation $\mathscr{L} w = w \diamond \xi$. In both cases, the solution will depend continuously on some suitably extended data that is constructed from ξ in a similar way as described in Definition 43. Moreover, the formal links between the three equations can be made rigorous. These results will be included in [GP14].

Bibliography

[BCD11] Hajer Bahouri, Jean-Yves Chemin and Raphael Danchin. *Fourier analysis and nonlinear partial differential equations*. Springer, 2011.

[BG97] Lorenzo Bertini and Giambattista Giacomin. Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.*, 183(3):571–607, 1997.

[BG13] G. Bal and Y. Gu. Limiting models for equations with large random potential; a review. Submitted, 2013.

[Bon81] 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.

[FV10] Peter Friz and Nicolas Victoir. *Multidimensional stochastic processes as rough paths. Theory and applications*. Cambridge University Press, 2010.

[GIP13] Massimiliano Gubinelli, Peter Imkeller and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *ArXiv preprint arXiv:1210.2684*, , 2013.

[GJ10] Patricia Goncalves and Milton Jara. Universality of KPZ equation. *ArXiv:1003.4478*, , mar 2010. ArXiv: 1003.4478.

[GJ13] M. Gubinelli and M. Jara. Regularization by noise and stochastic burgers equations. *Stochastic Partial Differential Equations: Analysis and Computations*, 1(2):325–350, jun 2013.

[GP14] Massimiliano Gubinelli and Nicolas Perkowski. KPZ reloaded. , , 2014. In preparation.

[Gub04] Massimiliano Gubinelli. Controlling rough paths. *J. Funct. Anal.*, 216(1):86–140, nov 2004.

[Gub10] Massimiliano Gubinelli. Ramification of rough paths. *J. Differential Equations*, 248(4):693–721, 2010.

[Hai13] Martin Hairer. Solving the KPZ equation. *Ann. Math.*, 178(2):559–664, 2013.

[Hai14] Martin Hairer. A theory of regularity structures. *Invent. math.*, , 2014. DOI 10.1007/s00222-014-0505-4.

[HPP13] Martin Hairer, Etienne Pardoux and Andrey Piatnitski. Random homogenisation of a highly oscillatory singular potential. *Stochastic Partial Differential Equations: Analysis and Computations*, 1(4):571–605, dec 2013.

[Hu02] Yaozhong Hu. Chaos expansion of heat equations with white noise potentials. *Potential Anal.*, 16(1):45–66, 2002.

[Jan97] Svante Janson. *Gaussian Hilbert spaces*, volume 129 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1997.

[KPZ86] Mehran Kardar, Giorgio Parisi and Yi-Cheng Zhang. Dynamic scaling of growing interfaces. *Physical Review Letters*, 56(9):889–892, 1986.

[LCL07] Terry J. Lyons, Michael Caruana and Thierry Lévy. *Differential equations driven by rough paths*, volume 1908 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.

[LQ02] Terry Lyons and Zhongmin Qian. *System control and rough paths*. Oxford University Press, 2002.

[Lyo98] Terry J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoam.*, 14(2):215–310, 1998.

[ST87] H.-J. Schmeisser and H. Triebel. *Topics in Fourier analysis and function spaces*, volume 42 of *Mathematik und ihre Anwendungen in Physik und Technik [Mathematics and its Applications in Physics and Technology]*. Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig, 1987.

[You36] Laurence C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.*, 67(1):251–282, 1936.