Paracontrolled distributions and singular PDEs *

Massimiliano Gubinelli†
CEREMADE & CNRS UMR 7534
Université Paris-Dauphine
and Institut Universitaire de France
gubinelli@ceremade.dauphine.fr

Peter Imkeller
Institut für Mathematik
Humboldt-Universität zu Berlin
imkeller@math.hu-berlin.de

Nicolas Perkowski‡
CEREMADE & CNRS UMR 7534
Université Paris-Dauphine
perkowski@ceremade.dauphine.fr

July 24, 2014

Abstract

We introduce an approach to study certain singular PDEs which is based on techniques from paradifferential calculus and on ideas from the theory of controlled rough paths. We illustrate its applicability on some model problems like differential equations driven by fractional Brownian motion, a fractional Burgers type SPDE driven by space-time white noise, and a non-linear version of the parabolic Anderson model with a white noise potential.

Keywords: Rough paths, Paraproducts, Besov spaces, Stochastic partial differential equations, Renormalization

Contents

* *Note:* This document has been produced using the GNU $\text{T}_\text{E}\text{X}_{\text{MACS}}$ text editor (see http://www.texmacs.org)

†M.G. is supported by a Junior fellowship of the Institut Universitaire de France (IUF) and by the ANR Project ECRU (ANR-09-BLAN-0114-01)

‡N.P. is supported by the Fondation Sciences Mathématiques de Paris (FSMP) and by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-LABX-0098).

1 Introduction

In this paper we introduce the notion of *paracontrolled distribution* and show how to use it to give a meaning to and solve partial differential equations involving non-linear operations on generalized functions. More precisely, we combine the idea of *controlled paths*, introduced in [Gub04], with the *paraproduct* introduced by Bony [Bon81] and the related paradifferential calculus, in order to develop a non-linear theory for a certain class of distributions.

The approach presented here works for generalized functions defined on an index set of arbitrary dimension and constitutes a flexible and lightweight generalization of Lyons' rough path theory [Lyo98]. In particular it allows to handle problems involving singular stochastic PDEs which were substantially out of reach with previously known methods.

In order to set the stage for our analysis let us list some of the problems which are amenable to be analyzed in the paracontrolled framework:

1. The *rough* differential equation (RDE) driven by a d–dimensional Gaussian process X:

$$\partial_t u(t) = F(u(t))\partial_t X(t),$$

where $F\colon \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^n)$ is a smooth vector-field . Typically, X will be a Brownian motion or a fractional Brownian motion with Hurst exponent $H \in (0,1)$. The paracontrolled analysis works up to $H > 1/3$. While we do not have any substantial new results for this problem, it is a useful pedagogical example on which we can easily describe our approach.

2. Generalizations of Hairer's Burgers-like SPDE (BURGERS):

$$Lu = G(u)\partial_x u + \xi.$$

Here $u\colon \mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}^n$, where $\mathbb{T} = (\mathbb{R}/2\pi\mathbb{Z})$ denotes the torus, $L = \partial_t + (-\Delta)^\sigma$, where $-(-\Delta)^\sigma$ is the fractional Laplacian with periodic boundary conditions and we will take $\sigma > 5/6$, and ξ is a space-time white noise with values in $\mathbb{R}^n$. Moreover, $G\colon \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is a smooth field of linear transformations.

3. A non-linear generalization of the parabolic Anderson model (PAM):

$$Lu = F(u) \diamond \xi,$$

where $u\colon \mathbb{R}_+ \times \mathbb{T}^2 \to \mathbb{R}$, $L = \partial_t - \Delta$ is the parabolic operator corresponding to the heat equation, and where ξ is a random potential which is sampled according to the law of the white noise on $\mathbb{T}^2$ and is therefore independent of the time variable. We allow for a general smooth function $F\colon \mathbb{R} \to \mathbb{R}$, the linear case $F(u) = u$ corresponding to the standard parabolic Anderson model. The symbol $\diamond$ stands for a renormalized product which is necessary to have a well defined problem.

4. The one-dimensional periodic Kardar–Parisi–Zhang equation (KPZ):
$$Lh = \text{“}(\partial_x h)^2\text{”} + \xi,$$
where $u\colon \mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}$, $L = \partial_t - \Delta$, and where ξ is a space-time white noise. Here “$(\partial_x h)^2$” denotes the necessity of an additive renormalization in the definition of the square of the distribution $\partial_x h$.

5. The three-dimensional, periodic, stochastic quantization equation for the $(\phi)^4_3$ euclidean quantum field (SQ):
$$L\phi = \text{“}\frac{\lambda}{4!}(\phi)^3\text{”} + \xi,$$
where $\phi\colon \mathbb{R}_+ \times \mathbb{T}^3 \to \mathbb{R}$, $L = \partial_t - \Delta$, ξ is a space-time white noise, and where “$(\phi)^3$” denotes a suitable renormalization of a cubic polynomial of ϕ and λ is the coupling constant of the scalar theory.

In this paper we will consider in detail the three cases RDE, BURGERS, PAM. In all cases we will exhibit a space of paracontrolled distributions where the equations are well posed (in a suitable sense), and admit a global solution which is unique. The three-dimensional stochastic quantization equation SQ is studied by R. Catellier and K. Chouk in [CC13] by applying the paracontrolled technique. The paracontrolled analysis of KPZ will be presented elsewhere [GP14].

The kind of results which will be obtained below can be exemplified by the following statement for RDEs. Below $\mathscr{C}^\alpha = B^\alpha_{\infty,\infty}$ stands for the Hölder-Besov space of index α on $\mathbb{R}$. Given two distributions $f \in \mathscr{C}^\alpha$ and $g \in \mathscr{C}^\beta$ with $\alpha + \beta > 0$ we can always consider a certain distribution $f \circ g$ which is obtained via a bilinear operation of f, g and which belongs to $\mathscr{C}^{\alpha+\beta}$.

Theorem 1.1. *Let $\xi : [0,1] \to \mathbb{R}^n$ be a continuous function and $F : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^d)$ be a family of smooth vector-fields. Let $u : [0,1] \to \mathbb{R}^d$ be a solution of the Cauchy problem*
$$\partial_t u(t) = F(u(t))\xi(t), \qquad u(0) = u_0,$$
where $u_0 \in \mathbb{R}^d$. Let ϑ be a solution to $\partial_t \vartheta = \xi$ and let $R\xi = (\xi, \vartheta \circ \xi)$. Then for all $\alpha \in (1/3, 1)$ there exists a continuous map $\Psi : \mathbb{R}^d \times \mathscr{C}^{\alpha-1} \times \mathscr{C}^{2\alpha-1} \to \mathscr{C}^\alpha$ such that $u = \Psi(u_0, R\xi)$ for all $\xi \in C([0,1]; \mathbb{R}^d)$.

In particular, this theorem provides a natural way of extending the solution map to data ξ which are merely distributions in $\mathscr{C}^{\alpha-1}$. It suffices to approximate ξ by a sequence of smooth functions (ξ^n) converging to ξ in $\mathscr{C}^{\alpha-1}$, and to prove that the “lifted” sequence $(R\xi^n)$ converges to some limit in $\mathscr{C}^{\alpha-1} \times \mathscr{C}^{2\alpha-1}$. The uniqueness of this limit is not guaranteed however, and each possible limit will give rise to a different notion of solution to the RDE, just like in standard rough path theory.

The space $\mathcal{X}$ obtained by taking the closure in $\mathscr{C}^{\alpha-1} \times \mathscr{C}^{2\alpha-1}$ of the set of all elements of the form $R\xi$ for smooth ξ replaces the space of (geometric) rough paths, and the above theorem is a partial restatement of Lyons' continuity result: namely that the (Itô) solution map Ψ, going from data to solution of the differential equation, is a continuous map from the rough path space $\mathcal{X}$ to $\mathscr{C}^\alpha$. The space $\mathcal{X}$ is fibered over $\mathscr{C}^{\alpha-1}$. It allows us to equip the driving distribution with enough information to control the continuity of the solution map to our RDE problem – and as we will see below, also the continuity of the solution maps to suitable PDEs. In various contexts the space $\mathcal{X}$ can take different forms, and in general it does not seem to have the rich geometrical and algebraic structure of standard rough paths.

The verification that suitable approximations (ξ^n) are such that their lifts $(R\xi^n)$ converge in $\mathscr{C}^{\alpha-1} \times \mathscr{C}^{2\alpha-1}$ depends on the particular form of ξ. In the case of ξ being a Gaussian stochastic process (like in all our examples above), this verification is the result of almost sure convergence results for elements in a fixed chaos of an underlying Gaussian process, and the proofs rely on elementary arguments on Gaussian random variables.

Even in the case of RDEs, the paracontrolled analysis leads to some interesting insights. For example, we have that a more general equation of the form

$$\partial_t u(t) = F(u(t))\xi(t) + F'(u(t))F(u(t))\eta(t), \qquad u(0) = u_0,$$

where $\eta \in C([0,1]; \mathbb{R}^n \times \mathbb{R}^n)$, has a solution map which depends continuously on $(\xi, \vartheta \circ \xi + \eta) \in \mathscr{C}^{\alpha-1} \times \mathscr{C}^{2\alpha-1}$. The remarkable fact here is that the solution map depends only on the combination $\vartheta \circ \xi + \eta$ and not on each term separately. Such structural features of the solution map, which can be easily seen using the paracontrolled analysis, are very important in situations where renomalizations are needed, as for example in the PAM model. In the RDE context we can simply remark that setting $\eta = -\vartheta \circ \xi$, the solution map becomes a continuous function of $\xi \in \mathscr{C}^{\alpha-1}$, without any further requirement on the bilinear object $\vartheta \circ \xi$. Thus, the equation

$$\partial_t u(t) = F(u(t))\xi(t) - F'(u(t))F(u(t))(\vartheta \circ \xi)(t), \qquad u(0) = u_0,$$

can be readily extended to any $\xi \in \mathscr{C}^{\alpha-1}$ by continuity. In that sense, this equation should be interpreted as a deterministic version of an Itô stochastic differential equation where, at the price of a modification of standard rules of calculus, we are able to solve more general problems than in the Stratonovich setting.

We remark that, even if only quite implicitly, paraproducts have been already exploited in the rough path context in the work of Unterberger on the renormalization of rough paths [Unt10a, Unt10b], where it is referred to as "Fourier normal-ordering", and in the related work of Nualart and Tindel [NT11].

In this paper we construct weak solutions for the SPDEs under consideration. For an approach using mild solutions see [Per14]. See also [GIP14], where we use the decomposition of continuous functions in a certain Fourier series and similar ideas as developed below, in order to give a new and relatively elementary approach to rough path integration.

Relevant literature. Before going into the details, let us describe the context of our study. Consider for example the RDE problem above. Schwartz' theory of distributions gives a robust framework for defining linear operations on irregular generalized functions. But when trying to handle non-linear operations, we quickly run into problems. For example, in Schwartz' theory, it is not possible to define the product $F(u)\partial_t X(t)$ in the case where X is the sample path of a Brownian motion. The standard analysis of this difficulty goes as follows: X is an α–Hölder continuous process for any $\alpha < 1/2$, but not better. The solution u has to have the same regularity, which is transferred to $F(u)$ if F is smooth. In this situation, the product $F(u)\partial_t X$ corresponds to the product of an α–Hölder continuous function with the distribution $\partial_t X$ which is of order $\alpha - 1$. A well known result of analysis (see Section 2.1 below) tells us that a necessary condition for this product to be well defined is that the sum of the orders is positive, that is

$2\alpha - 1 > 0$, which is barely violated in the Brownian setting. This is the classical problem which motivated Itô's theory of stochastic integrals.

Itô's integral has however quite stringent structural requirements: an "arrow of time" (i.e. a filtration and adapted integrands), a probability measure (it is defined as L^2–limit), and L^2–orthogonal increments of the integrator (the integrator needs to be a (semi-) martingale).

If one or several of these assumptions are violated, then Lyons' rough path integral [Lyo98, LQ02, LCL07, FV10] can be an effective alternative. For example, it allows to construct pathwise integrals for, among other processes, fractional Brownian motion, which is not a semimartingale.

In the last years, several other works applied rough path techniques to SPDEs. But they all relied on special features of the problem at hand in order to apply the integration theory provided by the rough path machinery.

A first series of works attempts to deal with "time"-like irregularities by adapting the standard rough path approach:

– Deya, Gubinelli, Lejay, and Tindel [GLT06, Gub12, DGT12] deal with SPDEs of the form

$$Lu(t, x) = \sigma(u(t, x))\eta(t, x),$$

where $x \in \mathbb{T}$, $L = \partial_t - \Delta$, the noise η is a space-time Gaussian distribution (for example white in time and colored in space), and σ is some non-linear coefficient. They interpret this as an evolution equation (in time), taking values in a space of functions (with respect to the space variable). They extend the rough path machinery to handle the convolution integrals that appear when applying the heat flow to the noise.

– Friz, Caruana, Diehl, and Oberhauser [CF09, CFO11, FO11, DF12] deal with fully non-linear stochastic PDEs with a special structure. Among others, of the form

$$\partial_t u(t, x) = F(u, \partial_x u, \partial_x^2 u) + \sigma(t, x)\partial_x u(t, x)\eta(t),$$

where the spatial index x can be multidimensional, but the noise η only depends on time. Such an SPDE can be reinterpreted as a standard PDE with random coefficients via a change of variables involving the flow of the stochastic characteristics associated to σ. This flow is handled using usual rough path results for RDEs.

– Teichmann [Tei11] studies semilinear SPDEs of the form

$$(\partial_t - A)u(t, x) = \sigma(u)(t, x)\eta(t, x),$$

where A is a suitable linear operator, in general unbounded, and σ is a general non-linear operation on the unknown u which however should satisfy some restrictive conditions. The SPDE is transformed into an SDE with bounded coefficients by applying a transformation based on the group generated by A on a suitable space.

The "arrow of time" condition of Itô's integral is typically violated if the index is a spatial variable and not a temporal variable. Another series of works applied rough path integrals to deal with situations involving irregularities in the "space" directions:

– Bessaih, Gubinelli, and Russo [BGR05] and Brzezniak, Gubinelli, and Neklyudov [BGN13] consider the vortex filament equation which describes the (approximate) motion of a closed vortex line $x(t, \cdot) \in C(\mathbb{T}, \mathbb{R}^3)$ in an incompressible three-dimensional fluid:

$$\partial_t x(t, \sigma) = u^{x(t,\cdot)}(x(t, \sigma)), \qquad u^{x(t,\cdot)}(y) = \int_{\mathbb{T}} K(y - x(t, \sigma))\partial_\sigma x(t, \sigma)\mathrm{d}\sigma,$$

where $K : \mathbb{R}^3 \to \mathcal{L}(\mathbb{R}^3, \mathbb{R}^3)$ is a smooth antisymmetric field of linear transformations of $\mathbb{R}^3$. In the modeling of turbulence it is interesting to study this equation with initial condition $x(0, \cdot)$ sampled according to the law of the three-dimensional Brownian bridge. In this case, the regularity of $x(t, \sigma)$ with respect to σ is no better than Brownian for any positive time, and thus the integral in the definition of the velocity field $u^{x(t,\cdot)}$ is not well defined. Rough path theory allows to make sense of this integral and then of the equation.

– Hairer, Maas, and Weber [Hai11, HW13, Hai13, HMW12] build on the insight of Hairer that rough path theory allows to make sense of SPDEs which are ill-defined in standard function spaces due to spatial irregularities. Hairer and Weber [HW13] extend the BURGERS type SPDE that we presented above to the case of multiplicative noise. Hairer, Maas, and Weber [HMW12] study approximations to this equation, where they discretize the spatial derivative as $\partial_x u(t,x) \simeq 1/\varepsilon(u(t,x+\varepsilon) - u(t,x))$. They show that in the limit $\varepsilon \to 0$, the approximation may introduce a Stratonovich type correction term to the equation. Finally, Hairer [Hai13] uses this approach to define and solve for the first time the Kardar–Parisi–Zhang (KPZ) equation, an SPDE of one spatial index variable that describes the random growth of an interface. The KPZ equation was introduced by Kardar, Parisi, and Zhang [KPZ86], and prior to Hairer's work it could only be solved by applying a spatial transform (the Cole-Hopf transform) which had the effect of linearizing the equation.

Alternative approaches. In all the papers cited above, the intrinsic one-dimensional nature of rough path theory severely limits possible improvements or applications to other contexts. To the best of our knowledge, the first attempt to remove these limitations is the still unpublished work by Chouk and Gubinelli [CG13], extending rough path theory to handle (fractional) Brownian sheets (Gaussian two-parameter stochastic processes akin to (fractional) Brownian motion).

In the recent paper [Hai14], Hairer has introduced a theory of *regularity structures* with the aim of giving a more general and versatile notion of regularity. Hairer's theory is also inspired by the theory of controlled rough paths, and it can also be considered a generalization of it to functions of a multidimensional index variable. The crucial insight is that the regularity of the solution to an equation driven by – say – Gaussian space-time white noise should not be described in the classical way. Usually we say that a function is smooth if it can be approximated around every point by a polynomial of a given degree (the Taylor polynomial). Since the solution to an SPDE does not look like a polynomial at all, this is not the correct way of describing its regularity. We rather expect that the solution locally looks like the driving noise (more precisely like the noise convoluted with the Green kernel of the linear part of the equation; so in the case of RDEs the time integral of the white noise, i.e. the Brownian motion). Therefore, in Hairer's theory a function is called smooth if it can locally be well approximated by this convolution (and higher order terms depending on the noise). Hairer's notion of smoothness induces a natural topology in which the solutions to semilinear SPDEs depend continuously on the driving signal. This approach is very general, and allows to handle more complicated problems than the ones we are currently able to treat in the paracontrolled approach. If there is a merit in our 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 harmonic analysis and rough paths.

Plan of the paper. Section 2 develops the calculus of paracontrolled distributions. In Section 3 we solve ordinary differential equations driven by suitable Gaussian processes such as the fractional Brownian motion with Hurst index $H > 1/3$. In Section 4 we solve a fractional

Burgers type equation driven by white noise, and in Section 5 we study a non-linear version of the parabolic Anderson model. In Appendix A we recall the main concepts of Littlewood-Paley theory and of Bony's paraproduct, and Appendix B contains a commutator estimate between paraproduct and time integral. We stress the fact that this paper is mostly self-contained, and in particular we will not need any results from rough path theory and just basic elements of the theory of Besov spaces.

Acknowledgments. The main part of the research was carried out while N.P. was employed by Humboldt-Universität zu Berlin. During an Oberwolfach workshop in the summer of 2012, M. Hairer discussed with one of us (M.G.) his approach to extend rough path theory and M.G. would like to thank M. Hairer for suggesting the application to the two-dimensional non-linear parabolic Anderson model discussed in this paper.

Notation and conventions. Throughout the paper, we use the notation $a \lesssim b$ if there exists a constant $c > 0$, independent of the variables under consideration, such that $a \leqslant c \cdot b$, and 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 $a(x) \lesssim_x b(x)$. For index variables i and j of Littlewood-Paley decompositions (see below) we write $i \lesssim j$ if $2^i \lesssim 2^j$, so in other words if there exists $N \in \mathbb{N}$, independent of j, such that $i \leqslant j + N$, and we write $i \sim j$ if $i \lesssim j$ and $j \lesssim i$.

An *annulus* is a set of the form $\mathscr{A} = \{x \in \mathbb{R}^d : a \leqslant |x| \leqslant b\}$ for some $0 < a < b$. A *ball* is a set of the form $\mathscr{B} = \{x \in \mathbb{R}^d : |x| \leqslant b\}$. $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ denotes the torus.

The Hölder-Besov space $B^{\alpha}_{\infty,\infty}(\mathbb{R}^d, \mathbb{R}^n)$ for $\alpha \in \mathbb{R}$ will be denoted by $\mathscr{C}^{\alpha}$, equipped with the norm $\|\cdot\|_{\alpha} = \|\cdot\|_{B^{\alpha}_{\infty,\infty}}$. The local space $\mathscr{C}^{\alpha}_{\mathrm{loc}}$ consists of all u which satisfy $\varphi u \in \mathscr{C}^{\alpha}$ for every infinitely differentiable φ of compact support. Given two Banach spaces X, Y we denote by $\mathcal{L}(X, Y)$ the Banach space of linear maps from X to Y, endowed with the operator norm $\|\cdot\|_{\mathcal{L}(X,Y)}$. More generally, given $k \in \mathbb{N}$ and Banach spaces $X_1, \ldots, X_k$, we write $\mathcal{L}^k(X_1 \times \ldots \times X_k, Y)$ for the space of k-linear maps from $X_1 \times \ldots \times X_k$ to Y, and $\|\cdot\|_{\mathcal{L}^k(X_1 \times \ldots \times X_k, Y)}$ for the operator norm. We denote by $C(X, Y)$ the Banach space of continuous maps from X to Y, endowed with the supremum norm $\|\cdot\|_{C(X,Y)}$. We write $C_T Y = C([0, T], Y)$ for the space of continuous maps from $[0, T]$ to Y, equipped with the supremum norm $\|\cdot\|_{C_T Y}$. If $\alpha \in (0, 1)$, then we also define $C^{\alpha}_T Y$ as the space of α-Hölder continuous functions from $[0, T]$ to Y, endowed with the seminorm

$$\|f\|_{C^{\alpha}_T Y} = \sup_{0 \leqslant s < t \leqslant T} \frac{\|f(t) - f(s)\|_Y}{|t - s|^{\alpha}}.$$

If f is a map from $A \subset \mathbb{R}$ to the linear space Y, then we write $f_{s,t} = f(t) - f(s)$, so that $\|f\|_{C^{\alpha}_T Y} = \sup_{0 \leqslant s < t \leqslant T} \|f_{s,t}\|_Y / |t - s|^{\alpha}$. For $f \in L^p(\mathbb{T})$ we write $\|f(x)\|^p_{L^p_x(\mathbb{T})} = \int_{\mathbb{T}} |f(x)|^p \mathrm{d}x$.

For a multi-index $\mu = (\mu_1, \ldots, \mu_d) \in \mathbb{N}^d$ we write $|\mu| = \mu_1 + \ldots + \mu_d$ and $\partial^{\mu} = \partial^{|\mu|} / \partial^{\mu_1}_{x_1} \cdots \partial^{\mu_d}_{x_d}$. $\mathrm{D}F$ or F' denote the total derivative of F. For $k \in \mathbb{N}$ we denote by $\mathrm{D}^k F$ the k-th order derivative of F. For $\alpha > 0$, $C^{\alpha}_b = C^{\alpha}_b(\mathbb{R}^d, \mathbb{R}^n)$ is the space of $\lfloor \alpha \rfloor$ times continuously differentiable functions, bounded with bounded partial derivatives, and with $(\alpha - \lfloor \alpha \rfloor)$–Hölder continuous partial derivatives of order $\lfloor \alpha \rfloor$, equipped with its usual norm $\|\cdot\|_{C^{\alpha}_b}$. We also write ∂_x for the partial derivative in direction x, and if $F \colon \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^n$, then we write $\mathrm{D}_x F(t, x)$ for its spatial derivative in the point $(t, x) \in \mathbb{R} \times \mathbb{R}^d$.

The space of real valued infinitely differentiable functions of compact support is denoted by $\mathscr{D}(\mathbb{R}^d)$ or $\mathscr{D}$. The space of Schwartz functions is denoted by $\mathscr{S}(\mathbb{R}^d)$ or $\mathscr{S}$. Its dual, the space of tempered distributions, is $\mathscr{S}'(\mathbb{R}^d)$ or $\mathscr{S}'$. If u is a vector of n tempered distributions on $\mathbb{R}^d$, then we write $u \in \mathscr{S}'(\mathbb{R}^d, \mathbb{R}^n)$. The Fourier transform is defined with the normalization

$$\mathscr{F} u(z) = \hat{u}(z) = \int_{\mathbb{R}^d} e^{-\iota \langle z, x \rangle} u(x) \mathrm{d}x,$$

so that the inverse Fourier transform is given by $\mathscr{F}^{-1}u(z) = (2\pi)^{-d}\mathscr{F}u(-z)$. If φ is a smooth function, such that φ and all its partial derivatives are at most of polynomial growth at infinity, then we define the Fourier multiplier $\varphi(\mathrm{D})$ by $\varphi(\mathrm{D})u = \mathscr{F}^{-1}(\varphi\mathscr{F}u)$ for any $u \in \mathscr{S}'$. More generally, we define $\varphi(\mathrm{D})u$ by this formula whenever the right hand side makes sense. The scaling operator Λ on $\mathscr{S}'$ is defined for $\lambda > 0$ by $\Lambda_\lambda u = u(\lambda\cdot)$.

Throughout the paper, (χ, ρ) will denote a dyadic partition of unity, and $(\Delta_j)_{j\geqslant -1}$ will denote the Littlewood-Paley blocks associated to this partition of unity, i.e. $\Delta_{-1} = \chi(\mathrm{D})$ and $\Delta_j = \rho(2^{-j}\mathrm{D})$ for $j \geqslant 0$. We will often write ρ_j, by which we mean χ if $j = -1$, and we mean $\rho(2^{-j}\cdot)$ if $j \geqslant 0$. We also use the notation $S_j = \sum_{i<j} \Delta_i$.

2 Paracontrolled calculus

2.1 Bony's paraproduct

Paraproducts are bilinear operations introduced by Bony [Bon81] in order to linearize a class of non-linear PDE problems. In this section we will introduce paraproducts to the extent of our needs. We will be using the Littlewood-Paley theory of Besov spaces. The reader can peruse Appendix A, where we summarize the basic elements of Besov space theory and Littlewood-Paley decompositions which will be needed in the remainder of the paper.

One of the simplest situations where paraproducts appear naturally is in the analysis of the product of two Besov distributions. In general, the product fg of two distributions $f \in \mathscr{C}^\alpha$ and $g \in \mathscr{C}^\beta$ is not well defined unless $\alpha + \beta > 0$. In terms of Littlewood–Paley blocks, the product fg can be (at least formally) decomposed as

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

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

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

We also introduce the notation

$$f \succcurlyeq g = f \succ g + f \circ g.$$

This decomposition behaves nicely with respect to Littlewood–Paley theory. Of course, it 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. Our choice of taking all (i, j) with $|i - j| \leqslant 1$ into the diagonal part corresponds to property iii. in the definition of dyadic partition of unity in Appendix A, where we assumed that $\mathrm{supp}(\rho(2^{-i}\cdot)) \cap \mathrm{supp}(\rho(2^{-j}\cdot)) = \emptyset$ for $|i - j| > 1$. This means that every term in the series

$$f \prec g = \sum_{j\geqslant -1}\sum_{i=-1}^{j-2} \Delta_i f \Delta_j g = \sum_{j\geqslant -1} S_{j-1} f \Delta_j g$$

has a Fourier transform which is supported in a suitable annulus, and of course the same holds true for $f \succ g$. On the other side, every term in the diagonal part $f \circ g$ has a Fourier transform that is supported in a ball. 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. In particular, if $\alpha > 0$ and $\beta \in \mathbb{R}$, then $(f, g) \mapsto f \prec g$ is a bounded bilinear operator from

$\mathscr{C}^\alpha \times \mathscr{C}^\beta$ to $\mathscr{C}^\beta$. Heuristically, $f \prec g$ behaves at large frequencies like g (and thus retains the same regularity), and f provides only a modulation of g at larger scales. The only difficulty in defining fg 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.

Lemma 2.1 (Paraproduct estimates, [Bon81])**.** *For any $\beta \in \mathbb{R}$ we have*

$$\|f \prec g\|_\beta \lesssim_\beta \|f\|_{L^\infty} \|g\|_\beta, \tag{1}$$

and for $\alpha < 0$ furthermore

$$\|f \prec g\|_{\alpha+\beta} \lesssim_{\alpha,\beta} \|f\|_\alpha \|g\|_\beta. \tag{2}$$

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

$$\|f \circ g\|_{\alpha+\beta} \lesssim_{\alpha,\beta} \|f\|_\alpha \|g\|_\beta. \tag{3}$$

Proof. Observe that 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 that for $f \in L^\infty$ we have

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

On the other side, if $\alpha < 0$ and $f \in \mathscr{C}^\alpha$, then

$$\|S_{j-1} f \Delta_j g\|_{L^\infty} \leqslant \sum_{i \leqslant j-2} \|\Delta_i f\|_{L^\infty} \|\Delta_j g\|_{L^\infty} \lesssim \|f\|_\alpha \|g\|_\beta \sum_{i \leqslant j-2} 2^{-i\alpha - j\beta} \lesssim \|f\|_\alpha \|g\|_\beta 2^{-j(\alpha+\beta)}.$$

By Lemma A.3, we thus obtain (1) and (2). To estimate $f \circ g$, observe that the term $u_j = \Delta_j f \sum_{i:|i-j| \leqslant 1} \Delta_i g$ has Fourier transform supported in a ball $2^j \mathscr{B}$, and that

$$\|u_j\|_{L^\infty} \lesssim \|\Delta_j f\|_{L^\infty} \sum_{i:|i-j| \leqslant 1} \|\Delta_i g\|_{L^\infty} \lesssim \|f\|_\alpha \|g\|_\beta 2^{-(\alpha+\beta)j}.$$

So if $\alpha + \beta > 0$, then we can apply the second part of Lemma A.3 to obtain that $f \circ g = \sum_{j \geqslant -1} u_j$ is an element of $\mathscr{C}^{\alpha+\beta}$ and that equation (3) holds. $\square$

A natural corollary is that the product fg of two elements $f \in \mathscr{C}^\alpha$ and $g \in \mathscr{C}^\beta$ is well defined as soon as $\alpha + \beta > 0$, and that it belongs to $\mathscr{C}^\gamma$, where $\gamma = \min\{\alpha, \beta, \alpha + \beta\}$.

2.2 Paracontrolled distributions and RDEs

Consider the RDE

$$\partial_t u = F(u) \xi, \qquad u(0) = u_0, \tag{4}$$

where $u_0 \in \mathbb{R}^d$, $u \colon \mathbb{R} \to \mathbb{R}^d$ is a continuous vector valued function, ∂_t is the time derivative, $\xi \colon \mathbb{R} \to \mathbb{R}^n$ is a vector valued distribution with values in $\mathscr{C}^{\alpha-1}$ for some $\alpha \in (1/3, 1)$, and $F \colon \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^d)$ is a family of vector fields on $\mathbb{R}^d$. A natural approach is to understand this equation as limit of the classical ODEs

$$\partial_t u^\varepsilon = F(u^\varepsilon) \xi^\varepsilon, \qquad u^\varepsilon(0) = u_0, \tag{5}$$

for a family of smooth approximations (ξ^ε) of ξ such that $\xi^\varepsilon \to \xi$ in $\mathscr{C}^{\alpha-1}$ as $\varepsilon \to 0$. In order to pass to the limit, we are looking for a priori estimates on u^ε which require only a control on the $\mathscr{C}^{\alpha-1}$ norm of ξ.

To avoid cumbersome notation, we will work at the level of equation (4) for smooth ξ, where it should be understood that our aim is to obtain a priori estimates for the solution, in order to safely pass to the limit and extend the solution map to a larger class of data. The natural regularity of u is $\mathscr{C}^\alpha$, since u should gain one derivative with respect to $F(u)\xi$, which will not behave better than ξ, and will therefore be in $\mathscr{C}^{\alpha-1}$.

We use the paraproduct decomposition to write the right hand side of (4) as a sum of the three terms

$$\underbrace{F(u) \prec \xi}_{\alpha-1} + \underbrace{F(u) \circ \xi}_{2\alpha-1} + \underbrace{F(u) \succ \xi}_{2\alpha-1} \tag{6}$$

(where the quantity indicated by the underbrace corresponds to the expected regularity of each term). Note however that unless $2\alpha - 1 > 0$, the resonant term $F(u) \circ \xi$ cannot be controlled using only the $\mathscr{C}^\alpha$–norm of u and the $\mathscr{C}^{\alpha-1}$–norm of ξ. If F is at least in C^2, we can use a paralinearization result (see Lemma 2.7 below) to rewrite this term as

$$F(u) \circ \xi = F'(u)(u \circ \xi) + \Pi_F(u, \xi), \tag{7}$$

where the remainder $\Pi_F(u, \xi)$ is well defined under the condition $3\alpha - 1 > 0$, provided that $u \in \mathscr{C}^\alpha$ and $\xi \in \mathscr{C}^{\alpha-1}$. In this case it belongs to $\mathscr{C}^{3\alpha-1}$. The difficulty is now localized in the linearized resonant product $u \circ \xi$. In order to control this term, we would like to exploit the fact that the function u is not a generic element of $\mathscr{C}^\alpha$ but that it has a specific structure, since its derivative $\partial_t u$ has to match the paraproduct decomposition given in (6). Thus, we postulate that the solution u is given by the following *paracontrolled ansatz:*

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

where $u^\vartheta, \vartheta \in \mathscr{C}^\alpha$ and the remainder $u^\sharp$ is in $\mathscr{C}^{2\alpha}$. This decomposition allows for a finer analysis of the resonant term $u \circ \xi$. Indeed, we have

$$u \circ \xi = (u^\vartheta \prec \vartheta) \circ \xi + u^\sharp \circ \xi = u^\vartheta(\vartheta \circ \xi) + C(u^\vartheta, \vartheta, \xi) + u^\sharp \circ \xi, \tag{8}$$

where the commutator is defined by $C(u^\vartheta, \vartheta, \xi) = (u^\vartheta \prec \vartheta) \circ \xi - u^\vartheta(\vartheta \circ \xi)$. Observe now that the term $u^\sharp \circ \xi$ does not pose any further problem, as it is bounded in $\mathscr{C}^{3\alpha-1}$. Moreover, we will show that the commutator is a bounded multilinear function of its arguments as long as the sum of their regularities is strictly positive, see Lemma 2.4 below. By assumption, we have $3\alpha - 1 > 0$, and therefore $C(u^\vartheta, \vartheta, \xi) \in \mathscr{C}^{3\alpha-1}$. The only problematic term which remains to be handled is thus $\vartheta \circ \xi$. Here we need to make the assumption that $\vartheta \circ \xi \in \mathscr{C}^{2\alpha-1}$ in order for the product $u^\vartheta(\vartheta \circ \xi)$ 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. Granting this, we have obtained that the right hand side of equation (4) is well defined and a continuous function of $(u, u^\vartheta, u^\sharp, \vartheta, \xi, \vartheta \circ \xi)$.

The paracontrolled ansatz and the Leibniz rule for the paraproduct now imply that (4) can be rewritten as

$$\partial_t u = \partial_t(u^\vartheta \prec \vartheta + u^\sharp) = \partial_t u^\vartheta \prec \vartheta + u^\vartheta \prec \partial_t \vartheta + \partial_t u^\sharp = F(u) \prec \xi + F(u) \circ \xi + F(u) \succ \xi.$$

If we choose ϑ such that $\partial_t \vartheta = \xi$ and we set $u^\vartheta = F(u)$, then we can use (7) and (8) to obtain the following equation for the remainder $u^\sharp$:

$$\begin{aligned}\partial_t u^\sharp = {} & F'(u)F(u)(\vartheta \circ \xi) + F(u) \succ \xi - (\partial_t F(u) \prec \vartheta) \\ & + F'(u)C(F(u), \vartheta, \xi) + F'(u)(u^\sharp \circ \xi) + \Pi_F(u, \xi).\end{aligned}$$

Together with the equation $u = F(u) \prec \vartheta + u^\sharp$, this completely describes the solution and allows us to obtain an a priori estimate on u in terms of $(u_0, \|\xi\|_{\alpha-1}, \|\vartheta \circ \xi\|_{2\alpha-1})$. With this estimate at hand, it is now easy to show that if $F \in C_b^3$, then u depends continuously on the data $(u_0, \xi, \vartheta \circ \xi)$, so that we can pass to the limit in (5) and make sense of the solution to (4) also for irregular $\xi \in \mathscr{C}^{\alpha-1}$ as long as $\alpha > 1/3$.

2.3 Commutator estimates and paralinearization

In this section we prove some lemmas which will allow us to perform algebraic computations with the paraproduct and the resonant term, thus justifying the analysis of the previous section.

Lemma 2.2 (see also Lemma 2.97 of [BCD11])**.** *Let $f \in \mathscr{C}^\alpha$ for $\alpha \in (0,1)$, and let $g \in L^\infty$. For any $j \geqslant -1$ we have*

$$\|[\Delta_j, f]g\|_{L^\infty} = \|\Delta_j(fg) - f\Delta_j g\|_{L^\infty} \lesssim 2^{-\alpha j}\|f\|_\alpha \|g\|_{L^\infty}.$$

This commutator lemma is easily proven by writing $\Delta_j = \rho_j(\mathrm{D})$ as a convolution operator, and using the embedding of $\mathscr{C}^\alpha$ in the space of Hölder continuous functions.

Lemma 2.3. *Assume that $\alpha \in (0,1)$ and $\beta \in \mathbb{R}$, and let $f \in \mathscr{C}^\alpha$ and $g \in \mathscr{C}^\beta$. Then*

$$\Delta_j(f \prec g) = f\Delta_j g + R_j(f,g),$$

for all $j \geqslant -1$, with a remainder $R_j(f,g)$ which satisfies $\|R_j(f,g)\|_{L^\infty} \lesssim 2^{-j(\alpha+\beta)}\|f\|_\alpha \|g\|_\beta$.

Proof. Note that $f \prec g = \sum_i f \prec \Delta_i g$, and that the Fourier transform of $f \prec \Delta_i g$ is supported in an annulus of the form $2^i \mathscr{A}$. Hence, we have $\Delta_j(f \prec \Delta_i g) \neq 0$ only if $j \sim i$, which leads to

$$\begin{aligned}\Delta_j(f \prec g) &= \sum_{i:i\sim j} \Delta_j(f \prec \Delta_i g) = \sum_{i:i\sim j} \Delta_j(f\Delta_i g) - \sum_{i:i\sim j} \Delta_j(f \succcurlyeq \Delta_i g) \\ &= \sum_{i:i\sim j} f\Delta_j\Delta_i g - \sum_{i:i\sim j} [\Delta_j, f]\Delta_i g - \sum_{i:i\sim j} \Delta_j(f \succcurlyeq \Delta_i g),\end{aligned}$$

where we recall that $[\Delta_j, f]\Delta_i g = \Delta_j(f\Delta_i g) - f\Delta_j\Delta_i g$ denotes the commutator. The sum over i with $i \sim j$ can be chosen to encompass enough terms so that $\Delta_j g = \sum_{i:i\sim j} \Delta_j \Delta_i g$, and therefore we conclude that

$$\|\Delta_j(f \prec g) - f\Delta_j g\|_{L^\infty} \leqslant \sum_{i:i\sim j} \|[\Delta_j, f]\Delta_i g\|_{L^\infty} - \sum_{i:i\sim j} \|\Delta_j(f \succcurlyeq \Delta_i g)\|_{L^\infty}.$$

We apply Lemma 2.2 to each term of the first sum, and the paraproduct estimates to each term of the second sum, to obtain

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

□

Using this result, it is easy to prove our basic commutator lemma.

Lemma 2.4. *Assume that $\alpha \in (0,1)$ and $\beta, \gamma \in \mathbb{R}$ are such that $\alpha+\beta+\gamma > 0$ and $\beta+\gamma < 0$. Then for smooth f, g, h, the trilinear operator*

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

allows for the bound

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

Thus, C can be uniquely extended to a bounded trilinear operator in $\mathcal{L}^3\left(\mathscr{C}^\alpha \times \mathscr{C}^\beta \times \mathscr{C}^\alpha, \mathscr{C}^{\alpha+\beta+\gamma}\right)$.

Proof. Let f, g, h be smooth functions and write

$$C(f,g,h) = ((f \prec g) \circ h) - f(g \circ h) = \sum_{j,k \geqslant -1} \sum_{i:|i-j|\leqslant 1} [\Delta_i(\Delta_k f \prec g)\Delta_j h - \Delta_k f \Delta_i g \Delta_j h].$$

Observe that for fixed k, the term $\Delta_k f \prec g$ has a Fourier transform supported outside of a ball $2^k \mathscr{B}$. Thus, we have $\Delta_i(\Delta_k f \prec g) = \mathbf{1}_{i \gtrsim k} \Delta_i(\Delta_k f \prec g)$. We can therefore apply Lemma 2.3 to obtain

$$\begin{aligned} C(f,g,h) &= \sum_{j,k \geqslant -1} \sum_{i:|i-j|\leqslant 1} [\mathbf{1}_{i \gtrsim k}(\Delta_k f \Delta_i g + R_i(\Delta_k f, g))\Delta_j h - \Delta_k f \Delta_i g \Delta_j h] \\ &= \sum_{j,k \geqslant -1} \sum_{i:|i-j|\leqslant 1} [\mathbf{1}_{i \gtrsim k} R_i(\Delta_k f, g)\Delta_j h - \mathbf{1}_{i \leqslant k-N} \Delta_k f \Delta_i g \Delta_j h] \end{aligned} \tag{9}$$

for some fixed $N \in \mathbb{N}$. We treat the two sums separately. First observe that for fixed k, the term $\sum_{j \geqslant -1} \sum_{i:|i-j|\leqslant 1} \mathbf{1}_{i \leqslant k-N} \Delta_k f \Delta_i g \Delta_j h$ has a Fourier transform which is supported in a ball $2^k \mathscr{B}$. Moreover,

$$\begin{aligned} \Big\| \sum_{j \geqslant -1} \sum_{i:|i-j|\leqslant 1} \mathbf{1}_{i \leqslant k-N} \Delta_k f \Delta_i g \Delta_j h \Big\|_{L^\infty} &\lesssim 2^{-k\alpha} \|f\|_\alpha \sum_{i=-1}^{k-N} 2^{-i(\beta+\gamma)} \|g\|_\beta \|h\|_\gamma \\ &\simeq 2^{-k(\alpha+\beta+\gamma)} \|f\|_\alpha \|g\|_\beta \|h\|_\gamma, \end{aligned}$$

where in the second step we used that $\beta + \gamma < 0$. Since $\alpha + \beta + \gamma > 0$, the estimate for the second series in (9) follows from Lemma A.3.

For the first series, recall that $R_i(\Delta_k f, g) = \Delta_i(\Delta_k f \prec g) - \Delta_k f \Delta_i g$. So for fixed j, the Fourier transform of $\sum_{k \geqslant -1} \sum_{i:|i-j|\leqslant 1} \mathbf{1}_{i \gtrsim k} R_i(\Delta_k f, g)\Delta_j h$ is supported in ball $2^j \mathscr{B}$. Furthermore, Lemma 2.3 yields

$$\begin{aligned} \Big\| \sum_{k \geqslant -1} \sum_{i:|i-j|\leqslant 1} \mathbf{1}_{i \gtrsim k} R_i(\Delta_k f, g)\Delta_j h \Big\|_{L^\infty} &= \Big\| \sum_{i:|i-j|\leqslant 1} R_i\Big(\sum_{k \lesssim i} \Delta_k f, g \Big) \Delta_j h \Big\|_{L^\infty} \\ \lesssim \sum_{i:|i-j|\leqslant 1} 2^{-i(\alpha+\beta)} \Big\| \sum_{k \lesssim i} \Delta_k f \Big\|_\alpha \|g\|_\beta 2^{-j\gamma} \|h\|_\gamma &\lesssim 2^{-j(\alpha+\beta+\gamma)} \|f\|_\alpha \|g\|_\beta \|h\|_\gamma, \end{aligned}$$

so that the claimed bound for $\|C(f,g,h)\|_{\alpha+\beta+\gamma}$ follows from another application of Lemma A.3.

Now we can uniquely extend C to a bounded trilinear operator on the closure of the smooth functions in $\mathscr{C}^\alpha \times \mathscr{C}^\beta \times \mathscr{C}^\gamma$. Unfortunately, this is a strict subset of $\mathscr{C}^\alpha \times \mathscr{C}^\beta \times \mathscr{C}^\gamma$. But we obtain similar bounds for C acting on $\mathscr{C}^{\alpha'} \times \mathscr{C}^{\beta'} \times \mathscr{C}^{\gamma'}$ for $\alpha' \in (0,1)$ and $\beta', \gamma' \in \mathbb{R}$, such that $\alpha' < \alpha$, $\beta' < \beta$, $\gamma' < \gamma$, and $\alpha' + \beta' + \gamma' > 0$. Since $\mathscr{C}^\alpha \times \mathscr{C}^\beta \times \mathscr{C}^\gamma$ is contained in the closure of the smooth functions in $\mathscr{C}^{\alpha'} \times \mathscr{C}^{\beta'} \times \mathscr{C}^{\gamma'}$, the extension of C to $\mathscr{C}^\alpha \times \mathscr{C}^\beta \times \mathscr{C}^\gamma$ is unique. □

Remark 2.5. 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 2.6 (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*

$$F(f) = F'(f) \prec f + R_F(f) \tag{10}$$

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 $i \geqslant 1$ 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 $i \geqslant 0$ and $\int K_{<i-1}(z)\mathrm{d}z = 1$ for $i \geqslant 1$. 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 A.3. 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 2.7. *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 $\Pi_F\colon \mathscr{C}^\alpha \times \mathscr{C}^\gamma \to \mathscr{C}^{\alpha+\beta+\gamma}$ such that*

$$F(f) \circ g = F'(f)(f \circ g) + \Pi_F(f,g) \tag{11}$$

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. Just 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 2.4 and Lemma 2.6. □

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

Lemma 2.8. *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 $\Pi_\times\colon \mathscr{C}^\alpha \times \mathscr{C}^\alpha \times \mathscr{C}^\gamma \to \mathscr{C}^{2\alpha+\gamma}$, such that*

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

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

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

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

□

3 Rough differential equations

Let us now resume the analysis of Section 2.2. We want to study the RDE

$$\partial_t u = F(u)\xi, \qquad u(0) = u_0, \tag{12}$$

where $u_0 \in \mathbb{R}^d$, $u\colon \mathbb{R}\to\mathbb{R}^d$ is a continuous vector valued function, $\xi\colon \mathbb{R}\to\mathbb{R}^n$ is a vector valued distribution with values in $\mathscr{C}^{\alpha-1}$ for some $\alpha \in (1/3,1)$, and $F\colon \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^n,\mathbb{R}^d)$ is a family of vector fields on $\mathbb{R}^d$.

In order to obtain concrete estimates, we have to localize the equation. Therefore, we introduce a smooth cut-off function φ with support on $[-2,2]$, which is equal to 1 on $[-1,1]$ and modify the equation as

$$\partial_t u = \varphi F(u)\xi, \qquad u(0) = u_0.$$

In the regular setting, if u is solution to this equation, it is also a solution of the original equation on $[-1,1]$, and thus it is sufficient to study the last equation for local bounds. To avoid problems with the fact that the paraproduct is a (mildly) non-local operation, we modify the paracontrolled ansatz as follows:

$$u = \varphi(F(u)\prec\vartheta) + u^\sharp. \tag{13}$$

If $F \in C_b^2$, an easy computation gives

$$\begin{aligned}\partial_t u^\sharp &= \varphi F(u)\xi - (\partial_t\varphi)(F(u)\prec\vartheta) - \varphi(\partial_t F(u)\prec\vartheta) - \varphi(F(u)\prec\xi)\\ &= \varphi\left[(F(u)\succ\xi) + F'(u)((u-u_0)\circ\xi) + \Pi_{F_{u_0}}(u-u_0,\xi) - (\partial_t F(u)\prec\vartheta)\right]\\ &\quad - (\partial_t\varphi)(F(u)\prec\vartheta),\end{aligned}$$

where we set $F_{u_0}(x) = F(u_0+x)$ and used that $(F_{u_0})'(x-u_0) = F'(x)$ for all $x\in\mathbb{R}^d$. We subtract the contribution of the initial condition, because this will eventually allow us to solve the equation on a small interval whose length does not depend on u_0. If we plug in the modified paracontrolled ansatz for u, then $F'(u)((u-u_0)\circ\xi)$ becomes

$$F'(u)((u-u_0)\circ\xi) = F'(u)((\varphi(F(u)\prec\vartheta))\circ\xi) + F'(u)((u^\sharp-u_0)\circ\xi).$$

For the first term on the right hand side we can further use that

$$(\varphi(F(u)\prec\vartheta))\circ\xi = \varphi((F(u)\prec\vartheta)\circ\xi) + (F(u)\prec\vartheta)(\varphi\circ\xi) + \Pi_\times(\varphi, F(u)\prec\vartheta,\xi),$$

where we recall that $\Pi_\times$ was defined in Lemma 2.8. So finally, an application of our commutator lemma yields

$$\begin{aligned}\partial_t u^\sharp = \varphi\Big[&(F(u)\succ\xi) + \Pi_{F_{u_0}}(u-u_0,\xi) + F'(u)((u^\sharp-u_0)\circ\xi) + (F(u)\prec\vartheta)(\varphi\circ\xi)\\ &+\Pi_\times(\varphi,F(u)\prec\vartheta,\xi) + \varphi C(F(u),\vartheta,\xi) + F'(u)\varphi F(u)(\vartheta\circ\xi) - (\partial_t F(u)\prec\vartheta)\Big]\\ &-(\partial_t\varphi)(F(u)\prec\vartheta)\\ =\ &\varphi\Phi^\sharp - (\partial_t\varphi)(F(u)\prec\vartheta), \end{aligned}\tag{14}$$

where $\Phi^\sharp$ is defined to be the term in the square large brackets. Let us summarize our observations so far.

Lemma 3.1. *Let ξ be a smooth path, let ϑ be such that $\partial_t\vartheta = \xi$, and let $F\in C^2_b$. Then u solves the ODE*

$$\partial_t u = \varphi F(u)\xi, \qquad u(0) = u_0,$$

if and only if $u = \varphi(F(u)\prec\vartheta) + u^\sharp$, where $u^\sharp$ solves

$$\partial_t u^\sharp = \varphi\Phi^\sharp - (\partial_t\varphi)(F(u)\prec\vartheta), \qquad u^\sharp(0) = u_0 - \varphi(F(u)\prec\vartheta)(0),$$

and where $\Phi^\sharp$ is defined in (14)*. Moreover, for $\alpha\in(1/3,1/2)$ we have the estimate*

$$\|\Phi^\sharp\|_{2\alpha-1} \lesssim C_F C_\xi(1 + \|u-u_0\|_\alpha + \|u-u_0\|_\alpha^2 + \|u^\sharp-u_0\|_{2\alpha}),$$

where

$$C_\xi = \|\xi\|_{\alpha-1} + \|\vartheta\|_\alpha + \|\vartheta\circ\xi\|_{2\alpha-1} + \|\vartheta\|_\alpha\|\xi\|_{\alpha-1} \qquad \text{and} \qquad C_F = \|F\|_{C^2_b} + \|F\|^2_{C^2_b}.$$

The estimate for $\Phi^\sharp$ follows from a somewhat lengthy but elementary calculation based on the decomposition (14), where we estimate the L^∞ norm rather than the $\mathscr{C}^{2\alpha-1}$ norm for each term where this is possible.

Plugging in the correct initial condition for $u^\sharp$ leads to

$$\begin{aligned}u^\sharp(t) &= u_0 - (F(u)\prec\vartheta)(0) + \int_0^t \partial_s u^\sharp(s)\mathrm{d}s\\ &= u_0 - (F(u)\prec\vartheta)(0) + \int_0^t(\varphi\Phi^\sharp)(s)\mathrm{d}s - \int_0^t(\partial_s\varphi)(s)(F(u)\prec\vartheta)(s)\mathrm{d}s.\end{aligned}$$

Now φ is compactly supported, and therefore Lemma A.10 gives estimates for the integrals appearing on the right hand side in terms of distributional norms of the integrands, and we obtain the bound

$$\|u^\sharp-u_0\|_{2\alpha} \lesssim \|F(u)\prec\vartheta\|_{2\alpha-1} + \|\Phi^\sharp\|_{2\alpha-1} \lesssim C_F C_\xi(1 + \|u-u_0\|_\alpha + \|u-u_0\|^2_\alpha + \|u^\sharp-u_0\|_{2\alpha}).$$

Using that $u = \varphi(F(u) \prec \vartheta) + u^\sharp$, we moreover have

$$\|u - u_0\|_\alpha \lesssim \|F\|_{L^\infty} \|\vartheta\|_{2\alpha} + \|u^\sharp - u_0\|_{2\alpha}.$$

From these two estimates we deduce that if C_F is small enough (depending only on C_ξ but not on $|u_0|$), then $\|u^\sharp\|_{2\alpha} \leqslant |u_0| + 1$. This is the required uniform estimate on the problem.

Similarly we can show that if $F \in C_b^3$ and if $\|F\|_{C_b^3}$ is small enough, then the map

$$(u_0, \xi, \vartheta, \xi \circ \vartheta) \mapsto (u, u^\sharp)$$

is locally Lipschitz continuous from $\mathscr{C}^{\alpha-1} \times \mathscr{C}^\alpha \times \mathscr{C}^{2\alpha-1} \times \mathbb{R}^d$ to $\mathscr{C}^\alpha \times \mathscr{C}^{2\alpha-1}$. To summarize:

Lemma 3.2. *Let $a > 0$ and let $\|F\|_{C_b^3}$ be small enough (depending on a). Let ξ, ϑ, and φ be smooth functions with $\xi = \partial_t \vartheta$ and such that φ has compact support. If $\alpha > 1/3$ and*

$$\max\{\|\xi\|_{\alpha-1}, \|\vartheta\|_\alpha, \|\xi \circ \vartheta\|_{2\alpha-1}, \|\varphi\|_{C_b^1}\} \leqslant a, \tag{15}$$

then for every $u_0 \in \mathbb{R}^d$ there exists a unique global solution u to

$$\partial_t u = \varphi F(u) \xi, \qquad u(0) = u_0.$$

For fixed φ and F, the solution u depends in a Lipschitz continuous way on $(u_0, \xi, \vartheta, \xi \circ \vartheta)$ satisfying (15)*.*

In order to ensure that $\|F\|_{C_b^3}$ is small enough we can use a dilation argument. Recall that the scaling operator Λ_λ is defined for $\lambda > 0$ by $\Lambda_\lambda u = u(\lambda \cdot)$. If we let $u^\lambda = \Lambda_\lambda u$ and $\xi^\lambda = \lambda^{1-\alpha} \Lambda_\lambda \xi$ for $\lambda > 0$, then u^λ solves

$$\partial_t u^\lambda = \lambda^\alpha F(u^\lambda) \xi^\lambda, \qquad u^\lambda(0) = u_0.$$

The rescaling of ξ^λ is chosen so that its $\mathscr{C}^\alpha$ norm is uniformly bounded by that of ξ as $\lambda \to 0$. Indeed, Lemma A.4 yields

$$\|\xi^\lambda\|_{\alpha-1} = \lambda^{1-\alpha} \|\Lambda_\lambda \xi\|_{\alpha-1} \lesssim (1 + \lambda^{1-\alpha}) \|\xi\|_{\alpha-1} \lesssim \|\xi\|_{\alpha-1}$$

for $\lambda \leqslant 1$. If moreover we let $\vartheta^\lambda = \lambda^{-\alpha} \Lambda_\lambda \vartheta$, then $\|\vartheta^\lambda \circ \xi^\lambda\|_{2\alpha-1} \lesssim \|\vartheta \circ \xi\|_{2\alpha-1} + \|\vartheta\|_\alpha \|\xi\|_{\alpha-1}$ by Lemma B.1 in Appendix B below. Thus, we deduce from Lemma 3.2 that for every φ of compact support there exists $\lambda > 0$, such that for all $u_0 \in \mathbb{R}^d$ we have a unique global solution u^λ to

$$\partial_t u^\lambda = \varphi \lambda^\alpha F(u^\lambda) \xi^\lambda, \qquad u^\lambda(0) = u_0.$$

The rescaled problem is equivalent to the original one upon the change $F \to \lambda^\alpha F$, $\xi \to \xi^\lambda$ and $\vartheta \circ \xi \to \vartheta^\lambda \circ \xi^\lambda$. So if we set $u = \Lambda_{\lambda^{-1}} u^\lambda$, then u is the unique global solution to

$$\partial_t u = \varphi_\lambda F(u) \xi, \qquad u(0) = u_0,$$

where we set $\varphi_\lambda(t) = \varphi(t/\lambda)$. In particular, if $\varphi \equiv 1$ on $[-1, 1]$, then u is the unique solution to the original RDE in the interval $[-\lambda, \lambda]$. Since λ can be chosen independently of u_0, we can now iterate on intervals of length 2λ, and obtain a global solution $u \in \mathscr{C}^\alpha_{\mathrm{loc}}$.

This analysis can be summarized in the following statement.

Theorem 3.3. *Let $\alpha > 1/3$. Assume that $(\xi^\varepsilon)_{\varepsilon > 0}$ is a family of smooth functions with values in $\mathbb{R}^n$, (u_0^ε) is a family of initial conditions in $\mathbb{R}^d$, and F is a family of C_b^3 vector fields on $\mathbb{R}^d$. Suppose that there exist $u_0 \in \mathbb{R}^d$, $\xi \in \mathscr{C}^{\alpha-1}$ and $\eta \in \mathscr{C}^{2\alpha-1}$ such that $(u_0^\varepsilon, \xi^\varepsilon, \vartheta^\varepsilon, (\vartheta^\varepsilon \circ \xi^\varepsilon))$ converges to $(u_0, \xi, \vartheta, \eta)$ in $\mathscr{C}^{\alpha-1} \times \mathscr{C}^\alpha \times \mathscr{C}^{2\alpha-1}$, where ϑ^ε and ϑ are solutions to $\partial_t \vartheta^\varepsilon = \xi^\varepsilon$ and*

$\partial_t\vartheta = \xi$, *respectively. Let for* $\varepsilon > 0$ *the function* u^ε *be the unique global solution to the Cauchy problem*

$$\partial_t u^\varepsilon = F(u^\varepsilon)\xi^\varepsilon, \qquad u^\varepsilon(0) = u_0^\varepsilon.$$

Then there exists $u \in \mathscr{C}^\alpha_{\mathrm{loc}}$ *such that* $u^\varepsilon \to u$ *in* $\mathscr{C}^\alpha_{\mathrm{loc}}$ *as* $\varepsilon \to 0$. *The limit* u *depends only on* $(u_0, \xi, \vartheta, \eta)$, *and not on the approximating family* $(u_0^\varepsilon, \xi^\varepsilon, \vartheta^\varepsilon, (\vartheta^\varepsilon \circ \xi^\varepsilon))$.

Proof. The only point which remains to be shown is the convergence of (u^ε) to u in $\mathscr{C}^\alpha_{\mathrm{loc}}$. A priori, we only know that for sufficiently small $\lambda > 0$, the solutions $\tilde{u}^\varepsilon$ to $\partial_t\tilde{u}^\varepsilon = \varphi_\lambda F(\tilde{u}^\varepsilon)\xi^\varepsilon$ with $\tilde{u}^\varepsilon(0) = u_0$ converge, as $\varepsilon \to 0$, in $\mathscr{C}^\alpha$ to a unique limit $\tilde{u}$. But since $\varphi_\lambda \equiv 1$ on $[-\lambda, \lambda]$, we have $\tilde{u}^\varepsilon|_{[-\lambda,\lambda]} = u^\varepsilon|_{[-\lambda,\lambda]}$. So if we define $u|_{[-\lambda,\lambda]} = \tilde{u}|_{[-\lambda,\lambda]}$, then $u|_{[-\lambda,\lambda]}$ does not depend on φ_λ. Moreover, for every $\psi \in \mathscr{D}$ with support contained in $[-\lambda, \lambda]$, we also have that $\|\psi(u^\varepsilon - u)\|_\alpha$ converges to zero as $\varepsilon \to 0$. Now we can iterate this construction of u on intervals of length 2λ. We end up with a distribution $u \in \mathscr{S}'$, which only depends on $(u_0, F, \xi, \vartheta, \eta)$, but not on φ_λ or on the approximating sequence $(u_0^\varepsilon, \xi^\varepsilon, \vartheta^\varepsilon, \xi^\varepsilon \circ \vartheta^\varepsilon)_{\varepsilon>0}$. If $\psi \in \mathscr{D}$, then it can be written as a finite sum of smooth functions with support contained in intervals of length 2λ, and therefore $\psi u = \lim_{\varepsilon\to} \psi u^\varepsilon$, where convergence takes places in $\mathscr{C}^\alpha$. □

Remark 3.4. By Lemma 2.7, it suffices if $F \in C_b^{2+\beta/\alpha}$ for some $\beta > 0$ with $2\alpha + \beta > 1$ to obtain existence and uniqueness solutions. If we only suppose $F \in C^{2+\beta/\alpha}$ and not that F and its derivatives are bounded, we still obtain local existence and uniqueness of solutions. In that case we may consider a function $G \in C_b^{2+\beta/\alpha}$ that coincides with F on $\{|x| \leqslant a\}$ for some $a > |u_0|$. The Cauchy problem

$$\partial_t v = G(v)\xi, \qquad v(0) = u_0,$$

then has a unique global solution in the sense of Theorem 3.3. If we stop v upon leaving the set $\{|x| \leqslant a\}$, we obtain a local solution to the RDE with vector field F.

3.1 Interpreting our RDE solutions

So far we showed that under the assumptions of Theorem 3.3 there exists a unique limit u of the solutions to the regularized equations, which does not depend on the particular approximating sequence. In that sense, one may formally call u the unique solution to

$$\partial_t u = F(u)\xi, \qquad u(0) = u_0.$$

But u is actually a weak solution to the equation if we interpret the product $F(u)\xi$ appropriately. Below we will introduce a map which extends the pointwise product $F(u)\xi$ from smooth ξ to $\xi \in \mathscr{C}^{\alpha-1}$ by a continuity argument. But first we present an auxiliary result which shows that the considered topologies and operators do not depend on the particular dyadic partition of unity that we use to describe them.

Lemma 3.5. *Let* $\alpha, \beta \in \mathbb{R}$. *Let* (χ, ρ) *and* $(\tilde{\chi}, \tilde{\rho})$ *be two dyadic partitions of unity and let* $(\prec, \succ, \circ)$ *and* $(\widetilde{\prec}, \widetilde{\succ}, \widetilde{\circ})$ *denote paraproducts and resonant term defined in terms of* (χ, ρ) *and* $(\tilde{\chi}, \tilde{\rho})$, *respectively. Then*

$$(u, v) \mapsto (u \prec v - u \mathbin{\widetilde{\prec}} v, u \circ v - u \mathbin{\widetilde{\circ}} v, u \succ v - u \mathbin{\widetilde{\succ}} v)$$

is a bounded bilinear operator from $\mathscr{C}^\alpha \times \mathscr{C}^\beta$ *to* $(\mathscr{C}^{\alpha+\beta})^3$.

Proof. The statement for $(u, v) \mapsto (u \prec v - u \mathbin{\widetilde{\prec}} v)$ (and thus for $(u, v) \mapsto (u \succ v - u \mathbin{\widetilde{\succ}} v)$) is shown in Bony [Bon81], Theorem 2.1. But for smooth functions u and v we have $u \circ v = uv - u \prec v - u \succ v$, and similarly for $u \mathbin{\widetilde{\circ}} v$. Thus, the bound on $u \circ v - u \mathbin{\widetilde{\circ}} v$ follows from the bounds on $u \prec v - u \mathbin{\widetilde{\prec}} v$ and on $u \succ v - u \mathbin{\widetilde{\succ}} v$ in combination with a continuity argument. □

Our commutator lemma states that if the product $g \circ h$ is given, then we can unambiguously make sense of the product $(f \prec g) \circ h$ for suitable f. This leads us to the following definition.

Definition 3.6. *Let $\alpha \in \mathbb{R}$, $\beta > 0$, and let $v \in \mathscr{C}^\alpha$. A pair of distributions $(u, u') \in \mathscr{C}^\alpha \times \mathscr{C}^\beta$ is called* paracontrolled *by v if*

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

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

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

According to Lemma 3.5, the space $\mathscr{D}^\beta$ does not depend on the specific partition of unity used to define it. To construct the product $F(u)\xi$, we could now show that smooth F preserve the paracontrolled structure of u. This can be achieved by combining Lemma 2.6 with another commutator lemma (Theorem 2.3 in [Bon81]). But we do not need the full strength of that result, let us just show that if u is paracontrolled by ϑ and F is smooth enough, then $F(u)\xi$ is well defined.

Theorem 3.7. *Let $\alpha \in (0,1)$, $\beta \in (0, \alpha]$, $\gamma < 0$ be such that $\alpha + \beta + \gamma > 0$. Let $F \in C^{1+\beta/\alpha}$ and let $v \in \mathscr{C}^\alpha$, $w \in \mathscr{C}^\gamma$, $\eta \in \mathscr{C}^{\alpha+\gamma}$ be such that there exist sequences of smooth functions (v_n), (w_n), converging to v and w respectively, such that $(v_n \circ w_n)$ converges to η. Then*

$$\begin{aligned}\mathscr{D}^\beta(v) \ni u \mapsto F(u)w = F(u) \succ w + F(u) \prec w + \Pi_F(u, w) + F'(u)(u^\sharp \circ w) \\ + F'(u)C(u', v, w) + F'(u)u'\eta \in \mathscr{C}^\gamma\end{aligned} \tag{16}$$

defines a locally Lipschitz continuous function. If w is a smooth function and $\eta = v \circ w$, then $F(u)w$ is simply the pointwise product.

The product $F(u)w$ does not depend on the specific dyadic partition used to construct it: if $(\widetilde{\prec}, \widetilde{\succ}, \widetilde{\circ})$ denote paraproducts and resonant term defined in terms of another partition unity, if

$$\tilde{\eta} = \eta + v \prec w + v \succ w - v \mathbin{\widetilde{\prec}} w - v \mathbin{\widetilde{\succ}} w,$$

and $\tilde{u}^\sharp = u' \mathbin{\widetilde{\prec}} v$, then $F(u)w$ is equal to the right hand side of (16) if we replace every operator by the corresponding operator defined in terms of $(\widetilde{\prec}, \widetilde{\succ}, \widetilde{\circ})$, and we replace $u^\sharp$ by $\tilde{u}^\sharp$ and η by $\tilde{\eta}$.

Proof. The local Lipschitz continuity of the product follows from its definition in combination with Lemma 2.4, Lemma 2.7, and the paraproduct estimates Lemma 2.1.

If w is smooth and $\eta = v \circ w$, then

$$F'(u)C(u', v, w) + F'(u)u'\eta = F'(u)((u' \prec v) \circ w),$$

and therefore

$$\Pi_F(u, w) + F'(u)(u^\sharp \circ w) + F'(u)C(u', v, w) + F'(u)u'\eta = \Pi_F(u, w) + F'(u)(u \circ w) = F(u) \circ w,$$

which shows that we recover $F(u) \prec w + F(u) \succ w + F(u) \circ w$, i.e. the pointwise product.

It remains to show that $F(u)w$ does not depend on the specific dyadic partition of unity. By continuity of the operators involved, we have

$$\begin{aligned}F(u)w &= \lim_{n\to\infty} \Big[F(u) \prec w_n + F(u) \succ w_n + \Pi_F(u, w_n) + F'(u)(u^\sharp \circ w_n) \\ &\qquad + F'(u)C(u', v_n, w_n) + F'(u)u'(v_n \circ w_n)\Big] \\ &= \lim_{n\to\infty} \Big[F(u)w_n + F'(u)((u' \prec (v_n - v)) \circ w_n)\Big].\end{aligned}$$

Assume now that we defined $F(u) \cdot w$ in terms of another partition of unity, as described above. Then Lemma 3.5 implies the convergence of $(v_n \,\widetilde{\circ}\, w_n)$ to $\tilde\eta$ in $\mathscr{C}^{\alpha+\gamma}$, and therefore

$$F(u) \cdot w = \lim_{n\to\infty} \Big[F(u)w_n + F'(u)((u' \,\widetilde{\prec}\, (v_n - v)) \,\widetilde{\circ}\, w_n)\Big].$$

Another application of Lemma 3.5 then yields $F(u)w = F(u) \cdot w$. □

Remark 3.8. If in the setting of Theorem 3.7 we let $\tilde v = v + f$ for some $f \in \mathscr{C}^{\alpha+\beta}$, then we have $\widetilde{\mathscr{D}^\beta(v)} = \mathscr{D}^\beta(\tilde v)$, and it is easy to see that if we set $\tilde\eta = \eta + f \circ w$, $\tilde u^\sharp = u - u' \prec \tilde v$, and define $\widetilde{F(u)}w$ like $F(u)w$, with $\tilde v$, $\tilde u^\sharp$, $\tilde\eta$ replacing $u^\sharp, v, \eta$, then $\widetilde{F(u)}w = F(u)w$.

With this product operator at hand, it is relatively straightforward to show that if ξ has compact support (which in general is necessary to have $u \in \mathscr{C}^\alpha$ and not just in $\mathscr{C}^\alpha_{\mathrm{loc}}$), then the solution u that we constructed in Theorem 3.3 is the unique element of $\mathscr{D}^\alpha$ which solves $\partial_t u = F(u)\xi$, $u(0) = u_0$, in the weak sense. Remark 3.8 explains why we did not fix the initial condition $\vartheta(0)$ in Theorem 3.3: it is of no importance whatsoever.

3.2 Alternative approach

We briefly describe an alternative approach to RDEs which avoids the paracontrolled ansatz. The idea is to control $u \circ \xi$ directly by exploiting that u solves the differential equation $\partial_t u = F(u)\xi$. Indeed, let as above ϑ be a solution to $\partial_t \vartheta = \xi$ and observe that the Leibniz rule yields

$$u \circ \xi = u \circ \partial_t \vartheta = \partial_t(u \circ \vartheta) - \partial_t u \circ \vartheta = \partial_t(u \circ \vartheta) - (F(u)\xi) \circ \vartheta.$$

Now the second term on the right hand side can be rewritten as

$$\begin{aligned}(F(u)\xi) \circ \vartheta &= (F(u) \prec \xi) \circ \vartheta + (F(u) \circ \xi) \circ \vartheta + (F(u) \succ \xi) \circ \vartheta \\ &= F(u)(\xi \circ \vartheta) + C(F(u), \xi, \vartheta) + (F'(u)(u \circ \xi)) \circ \vartheta + \Pi_F(u, \xi) \circ \vartheta + (F(u) \succ \xi) \circ \vartheta.\end{aligned}$$

Combining these two equations, we see that

$$\begin{gathered} u \circ \xi = \Phi - (F'(u)(u \circ \xi)) \circ \vartheta, \qquad \text{where} \\ \Phi = \partial_t(u \circ \vartheta) - F(u)(\xi \circ \vartheta) - C(F(u), \xi, \vartheta) - \Pi_F(u, \xi) \circ \vartheta - (F(u) \succ \xi) \circ \vartheta.\end{gathered}$$

This is an implict equation for $u \circ \xi$ which can be solved by fixed point methods. For example, it is easy to obtain the estimate

$$\|u \circ \xi\|_{2\alpha-1} \lesssim \|\Phi\|_{2\alpha-1} + C_F \|u \circ \xi\|_{2\alpha-1} \|\vartheta\|_\alpha,$$

and if C_F is small enough this leads to $\|u \circ \xi\|_{2\alpha-1} \lesssim \|\Phi\|_{2\alpha-1}$. Moreover, we have $\|\Phi\|_{2\alpha-1} \lesssim C_\xi[\|u\|_\alpha + C_F(1 + \|u\|_\alpha)^2]$. These estimates can be reinjected into the equation

$$\partial_t u = F(u)\xi = F(u) \prec \xi + F'(u)(u \circ \xi) + F(u) \succ \xi + R(F'(u), u, \xi)$$

to obtain a local estimate for u.

3.3 Structure of solutions to RDEs

In this section, we would like to discuss how the combination of analytic and algebraic requirements generate very interesting phenomena in the context of irregular PDEs. We will discuss only the simple case of RDEs but similar considerations apply also to the other models. These remarks are intentionally sketchy and have only a heuristic purpose, we plan to come back to them more systematically in a further publication.

The stable form of an RDE. As we have seen, the solution to the RDE (12) for smooth ξ can be understood as a regular function of $(u_0, \xi, \vartheta \circ \xi) \in \mathbb{R}^d \times \mathscr{C}^{\alpha-1}(\mathbb{R}, \mathbb{R}^n) \times \mathscr{C}^{2\alpha-1}(\mathbb{R}, \mathbb{R}^n \times \mathbb{R}^n)$, where ϑ is a solution to $\partial_t \vartheta = \xi$. Let us denote this function by $u = \Psi(u_0, \xi, \vartheta \circ \xi)$. The structure of the solution which we derived above shows that for $\eta \in \mathscr{C}^{2\alpha-1}$, the RDE

$$\partial_t v = F(v)\xi + F'(v)F(v)\eta, \qquad v(0) = u_0, \tag{17}$$

has a solution which can be obtained from the same map by replacing $\vartheta \circ \xi$ with $\vartheta \circ \xi + \eta$, i.e.

$$v = \Psi(u_0, \xi, \vartheta \circ \xi + \eta).$$

Moreover, the solution v depends continuously on the data $\vartheta \circ \xi + \eta \in \mathscr{C}^{2\alpha-1}$. First note that this regularity hypothesis concerns only the combination $\vartheta \circ \xi + \eta$, and second note that many equations share a similar structure and one can pass from solutions of one equation to solutions of other equations via a transformation of the (extended) data of the problem.

For many reasons, all these differential equations should be considered to be the same object, especially when dealing with data of low regularity. To understand this point of view, consider two different families $(\xi^\varepsilon)_{\varepsilon>0}$ and $(\tilde{\xi}^\varepsilon)_{\varepsilon>0}$ of smooth functions, such that both converge to ξ in $\mathscr{C}^{\alpha-1}$. The corresponding solutions (u^ε) and $(\tilde{u}^\varepsilon)$ to equation (12) are given by $u^\varepsilon = \Psi(u_0, \xi^\varepsilon, \vartheta^\varepsilon \circ \xi^\varepsilon)$ and $\tilde{u}^\varepsilon = \Psi(u_0, \tilde{\xi}^\varepsilon, \tilde{\vartheta}^\varepsilon \circ \tilde{\xi}^\varepsilon)$. Let us take the limit as $\varepsilon \to 0$ and assume that $\vartheta^\varepsilon \circ \xi^\varepsilon \to \vartheta$ in $\mathscr{C}^{2\alpha-1}$ and also that $\tilde{\vartheta}^\varepsilon \circ \tilde{\xi}^\varepsilon \to \tilde{\vartheta}$ in $\mathscr{C}^{2\alpha-1}$, where $\vartheta \neq \tilde{\vartheta}$. This could happen in principle, and it is not difficult to find specific and relevant examples of this multiplicity of limits. Of course, we have

$$\vartheta^\varepsilon \circ \xi^\varepsilon - \tilde{\vartheta}^\varepsilon \circ \tilde{\xi}^\varepsilon = \vartheta^\varepsilon \xi^\varepsilon - \tilde{\vartheta}^\varepsilon \tilde{\xi}^\varepsilon - \vartheta^\varepsilon \prec \xi^\varepsilon - \vartheta^\varepsilon \succ \xi^\varepsilon + \tilde{\vartheta}^\varepsilon \prec \tilde{\xi}^\varepsilon + \tilde{\vartheta}^\varepsilon \succ \tilde{\xi}^\varepsilon.$$

Since (ϑ^ε) and $(\tilde{\vartheta}^\varepsilon)$ both converge to ϑ, where $\partial_t \vartheta = \xi$, we deduce from the continuity of the paraproduct that all the terms on the right hand side cancel, except the first two, and we remain with

$$\eta = \vartheta - \vartheta' = \lim_{n \to \infty} (\vartheta^\varepsilon \xi^\varepsilon - \tilde{\vartheta}^\varepsilon \tilde{\xi}^\varepsilon).$$

Incidentally, this line of reasoning also shows that if the limit exists, it does not depend on the particular Littlewood-Paley decomposition we use to compute it. From the continuity of Ψ it moreover follows that $u^\varepsilon \to u$ and $\tilde{u}^\varepsilon \to \tilde{u}$ where $u = \Psi(u_0, \xi, \vartheta)$ and $\tilde{u} = \Psi(u_0, \xi, \vartheta + \eta)$. That is, different approximations of the same equation could lead to different equations in the limit. In particular, if ξ is smooth enough (but (ξ^ε) does not converge to ξ in a space of sufficiently high regularity), we can interpret u as a classical solution to a differential equation, and in this case u' will solve a modified equation.

Rough paths as a transformation group. We can therefore identify $\mathscr{C}^{\alpha-1} \times \mathscr{C}^{2\alpha-1}$ with a transformation group $(T_{f,g})_{f,g \in \mathscr{C}^{\alpha-1} \times \mathscr{C}^{2\alpha-1}}$, which acts on solutions to ODEs via

$$T_{f,g}\Psi(u_0, \xi, \eta) = \Psi(u_0, \xi + f, \eta + \vartheta \circ f + \Phi \circ \xi + \Phi \circ f + g).$$

where Φ solves $\partial_t \Phi = f$. In particular, $\Psi(u_0, \xi, \eta) = T_{\xi, \eta - \vartheta \circ \xi}\Psi(u_0, 0, 0)$. The neutral element is $T_{0,0}$ and the group operation is $T_{f,g}T_{f',g'} = T_{f+f', g+g'}$, so that the group is abelian. A simple distance is given by

$$d(T_{f,g}, T_{f',g'}) = d(T_{f-f', g-g'}, T_{0,0}) = \|f - f'\|_{\alpha-1} + \|g - g' + (\Phi - \Phi') \circ (f - f')\|_{2\alpha-1},$$

where Φ and Φ' are the definite integrals with $\Phi(0) = \Phi'(0) = 0$ of f and f' respectively.

Geometric conditions. Note that if J_S and J_A denote respectively the projections onto symmetric and antisymmetric tensors, then for smooth ξ, ϑ we have

$$J_S(\vartheta \circ \xi) = J_S(\vartheta \circ \partial_t \vartheta) = \frac{1}{2}\partial_t J_S(\vartheta \circ \vartheta),$$

where the right hand side is now well defined for all $\vartheta \in \mathscr{C}^\alpha$ and defines a distribution in $\mathscr{C}^{2\alpha-1}$. So for any family $\xi^\varepsilon \to \xi$ in $\mathscr{C}^{\alpha-1}$ we have that

$$J_S(\vartheta^\varepsilon \circ \xi^\varepsilon) = \frac{1}{2}\partial_t J_S(\vartheta^\varepsilon \circ \vartheta^\varepsilon) \to \frac{1}{2}\partial_t J_S(\vartheta \circ \vartheta)$$

as $\varepsilon \to 0$, and only the antisymmetric part $J_A(\vartheta^\varepsilon \circ \xi^\varepsilon)$ could feature multiple accumulation points. To highlight the geometric meaning of the symmetric and antisymmetric parts of $\vartheta \circ \xi$ we can submit the RDE to a nontrivial transformation given by the application of a smooth diffeomorphism $\phi \colon \mathbb{R}^d \to \mathbb{R}^d$. If u is a solution to the RDE $\partial_t u = F(u)\xi$, then $v = \phi(u)$ solves

$$\partial_t v = \phi'(u)\partial_t u = \phi'(u)F(u)\xi,$$

i.e. an RDE with vector field $(\phi' F) \circ \phi^{-1}$. We have

$$\partial_t \phi(T_{f,g}u) = \phi'(T_{f,g}u)\partial_t(T_{f,g}u) = \phi'(T_{f,g}u)(F(T_{f,g}u)(\xi + f) + F'(T_{f,g}u)F(T_{f,g}u)g).$$

On the other hand, we obtain for $T_{f,g}v = T_{f,g}\phi(u)$ that

$$\partial_t T_{f,g}v = ((\phi' F) \circ \phi^{-1})(T_{f,g}v)(\xi + f) + \mathrm{D}((\phi' F) \circ \phi^{-1})(T_{f,g}v)((\phi' F) \circ \phi^{-1})(T_{f,g}v)g,$$

and we have $\mathrm{D}((\phi' F) \circ \phi^{-1}) = ((\phi'' F + \phi' F') \circ \phi^{-1})(\phi')^{-1} \circ \phi^{-1}$, where we slightly abuse notation by writing ϕ^{-1} for the inverse function of ϕ and $(\phi')^{-1}$ for the inverse matrix of ϕ'. Thus, we obtain

$$\partial_t T_{f,g}v = ((\phi' F) \circ \phi^{-1})(T_{f,g}v)(\xi + f) + ((\phi'' F F + \phi' F' F) \circ \phi^{-1})(T_{f,g}v)g,$$

so that $\phi(T_{f,g}u) = T_{f,g}(\phi(u))$ only if $(\phi'' F F)(\phi^{-1}(T_{f,g}v))g = 0$. If this holds for all F and ϕ, then the symmetric part $J_S g$ of g must be 0. In this sense, $T_{f,g}$ acts geometrically only if $J_S g = 0$.

3.4 Connections to rough paths and existence of the area

We saw in the previous section that the solution u to an RDE of the form $\partial_t u = F(u)\xi$ depends on the driving signal in a continuous way, provided that we not only keep track of ξ but also of $\vartheta \circ \xi$. From the theory of rough paths it is well known that the same holds true if we keep track of ϑ and its iterated integrals $\int_s^t \int_s^{r_2} \mathrm{d}\vartheta(r_1)\mathrm{d}\vartheta(r_2)$. But in fact the convergence of $(\vartheta^\varepsilon \circ \xi^\varepsilon)$ is equivalent to the convergence of the iterated integrals $\int \int \mathrm{d}\vartheta^\varepsilon \mathrm{d}\vartheta^\varepsilon$:

Corollary 3.9. *Let $(u^\varepsilon, v^\varepsilon)_{\varepsilon > 0}$ be a family of smooth functions on $\mathbb{R}$. Define for every $\varepsilon > 0$ the "area"*

$$A^\varepsilon_{s,t} = \int_s^t \int_s^{r_2} \mathrm{d}u^\varepsilon(r_1)\mathrm{d}v^\varepsilon(r_2), \qquad s < t \in \mathbb{R}.$$

Let $\alpha, \beta \in (0,1)$ with $\alpha + \beta < 1$ and let $u \in \mathscr{C}^\alpha, v \in \mathscr{C}^\beta, \eta \in \mathscr{C}^{\alpha+\beta-1}$. Then $(u^\varepsilon, v^\varepsilon, u^\varepsilon \circ \partial_t v^\varepsilon)$ converges to (u, v, η) in $\mathscr{C}^\alpha \times \mathscr{C}^\beta \times \mathscr{C}^{\alpha+\beta-1}$ if and only if $(u^\varepsilon, v^\varepsilon)$ converges to (u, v) in $\mathscr{C}^\alpha \times \mathscr{C}^\beta$, and if moreover

$$\lim_{\varepsilon \to 0} \left(\sup_{s \neq t \in \mathbb{R}, |s-t| \leqslant 1} \frac{|A_{s,t} - A^\varepsilon_{s,t}|}{|t-s|^{\alpha+\beta}} \right) = 0, \tag{18}$$

where we set $A_{s,t} = \int_s^t (\eta + (u \prec \partial_t v) + (u \succ \partial_t v))(r)\mathrm{d}r - u(s)(v(t) - v(s))$ for $s, t \in \mathbb{R}$.

Proof. First suppose that $(u^\varepsilon, v^\varepsilon, u^\varepsilon \circ \partial_t v^\varepsilon)$ converges to (u, v, η) in $\mathscr{C}^\alpha \times \mathscr{C}^\beta \times \mathscr{C}^{\alpha+\beta-1}$, and let $s, t \in \mathbb{R}$ with $|s-t| \leqslant 1$. We have

$$
\begin{aligned}
A_{s,t} - A^\varepsilon_{s,t} = & \int_s^t (\eta + u \succ \partial_t v - u^\varepsilon \circ \partial_t v^\varepsilon - u^\varepsilon \succ \partial_t v^\varepsilon)(r)\mathrm{d}r \\
& + \int_s^t ((u^\varepsilon - u) \prec \partial_t v^\varepsilon)(r)\mathrm{d}r - (u^\varepsilon - u)(s)(v^\varepsilon(t) - v^\varepsilon(s)) \\
& + \int_s^t (u \prec \partial_t (v^\varepsilon - v))(r)\mathrm{d}r - (u)(s)((v^\varepsilon - v)(t) - (v^\varepsilon - v)(s)). \qquad (19)
\end{aligned}
$$

The first term on the right hand side can be estimated with the help of Lemma A.10, which allows us to bound increments of the integral in terms of Besov norms of the integrand. We get

$$
\begin{aligned}
& \left| \int_s^t (\eta + u \succ \partial_t v - u^\varepsilon \circ \partial_t v^\varepsilon - u^\varepsilon \succ \partial_t v^\varepsilon)(r)\mathrm{d}r \right| \\
& \qquad \lesssim (\|\eta - u^\varepsilon \circ \partial_t v^\varepsilon\|_{\alpha+\beta-1} + \|u - u^\varepsilon\|_\alpha \|\partial_t v\|_{\beta-1} + \|u^\varepsilon\|_\alpha \|\partial_t (v^\varepsilon - v)\|_{\beta-1}) |t-s|^{\alpha+\beta},
\end{aligned}
$$

using also the paraproduct estimates. Since $\|\partial_t (v^\varepsilon - v)\|_{\beta-1} \lesssim \|v^\varepsilon - v\|_\beta$, the right hand side goes to zero if we divide it by $|t-s|^{\alpha+\beta}$ and let $\varepsilon \to 0$.

The second term on the right hand side of (19) can be estimated using Lemma B.2, which roughly states that time integral and paraproduct commute with each other, at the price of introducing a smoother remainder term:

$$
\left| \int_s^t ((u^\varepsilon - u) \prec \partial_t v^\varepsilon)(r)\mathrm{d}r - (u^\varepsilon - u)(s)(v^\varepsilon(t) - v^\varepsilon(s)) \right| \lesssim |t-s|^{\alpha+\beta} \|u^\varepsilon - u\|_\alpha \|v^\varepsilon\|_\beta,
$$

The third term on the right hand side of (19) is of the same type as the second term, and therefore the convergence in (18) follows.

Conversely, assume that $(u^\varepsilon, v^\varepsilon)$ converges to (u, v) in $\mathscr{C}^\alpha \times \mathscr{C}^\beta$, and that the convergence in (18) holds. It follows from the representation (19) and the convergence of $(u^\varepsilon, v^\varepsilon)$ to (u, v) in $\mathscr{C}^\alpha \times \mathscr{C}^\beta$, that also

$$
\lim_{\varepsilon \to 0} \left(\sup_{s \neq t \in \mathbb{R}, |s-t| \leqslant 1} \frac{|\int_s^t (\eta - u^\varepsilon \circ \partial_r v^\varepsilon)(r)\mathrm{d}r|}{|t-s|^{\alpha+\beta}} \right) = 0.
$$

Due to the restriction $|s-t| \leqslant 1$, it is not entirely obvious that this implies the convergence of $u^\varepsilon \circ \partial_r v^\varepsilon$ to η in $\mathscr{C}^{\alpha+\beta-1}$. However, here we can use an alternative characterization of Besov spaces in terms of local means. Let k^0 and k be infinitely differentiable functions on $\mathbb{R}$ with support in $(-1, 1)$, such that $\mathcal{F}k^0(0) \neq 0$, and such that there exists $\delta > 0$ with $\mathcal{F}k(z) \neq 0$ for all $0 < |z| < \delta$. Then an equivalent norm on $\mathscr{C}^{\alpha+\beta-1}(\mathbb{R})$ is given by

$$
\|w\|_{\alpha+\beta-1} \simeq \max \Big\{ \|k^0 * w\|_{L^\infty}, \sup_{j \geqslant 0} 2^{j(\alpha+\beta-1)} \|2^j k(2^j \cdot) * w\|_{L^\infty} \Big\},
$$

see [Tri06], Theorem 1.10. Let us write $f = \int_0^\cdot (\eta - u^\varepsilon \circ \partial_r v^\varepsilon)(r)\mathrm{d}r$ and let $t \in \mathbb{R}$ and $j \geqslant 0$. Then

$$
\begin{aligned}
|2^j k(2^j \cdot) * (\partial_t f)(t)| = 2^{2j} \left| \int_{\mathbb{R}} (\partial_t k)(2^j (t-s))(f(t) - f(s))\mathrm{d}s \right| \\
\lesssim 2^{2j} \int_{\mathbb{R}} |(\partial_t k)(2^j (t-s))| |t-s|^{\alpha+\beta} \mathrm{d}s \sup_{|a-b| \leqslant 1} \frac{|f(b) - f(a)|}{|b-a|^{\alpha+\beta}} \lesssim 2^{-j(\alpha+\beta-1)} \sup_{|a-b| \leqslant 1} \frac{|f(b) - f(a)|}{|b-a|^{\alpha+\beta}},
\end{aligned}
$$

where we used that $\int_{\mathbb{R}} \partial_t k(t)\mathrm{d}t = 0$, and that k is supported in $(-1,1)$. Similarly, we obtain

$$|k^0 * (\partial_t f)(t)| \lesssim \int_{\mathbb{R}} |\partial_t k^0(t-s)||t-s|^{\alpha+\beta}\mathrm{d}s \sup_{|a-b|\leqslant 1} \frac{|f(b)-f(a)|}{|b-a|^{\alpha+\beta}} \lesssim \sup_{|a-b|\leqslant 1} \frac{|f(b)-f(a)|}{|b-a|^{\alpha+\beta}},$$

from where the convergence of $u^\varepsilon \circ \partial_t v^\varepsilon$ to η in $\mathscr{C}^{\alpha+\beta-1}$ follows. □

Corollary 3.10. *Let X be an n–dimensional centered Gaussian process with independent components and measurable trajectories, whose covariance function satisfies for some $H \in (1/4, 1)$ the inequalities*

$$\begin{aligned} \mathbb{E}[|X_t - X_s|^2] &\lesssim |t-s|^{2H} \qquad \text{and} \\ |\mathbb{E}[(X_{s+r} - X_s)(X_{t+r} - X_t)]| &\lesssim |t-s|^{2H-2} r^2 \end{aligned} \tag{20}$$

for all $s,t \in \mathbb{R}$ and all $r \in [0, |t-s|)$. Then $\varphi X \in \mathscr{C}^\alpha$ for all $\alpha < H$ and all $\varphi \in \mathscr{D}$, and there exists $\eta \in \mathscr{C}^{2\alpha-1}$ such that for every $\psi \in \mathscr{S}$ with $\int \psi \mathrm{d}t = 1$ and for every $\delta > 0$ we have

$$\lim_{\varepsilon \to 0} \mathbb{P}\left(\|\psi^\varepsilon * (\varphi X) - (\varphi X)\|_\alpha + \|(\psi^\varepsilon * (\varphi X)) \circ \partial_t(\psi^\varepsilon * (\varphi X)) - \eta\|_{\mathscr{C}^{2\alpha-1}} > \delta\right) = 0,$$

where we define $\psi^\varepsilon = \varepsilon^{-1}\psi(\varepsilon^{-1}\cdot)$.

Proof. Since φ is smooth and of compact support, it is easy to see that also the Gaussian process φX satisfies the covariance condition (3.10), and using Gaussian hypercontractivity we obtain $\mathbb{E}[|\varphi(t)X_t - \varphi(s)X_s|^{2p}] \lesssim |t-s|^{2Hp}$ for all $p \geqslant 1$. Using the fact that X has measurable trajectories, we can apply this estimate to show that $\mathbb{E}[\|\varphi X\|^{2p}_{B^\alpha_{2p,2p}}] < \infty$ for all $p \geqslant 1$, $\alpha < H$. Now it suffices to apply Besov embedding, Lemma A.2, to obtain that $\varphi X \in \mathscr{C}^\alpha$.

Moreover, φX has compact support. So by Theorem 15.45 of [FV10], for every $p \geqslant 1$, the iterated integrals $\int_s^t \int_s^{r_2} \mathrm{d}\psi^\varepsilon * (\varphi X)(r_1)\mathrm{d}\psi^\varepsilon * (\varphi X)(r_2)$ converge in L^p in the sense of (18). The statement then follows from Corollary 3.9. □

Remark 3.11. The proof of Corollary 3.9 actually shows more than the equivalence of the convergence of A^ε and of $u^\varepsilon \circ \partial_t v^\varepsilon$: it shows that the norm of $(u^\varepsilon \circ \partial_t v^\varepsilon - \eta)$ can be controlled by a polynomial of the norms of $(A^\varepsilon - A)$, $(u^\varepsilon - u)$, and $(v^\varepsilon - v)$. So in fact we have L^p–convergence in Corollary 3.10, and not just convergence in probability. Alternatively, the L^p–convergence is obtained from the convergence in probability because we are considering random variables living in a fixed Gaussian chaos, see Theorem 3.50 of [Jan97].

Combining Corollary 3.10 with Theorem 3.3, we obtain the following corollary:

Corollary 3.12. *Let X be a n–dimensional centered Gaussian process satisfying the conditions of Corollary 3.9 for some $H > 1/3$, and let $\varphi \in \mathscr{D}$ and $F \in C^3_b$. Then there exists a unique solution u to*

$$\partial_t u = F(u)\partial_t(\varphi X), \qquad u(0) = u_0,$$

in the following sense: If $\psi \in \mathscr{S}$ with $\int \psi \mathrm{d}t = 1$ and if for $\varepsilon > 0$ the function u^ε solves

$$\partial_t u^\varepsilon = F(u^\varepsilon)\partial_t(\varphi X)^\varepsilon, \qquad u(0) = u_0,$$

where $(\varphi X)^\varepsilon = \varepsilon^{-1}\psi(\varepsilon\cdot) * (\varphi X)$, then u^ε converges to u in probability in $\mathscr{C}^\alpha$ for all $\alpha < H$.

4 Rough Burgers equation

Fix now $\sigma > 5/6$ and consider the following PDE on $[0,T] \times \mathbb{T}$ for some fixed $T > 0$:

$$Lu = G(u)\partial_x u + \xi, \qquad u(0) = u_0, \tag{21}$$

where $L = \partial_t + (-\Delta)^\sigma$ and $u_0 \in \mathscr{C}^\alpha$ for a suitable α. We would like to consider solutions u in the case of a distributional ξ, and in particular we want to allow ξ to be a typical realization of a space-time white noise. We will see below that in this case the solution ϑ to the linear equation $L\vartheta = \xi$, $\vartheta(0) = 0$, belongs (locally in time) to $\mathscr{C}^\alpha(\mathbb{T})$ for any $\alpha < \sigma - 1/2$, but it is not better than that. This is also the regularity to be expected from the solution u of the non-linear problem (21), and so for $\sigma \leqslant 1$ the term $G(u(t))\partial_x u(t)$ is not well defined since $G(u(t)) \in \mathscr{C}^\alpha(\mathbb{T})$ and $\partial_x u(t) \in \mathscr{C}^{\alpha-1}(\mathbb{T})$, and the sum of their regularities fails to be positive.

For $\sigma = 1$, this equation has been solved by Hairer [Hai11], who used rough path integrals to define the product $G(u)\partial_x u$. In the following, we will show how to solve the equation using paracontrolled distributions.

While in general it is possible to set up the equation in a space-time Besov space, the fact that the distribution ξ (which is a genuine space-time distribution) enters the problem linearly allows for a small simplification. Indeed, if we let $w = u - \vartheta$, then w solves the PDE

$$Lw = G(\vartheta + w)\partial_x(\vartheta + w), \tag{22}$$

which can be studied as an evolution equation for a continuous function of time with values in a suitable Hölder-Besov space:

Recall that for $T > 0$ and $\beta \in \mathbb{R}$ we defined the spaces $C_T\mathscr{C}^\beta = C([0,T], \mathscr{C}^\beta(\mathbb{T}^d, \mathbb{R}^n))$ with norm $\|u\|_{C_T\mathscr{C}^\beta} = \sup_{0\leqslant s\leqslant T} \|u(s)\|_\beta$. By the regularity theory for L we expect $w \in C_T\mathscr{C}^{\alpha-1+2\sigma}$ whenever $G(\vartheta + w)\partial_x(\vartheta + w) \in C_T\mathscr{C}^{\alpha-1}$ (at least in the sense of uniform estimates as the regularization goes to zero). The paraproduct allows us to decompose the right hand side of (22) as

$$G(\vartheta + w)\partial_x(\vartheta + w) = G(\vartheta + w) \prec \partial_x\vartheta + G(\vartheta + w) \circ \partial_x\vartheta + G(\vartheta + w) \succ \partial_x\vartheta + G(\vartheta + w)\partial_x w,$$

where we have expanded only the term containing $\partial_x\vartheta$ since the one linear in $\partial_x w$ is well defined under the hypothesis that $w \in C_T\mathscr{C}^{\alpha-1+2\sigma}$. Note that here we only let the paraproduct act on the spatial variables, i.e. $G(\vartheta + w) \prec \partial_x\vartheta$ should really be understood as

$$t \mapsto G(\vartheta(t) + w(t)) \prec \partial_x\vartheta(t),$$

an element of $C_T\mathscr{C}^{\alpha-1}$. A simple modification of the proof of Lemma 2.6 shows that, for $\alpha \in (0, 1/2)$, we have

$$\|G(\vartheta+w) - G'(\vartheta+w) \prec \vartheta\|_{2\alpha} \lesssim \|G\|_{C_b^2}(1+\|\vartheta\|_\alpha^2)(1+\|w\|_{2\alpha}) \lesssim \|G\|_{C_b^2}(1+\|\vartheta\|_\alpha^2)(1+\|w\|_{\alpha-1+2\sigma}),$$

where we used that $\alpha - 1 + 2\sigma > 2\alpha$, which holds because $\alpha < \sigma - 1/2 < 2\sigma - 1$. The linear dependence on the norm of w will be crucial for obtaining global solutions. We can now rewrite

$$G(\vartheta + w) \circ \partial_x\vartheta = (G(\vartheta + w) - G'(\vartheta + w) \prec \vartheta) \circ \partial_x\vartheta + C(G'(\vartheta + w), \vartheta, \partial_x\vartheta) + G'(\vartheta + w)(\vartheta \circ \partial_x\vartheta).$$

So if we assume that $(\vartheta \circ \partial_x\vartheta) \in C_T\mathscr{C}^{2\alpha-1}$, then we have a well behaved representation of the resonant term $G(\vartheta + w) \circ \partial_x\vartheta$, and

$$\begin{aligned}\|G(\vartheta + w) \circ \partial_x\vartheta\|_{2\alpha-1} &\lesssim \|G\|_{C_b^2}(1 + \|\vartheta\|_\alpha^2)(1 + \|w\|_{\alpha-1+2\sigma})\|\partial_x\vartheta\|_{\alpha-1} + \|G'(\vartheta + w)\|_\alpha\|\vartheta\|_\alpha^2 \\ &\quad + \|G'(\vartheta + w)\|_\alpha\|\vartheta \circ \partial_x\vartheta\|_{2\alpha-1} \lesssim C_G C_\vartheta(1 + \|w\|_{\alpha-1+2\sigma}),\end{aligned} \tag{23}$$

where we set

$$C_G = \|G\|_{C_b^2} \qquad \text{and} \qquad C_\vartheta = \sup_{t\in[0,T]} (1 + \|\vartheta(t)\|_\alpha)^3 (1 + \|\vartheta(t) \circ \partial_x \vartheta(t)\|_{2\alpha-1}).$$

Let us now define

$$\Phi = G(\vartheta + w)\partial_x \vartheta = G(\vartheta + w) \prec \partial_x \vartheta + G(\vartheta + w) \succ \partial_x \vartheta + G(\vartheta + w) \circ \partial_x \vartheta,$$

so that (23) and the paraproduct estimates yield

$$\|\Phi\|_{\alpha-1} \lesssim C_G C_\vartheta (1 + \|w\|_{\alpha-1+2\sigma}), \tag{24}$$

and w satisfies $Lw = \Phi + G(\vartheta + w)\partial_x w$. So if we denote by $(P_t)_{t\geqslant 0}$ the semigroup generated by $-(-\Delta)^\sigma$, then

$$w(t) = P_t u_0 + \int_0^t P_{t-s}\Phi(s)\mathrm{d}s + \int_0^t P_{t-s}(G(\vartheta(s) + w(s))\partial_x w(s))\mathrm{d}s, \tag{25}$$

where we assumed that $\vartheta(0) = 0$. Applying the Schauder estimates for the fractional Laplacian (Lemma A.9 and Lemma A.7) to (25), we obtain for all $t > 0$ that

$$\begin{aligned} \|w(t)\|_{\alpha-1+2\sigma} &= \left\| P_t u_0 + \int_0^t P_{t-s}\Phi(s)\mathrm{d}s + \int_0^t P_{t-s}(G(\vartheta(s) + w(s))\partial_x w(s))\mathrm{d}s \right\|_{\alpha-1+2\sigma} \\ &\leqslant c t^{-(2\sigma-1)/2\sigma} \Big(\|u_0\|_\alpha + \sup_{s\in[0,t]} (s^{(2\sigma-1)/(2\sigma)} \|\Phi(s)\|_{\alpha-1}) \Big) \\ &\quad + c \int_0^t \frac{\|G(\vartheta(s) + w(s))\partial_x w(s)\|_{L^\infty}}{(t-s)^{(\alpha-1+2\sigma)/(2\sigma)}} \mathrm{d}s, \end{aligned}$$

where $c > 0$ is a generic constant whose value may change in every step. But now recall from (24) that $\|\Phi(s)\|_{\alpha-1} \lesssim C_G C_\vartheta (1 + \|w(s)\|_{\alpha-1+2\sigma})$. Moreover, if we choose $\alpha \in (1/3, \sigma - 1/2)$ close enough to $\sigma - 1/2$, then $\alpha + 2\sigma - 2 > 0$ (recall that $\sigma > 5/6$), and therefore

$$\|G(\vartheta(s) + w(s))\partial_x w(s)\|_{L^\infty} \lesssim \|G\|_{L^\infty} \|\partial_x w(s)\|_{\alpha-2+2\sigma} \lesssim \|G\|_{L^\infty} \|w(s)\|_{\alpha-1+2\sigma}.$$

Thus, we get for all $t \in [0, T]$ that

$$\begin{aligned} (t^{1-1/(2\sigma)} \|w(t)\|_{\alpha-1+2\sigma}) &\leqslant c\|u_0\|_\alpha + c C_\vartheta C_G (1 + \sup_{s\in[0,t]} (s^{(2\sigma-1)/(2\sigma)} \|w(s)\|_{\alpha-1+2\sigma})) \\ &\quad + c C_G t^{1-1/(2\sigma)} \int_0^t (s^{1-1/(2\sigma)} \|w(s)\|_{\alpha+1}) \frac{\mathrm{d}s}{(t-s)^{(\alpha-1+2\sigma)/(2\sigma)} s^{1-1/(2\sigma)}} \end{aligned}$$

Since $(\alpha - 1 + 2\sigma)/(2\sigma) < 1$, we have

$$t^{1-1/(2\sigma)} \int_0^t \frac{\mathrm{d}s}{(t-s)^{(\alpha-1+2\sigma)/(2\sigma)} s^{1-1/(2\sigma)}} \lesssim t^{1-(\alpha-1+2\sigma)/(2\sigma)} \lesssim 1$$

for $t \in [0, T]$. Putting everything together, we conclude that

$$(t^{1-1/(2\sigma)} \|w(t)\|_{\alpha-1+2\sigma}) \leqslant c\|u_0\|_\alpha + c C_\vartheta C_G (1 + \sup_{s\in[0,t]} (s^{(2\sigma-1)/(2\sigma)} \|w(s)\|_{\alpha-1+2\sigma})).$$

In order to turn this into a bound on $\|w\|_{C_T \mathscr{C}^{\alpha-1+2\sigma}}$, we use again a scaling argument. We extend the scaling transformation to the time variable in such a way that it leaves the operator L invariant. More precisely, for $\lambda > 0$ we set $\Lambda_\lambda u(t, x) = u(\lambda^{2\sigma} t, \lambda x)$, so that $L\Lambda_\lambda = \lambda^{2\sigma} \Lambda_\lambda L$.

Now let $u^\lambda = \Lambda_\lambda u$, $w^\lambda = \Lambda_\lambda w$, and $\vartheta^\lambda = \Lambda_\lambda \vartheta$. Note that $u^\lambda \colon [0, T/\lambda^{2\sigma}] \times \mathbb{T}_\lambda \to \mathbb{R}$, where $\mathbb{T}_\lambda = \mathbb{R}/(2\pi\lambda^{-1}\mathbb{Z})$ is a rescaled torus, and that w^λ solves the equation

$$Lw^\lambda = \lambda^{2\sigma}\Lambda_\lambda Lw = \lambda^{2\sigma}\Lambda_\lambda(\Phi + G(w+\vartheta)\partial_x w) = \lambda^{2\sigma}\Lambda_\lambda\Phi + \lambda^{2\sigma-1}G(w^\lambda + \vartheta^\lambda)\partial_x w^\lambda.$$

The same derivation as above shows that

$$\|\Lambda_\lambda \Phi(t)\|_{\alpha-1} = \|G(\vartheta^\lambda(t) + w^\lambda(t))\Lambda_\lambda(\partial_x\vartheta)(t)\|_{\alpha-1} \lesssim C_G C_{\vartheta^\lambda}(1 + \|w^\lambda(t)\|_{\alpha-1+2\sigma}),$$

where we get using Lemma A.4 and Lemma B.1

$$C_{\vartheta^\lambda} = \sup_{t\in[0,T]} (1 + \|\vartheta^\lambda(t)\|_\alpha)^3 (1 + \|\vartheta^\lambda \circ \Lambda_\lambda(\partial_x\vartheta)(t)\|_{2\alpha-1}) \lesssim \lambda^{2\alpha-1}C_\vartheta^2 \leqslant \lambda^{-1}C_\vartheta^2$$

as long as $\lambda \in (0,1]$. Thus, we finally conclude that

$$\begin{aligned}(t^{1-1/(2\sigma)}\|w^\lambda(t)\|_{\alpha-1+2\sigma}) &\leqslant c\|\Lambda_\lambda u_0\|_\alpha + \lambda^{2\sigma-1} c C_\vartheta^2 C_G (1 + \sup_{s\in[0,t]}(s^{(2\sigma-1)/(2\sigma)}\|w^\lambda(s)\|_{\alpha-1+2\sigma})) \\ &\leqslant c\|u_0\|_\alpha + \lambda^{2\sigma-1} c C_\vartheta^2 C_G (1 + \sup_{s\in[0,t]}(s^{(2\sigma-1)/(2\sigma)}\|w^\lambda(s)\|_{\alpha-1+2\sigma}))\end{aligned}$$

for all $\lambda \in (0,1]$. Since $2\sigma - 1 > 0$, we get for small enough $\lambda > 0$, depending only on C_ϑ and C_G but not on u_0, that

$$\sup_{t\in[0,T]} (t^{1-1/(2\sigma)}\|w^\lambda(t)\|_{\alpha-1+2\sigma}) \leqslant 2(c\|u_0\|_\alpha + 1).$$

But $u = \Lambda_{\lambda^{-1}}(w^\lambda + \vartheta^\lambda)$, and therefore

$$\sup_{t\in[0,\lambda^{2\sigma}T]} \|u(t)\|_\alpha \lesssim_\lambda \|u_0\| + C_\vartheta.$$

This provides the key ingredient for obtaining a uniform estimate on the full time interval $[0,T]$, and then the existence of global solutions to the Burgers equation.

Uniqueness in the space of solutions u with decomposition $u = \vartheta + w$ with $w \in C_T\mathscr{C}^{\alpha-1+2\sigma}$ can be handled easily along the lines above, and we obtain the following result:

Theorem 4.1. *Let $\sigma > 5/6$, $\alpha \in (1/3, \sigma - 1/2)$, let $T > 0$, and assume that $(\xi^\varepsilon)_{\varepsilon>0}$ is a family of smooth functions on $[0,T]\times\mathbb{T}$ with values in $\mathbb{R}^n$, and $G \in C_b^3(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n,\mathbb{R}^n))$. Suppose that there exist $\vartheta \in C_T\mathscr{C}^\alpha$ and $\eta \in C_T\mathscr{C}^{2\alpha-1}$ such that $(\vartheta^\varepsilon, (\vartheta^\varepsilon \circ \partial_x\vartheta^\varepsilon))$ converges to (ϑ,η) in $C_T\mathscr{C}^{\alpha-1}\times C_T\mathscr{C}^{2\alpha-1}$, where ϑ^ε are solutions to $L\vartheta^\varepsilon = \xi^\varepsilon$ and $\vartheta^\varepsilon(0) = 0$, and where $L = \partial_t + (-\Delta)^\sigma$. Let for $\varepsilon > 0$ the function u^ε be the unique global solution to the Cauchy problem*

$$Lu^\varepsilon = G(u^\varepsilon)\partial_x u^\varepsilon + \xi^\varepsilon, \qquad u^\varepsilon(0) = u_0,$$

where $u_0 \in \mathscr{C}^\alpha$. Then there exists $u \in C_T\mathscr{C}^\alpha$ such that $u^\varepsilon \to u$ in $C_T\mathscr{C}^\alpha$. The limit u depends only on (u_0,ϑ,η), and not on the approximating family $(\vartheta^\varepsilon, (\vartheta^\varepsilon\circ\partial_x\vartheta^\varepsilon))$.

Remark 4.2. As for RDEs, the limit u of the regularized solutions u^ε actually solves the equation

$$Lu = G(u)\partial_x u + \xi, \qquad u(0) = u_0$$

in the weak sense as long as we interpret the product $G(u)\partial_x u$ correctly. According to Remark 3.8, it is not important that $\vartheta(0) = 0$, and we could consider any other initial condition in $\vartheta(0) \in \mathscr{C}^\gamma$ for some $\gamma \in \mathbb{R}$ to obtain the same solution u. However, of course the right choice of $\vartheta(0)$ may facilitate the proof of existence and uniqueness of paracontrolled solutions.

Remark 4.3. Of course, the solution u to the fractional Burgers type equation also depends continuously on the initial condition u_0.

4.1 Construction of the area

It remains to show that if ξ is a space-time white noise, then the solution ϑ to $L\vartheta = \xi$, $\vartheta(0) = 0$, is in $C_T\mathscr{C}^\alpha$ for all $\alpha < \sigma - 1/2$, and that the area $\vartheta \circ \partial_x \vartheta$ is in $C_T\mathscr{C}^{2\alpha-1}$. Some general results on the existence of the area for Gaussian processes indexed by a one-dimensional spatial variable are shown in [FGGR12]. However, in the present setting it is relatively straightforward to construct the area "by hand", using Fourier analytic methods.

In this section, we use $\mathscr{F}$ to denote the spatial Fourier transform, i.e. $\mathscr{F}u(t,\cdot)(k) = \int_{\mathbb{T}} e^{-\imath kx} u(t,x)\mathrm{d}x$. Recall that $\mathscr{F}\xi$ is a complex valued, centered Gaussian space-time distribution, whose covariance is formally given by

$$\mathbb{E}[\mathscr{F}\xi^i(t,\cdot)(k)\overline{\mathscr{F}\xi^{i'}(t',\cdot)(k')}] = 2\pi \mathbf{1}_{i=i'}\mathbf{1}_{k=k'}\delta(t-t')$$

for $i, i' \in \{1,\dots,n\}$, $t,t' \in [0,T]$, $k,k' \in \mathbb{Z}$, where δ denotes the Dirac delta. If $(P_t)_{t\geqslant 0} = (e^{-t|\cdot|^{2\sigma}}(\mathrm{D}))_{t\geqslant 0}$ denotes the semigroup generated by $-(-\Delta)^\sigma$, then $\vartheta(t,x) = \int_0^t (P_{t-s}\xi)(x)\mathrm{d}s$, $t \in [0,T]$, from where a straightforward calculation yields the following result:

Lemma 4.4. *The spatial Fourier transform $\mathscr{F}\vartheta$ of ϑ is a complex-valued Gaussian process with zero mean and covariance*

$$\mathbb{E}[\mathscr{F}\vartheta^i(t,\cdot)(k)\overline{\mathscr{F}\vartheta^{i'}(t',\cdot)(k')}] = \begin{cases} 2\pi\mathbf{1}_{i=i'}\mathbf{1}_{k=k'}(e^{-|t'-t||k|^{2\sigma}} - e^{-(t+t')|k|^{2\sigma}})/(2|k|^{2\sigma}), & k \neq 0, \\ 2\pi\mathbf{1}_{i=i'}\mathbf{1}_{k=k'}t\wedge t', & k = 0, \end{cases}$$

for $i,i' \in \{1,\dots,n\}$, $k,k' \in \mathbb{Z}$, and $t,t' \in [0,T]$. As a consequence, $\mathbb{E}[\mathscr{F}\vartheta^i_{s,t}(0)\overline{\mathscr{F}\vartheta^i_{s,t}(k')}] = 2\pi\mathbf{1}_{i=i'}\mathbf{1}_{k'=0}|t-s|$, and for $k \neq 0$

$$\mathbb{E}[\mathscr{F}\vartheta^i_{s,t}(k)\overline{\mathscr{F}\vartheta^{i'}_{s,t}(k')}] = \pi\mathbf{1}_{i=i'}\mathbf{1}_{k=k'}(2 - e^{-2s|k|^{2\sigma}} - e^{-2t|k|^{2\sigma}} - 2e^{-2|t-s||k|^{2\sigma}} + 2e^{-(s+t)|k|^{2\sigma}})/|k|^{2\sigma},$$

where we write $\mathscr{F}\vartheta^i_{s,t}(k) = \mathscr{F}\vartheta^i(t,\cdot)(k) - \mathscr{F}\vartheta^i(s,\cdot)(k)$ for all $0 \leqslant s < t \leqslant T$. In particular,

$$|\mathbb{E}[\mathscr{F}\vartheta^i_{s,t}(k)\overline{\mathscr{F}\vartheta^i_{s,t}(k)}]| \lesssim |t-s|^\delta |k|^{-2\sigma(1-\delta)} \tag{26}$$

for all $\delta \in [0,1]$ and all $k \neq 0$.

Our first concern is to study the Hölder-Besov regularity of the process ϑ.

Lemma 4.5. *For any $\alpha < \sigma - 1/2$ and any $p \geqslant 1$, the process ϑ satisfies*

$$\mathbb{E}[\|\vartheta\|^p_{C_T\mathscr{C}^\alpha(\mathbb{T})}] < \infty.$$

Proof. Let $s,t \in [0,T]$ and $\ell \geqslant 0$. The case $\ell = 0$ can be treated using essentially the same arguments, except that then we need to distinguish the cases $k = 0$ and $k \neq 0$, where k is the argument in the Fourier transform. Using Gaussian hypercontractivity ([Jan97], Theorem 3.50), we obtain for $p \geqslant 1$ that

$$\mathbb{E}[\|\Delta_\ell\vartheta_{s,t}\|^{2p}_{L^{2p}(\mathbb{T})}] \lesssim_p \|\mathbb{E}[|\Delta_\ell\vartheta_{s,t}(x)|^2]\|^p_{L^p_x(\mathbb{T})}. \tag{27}$$

Using Fourier inversion and Lemma 4.4, we have

$$\begin{aligned}\mathbb{E}[|\Delta_\ell\vartheta_{s,t}(x)|^2] &= (2\pi)^{-2}\sum_{k,k'\in\mathbb{Z}} \rho_\ell(k)\rho_\ell(k')e^{\imath(k-k')x}\mathbb{E}[\mathscr{F}\vartheta_{s,t}(k)\overline{\mathscr{F}\vartheta_{s,t}(k')}] \\ &\lesssim \sum_{k\in\mathbb{Z}}\rho_\ell^2(k)|t-s|^\delta|k|^{2\sigma(\delta-1)} \lesssim |t-s|^\delta \sum_{k\in\mathrm{supp}(\rho_\ell)}|k|^{2\sigma(\delta-1)} \lesssim |t-s|^\delta 2^{\ell(1-2\sigma(1-\delta))}\end{aligned}$$

for all $\delta \in (0,1]$. Hence, we obtain from (27) that

$$\mathbb{E}[\|\vartheta(t,\cdot)-\vartheta(s,\cdot)\|_{B^{\alpha}_{2p,2p}(\mathbb{T})}^{2p}] \lesssim \sum_{\ell\geqslant -1} 2^{\ell\alpha 2p}\mathbb{E}[\|\Delta_\ell \vartheta_{s,t}\|_{L^{2p}(\mathbb{T})}^{2p}] \lesssim \sum_{\ell \geqslant -1} 2^{\ell\alpha 2p}\Big(|t-s|^{\delta} 2^{2\ell(1/2-\sigma(1-\delta))}\Big)^{p}$$

for any $\alpha \in \mathbb{R}$ and any $p \geqslant 1$. For $\alpha < \sigma - 1/2$ there exists $\delta \in (0,1]$ small enough so that the series converges. Since we can choose p arbitrarily large, Kolmogorov's continuity criterion implies that ϑ has a continuous version with $\mathbb{E}[\|\vartheta\|_{C_T B^{\alpha}_{2p,2p}(\mathbb{T})}^{2p}] < \infty$ for all $\alpha < \sigma - 1/2$. Now we use again that p can be chosen arbitrarily large, so that the Besov embedding theorem, Lemma A.2, shows that this continuous version takes its values in $C_T\mathscr{C}^{\alpha}(\mathbb{T})$ for all $\alpha < \sigma - 1/2$. □

Next, we construct the area $\vartheta \circ \partial_x \vartheta$.

Lemma 4.6. *Define*

$$\vartheta \circ \partial_x \vartheta = (\vartheta^k \circ \partial_x \vartheta^\ell)_{1\leqslant k,\ell\leqslant n} = \Big(\sum_{|i-j|\leqslant 1} \Delta \vartheta^i \Delta_j \partial_x \vartheta^j\Big)_{1\leqslant k,\ell\leqslant n}.$$

Then almost surely $\vartheta \circ \partial_x \vartheta \in C_T\mathscr{C}^{2\alpha-1}(\mathbb{T};\mathbb{R}^{n\times n})$ for all $\alpha < \sigma - 1/2$. Moreover, if $\psi \in \mathscr{S}$ is such that $\int \psi(x)\mathrm{d}x = 1$ and $\vartheta^\varepsilon = \psi^\varepsilon * \vartheta$, where $\psi^\varepsilon = \varepsilon^{-1}\psi(\varepsilon^{-1}\cdot)$, then we have for all $p \geqslant 1$ that

$$\lim_{\varepsilon\to 0} \mathbb{E}[\|\vartheta^\varepsilon \circ \partial_x \vartheta^\varepsilon - \vartheta \circ \partial_x \vartheta\|_{C_T\mathscr{C}^{2\alpha-1}}^{p}] = 0. \tag{28}$$

Proof. Without loss of generality we can argue for $\vartheta^1 \circ \partial_x \vartheta^2$. The case $\vartheta^1 \circ \partial_x \vartheta^1$ is easy, because Leibniz' rule yields $\vartheta^1 \circ \partial_x \vartheta^1 = \frac{1}{2}\partial_x(\vartheta^1 \circ \vartheta^1)$.

Let $\ell \in \mathbb{N}$. Note that if i is smaller than $\ell - N$ for a suitable N, and if $|i-j| \leqslant 1$, then $\Delta_\ell(\Delta_i f \Delta_j g) = 0$ for all $f, g \in \mathscr{S}'$. Hence, the projection of $\vartheta^1 \circ \partial_x \vartheta^2$ onto the ℓ–th dyadic Fourier block is given by

$$\Delta_\ell(\vartheta^1 \circ \partial_x \vartheta^2) = \sum_{|i-j|\leqslant 1} \Delta_\ell(\Delta_i \vartheta^1 \Delta_j \partial_x \vartheta^2) = \sum_{|i-j|\leqslant 1} \mathbf{1}_{\ell \lesssim i} \Delta_\ell(\Delta_i \vartheta^1 \Delta_j \partial_x \vartheta^2).$$

To avoid case distinctions, we only argue for $\ell \geqslant N$, so that we can always assume $i, j \geqslant 0$. The case $\ell < N$ can be handled using essentially the same arguments.

We can apply Gaussian hypercontractivity to obtain

$$\begin{aligned} &\mathbb{E}[\|(\Delta_\ell(\vartheta^1 \circ \partial_x \vartheta^2 - \vartheta^{1,\varepsilon} \circ \partial_x \vartheta^{2,\varepsilon}))_{s,t}\|_{L^{2p}(\mathbb{T})}^{2p}] \\ &\lesssim \Big\|\mathbb{E}\Big[\Big|\sum_{|i-j|\leqslant 1} \mathbf{1}_{\ell\lesssim i}(\Delta_\ell(\Delta_i \vartheta^1 \Delta_j \partial_x \vartheta^2 - \Delta_i \vartheta^{1,\varepsilon} \Delta_j \partial_x \vartheta^{2,\varepsilon})(x))_{s,t}\Big|^2\Big]\Big\|_{L^p_x(\mathbb{T})}^{p}, \end{aligned} \tag{29}$$

where we write $\vartheta^{1,\varepsilon} = \psi^\varepsilon * \vartheta$ and similarly for $\vartheta^{2,\varepsilon}$.

Let us start by estimating

$$\begin{aligned} &\mathbb{E}\Big[\Big|\sum_{|i-j|\leqslant 1} \mathbf{1}_{\ell\lesssim i}\Delta_\ell(\Delta_i \vartheta^1(t,\cdot)\Delta_j \partial_x \vartheta^2_{s,t} - \Delta_i \vartheta^{1,\varepsilon}(t,\cdot)\Delta_j \partial_x \vartheta^{2,\varepsilon}_{s,t})(x)\Big|^2\Big] \\ &\quad = \sum_{|i-j|\leqslant 1}\sum_{|i'-j'|\leqslant 1} \mathbf{1}_{\ell\lesssim i}\mathbf{1}_{\ell\lesssim i'}\mathbb{E}[\Delta_\ell(\Delta_i \vartheta^1(t,\cdot)\Delta_j \partial_x \vartheta^2_{s,t} - \Delta_i \vartheta^{1,\varepsilon}(t,\cdot)\Delta_j\partial_x\vartheta^{2,\varepsilon}_{s,t})(x) \\ &\qquad\qquad \times \overline{\Delta_\ell(\Delta_{i'}\vartheta^1(t,\cdot)\Delta_{j'}\partial_x\vartheta^2_{s,t} - \Delta_{i'}\vartheta^{1,\varepsilon}(t,\cdot)\Delta_{j'}\partial_x\vartheta^{2,\varepsilon}_{s,t})(x)}]. \end{aligned} \tag{30}$$

Taking the infinite sums outside of the expectation can be justified a posteriori, because for every finite partial sum we will obtain a bound on the L^2–norm below, which does not depend on the number of terms that we sum up. The Gaussian hypercontractivity (29) then provides a uniform L^p–bound for all $p \geqslant 2$, which implies that the squares of the partial sums are uniformly integrable, and thus allows us to exchange summation and expectation.

Recall that $\mathscr{F}(uv)(k) = (2\pi)^{-1} \sum_{k'} \mathscr{F}u(k')\mathscr{F}v(k-k')$, and $\mathscr{F}(\partial_x u)(k) = \imath k \mathscr{F}(u)(k)$, and therefore

$$
\begin{aligned}
&\Delta_\ell(\Delta_i \vartheta^1(t,\cdot)\Delta_j \partial_x \vartheta^2_{s,t} - \Delta_i \vartheta^{1,\varepsilon}(t,\cdot)\Delta_j\partial_x \vartheta^{2,\varepsilon}_{s,t})(x) \\
&= (2\pi)^{-1} \sum_{k\in\mathbb{Z}} \rho_\ell(k) e^{\imath\langle k,x\rangle} \mathscr{F}(\Delta_i \vartheta^1(t,\cdot)\Delta_j\partial_x\vartheta^2_{s,t} - \Delta_i\vartheta^{1,\varepsilon}(t,\cdot)\Delta_j\partial_x\vartheta^{2,\varepsilon}_{s,t})(k) \\
&= (2\pi)^{-2} \sum_{k,k'\in\mathbb{Z}} \rho_\ell(k)e^{\imath\langle k,x\rangle}\rho_i(k')\rho_j(k-k')\imath(k-k')\mathscr{F}\vartheta^1(t,\cdot)(k')\mathscr{F}\vartheta^2_{s,t}(k-k') \\
&\qquad \times(1-\mathscr{F}\psi(\varepsilon k')\mathscr{F}\psi(\varepsilon(k-k'))).
\end{aligned}
$$

From this expression it is clear that if we can show $\mathbb{E}[\|\vartheta^\varepsilon \circ \partial_x \vartheta^\varepsilon\|^p_{C_T\mathscr{C}^{2\alpha-1}}] < \infty$, then the convergence result in (28) will follow by dominated convergence, because $\mathscr{F}\psi$ is bounded and $\mathscr{F}\psi(0) = 1$ by assumption.

Using the covariance of $\mathscr{F}\vartheta$ that we calculated in Lemma 4.4, we obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\Big|\sum_{|i-j|\leqslant 1} \mathbf{1}_{\ell\lesssim i}\Delta_\ell(\Delta_i\vartheta^1(t,\cdot)\Delta_j\partial_x\vartheta^2_{s,t})(x)\Big|^2\Big] \\
&\quad\lesssim \sum_{|i-j|\leqslant 1}\sum_{|i'-j'|\leqslant 1} \mathbf{1}_{\ell\lesssim i}\mathbf{1}_{\ell\lesssim i'} \sum_{k,k'\in\mathbb{Z}^d} \rho^2_\ell(k+k')\rho_i(k)\rho_{i'}(k)\rho_j(k')\rho_{j'}(k') \\
&\qquad\qquad \times \frac{1-e^{-2t|k|^{2\sigma}}}{2|k|^{2\sigma}}|k'|^2|t-s|^\delta|k'|^{-2\sigma(1-\delta)} \\
&\quad\lesssim \sum_{|i-j|\leqslant 1}\mathbf{1}_{\ell\lesssim i}\sum_{k\in\operatorname{supp}(\rho_i),k'\in\operatorname{supp}(\rho_j)} \rho^2_\ell(k+k')2^{2i(1-2\sigma+\sigma\delta)}|t-s|^\delta \\
&\quad\lesssim \sum_{i\gtrsim\ell} 2^\ell 2^{2i(1+1/2-2\sigma+\sigma\delta)}|t-s|^\delta
\end{aligned}
$$

for all $\delta\in[0,1]$. Since $\sigma > 5/6$, there exists $\delta > 0$ small enough so that the sum is finite, and we obtain

$$
\mathbb{E}\Big[\Big|\sum_{|i-j|\leqslant 1}\mathbf{1}_{\ell\lesssim i}\Delta_\ell(\Delta_i\vartheta^1(t,\cdot)\Delta_j\partial_x\vartheta^2_{s,t})(x)\Big|^2\Big] \lesssim 2^{2i(2-2\sigma+\sigma\delta)}|t-s|^\delta,
$$

and by the same arguments

$$
\mathbb{E}\Big[\Big|\sum_{|i-j|\leqslant 1}\mathbf{1}_{\ell\lesssim i}\Delta_\ell(\Delta_i\vartheta^1_{s,t}\Delta_j\partial_x\vartheta^2(s,\cdot))(x)\Big|^2\Big] \lesssim 2^{2i(2-2\sigma+\sigma\delta)}|t-s|^\delta.
$$

Since

$$
\Delta_i\vartheta^1(t,\cdot)\Delta_j\partial_x\vartheta^2(t,\cdot) - \Delta_i\vartheta^1(s,\cdot)\Delta_j\partial_x\vartheta^2(s,\cdot) = \Delta_i\vartheta^1(t,\cdot)\Delta_j\partial_x\vartheta^2_{s,t} + \Delta_i\vartheta^1_{s,t}\Delta_j\partial_x\vartheta^2(s,\cdot),
$$

we get for sufficiently small $\delta > 0$ and for arbitrarily large $p\geqslant 1$ that

$$
\mathbb{E}[\|\Delta_\ell(\vartheta^1\circ\partial_x\vartheta^2)_{s,t}\|^{2p}_{L^{2p}(\mathbb{T})}] \lesssim 2^{-2\ell(2\sigma-2-\sigma\delta)p}|t-s|^{\delta p}.
$$

Now we use the same arguments as in the proof of Lemma 4.5 to obtain the required L^p-bound for $\|\vartheta^1\circ\partial_x\vartheta^2\|_{C_T\mathscr{C}^{2\alpha-1}}$ with $\alpha<\sigma-1/2$. $\square$

Now Lemma 4.6 and Theorem 4.1 give us the existence and uniqueness of solutions to the fractional Burgers type equation driven by space-time white noise:

Corollary 4.7. *Let $\sigma > 5/6$, $\alpha \in (1/3, \sigma - 1/2)$, $T > 0$, $G \in C_b^3$, $u_0 \in \mathscr{C}^\alpha(\mathbb{T})$, $L = \partial_t + (-\Delta)^\sigma$, and let ξ be a space-time white noise on $[0,T] \times \mathbb{T}$ with values in $\mathbb{R}^n$. Then there exists a unique solution u to*

$$Lu = G(u)\partial_x u + \xi, \qquad u(0) = u_0,$$

in the following sense: If $\psi \in \mathscr{S}$ with $\int \psi \mathrm{d}t = 1$ and if for $\varepsilon > 0$ the function u^ε solves

$$Lu^\varepsilon = G(u^\varepsilon)\partial_x u^\varepsilon + \xi^\varepsilon, \qquad u(0) = u_0,$$

where $\xi^\varepsilon = \varepsilon^{-1}\psi(\varepsilon \cdot) * \xi$, then u^ε converges in probability in $C_T\mathscr{C}^\alpha$ to u.

Remark 4.8. There is no problem in considering the equation on $\mathbb{T}$ rather than on $\mathbb{T}^d$, and the analysis works exactly as in the one-dimensional case. The proof of Lemma 4.5 shows that if ξ is a space-time white noise on $[0,T] \times \mathbb{T}^d$, then the solution ϑ to $L\vartheta = \xi$, $\vartheta(0) = 0$, will be in $C_T\mathscr{C}^\alpha(\mathbb{T}^d)$ for every $\alpha < \sigma - d/2$. So as long as $\sigma - d/2 > 1/3$, we can solve the Burgers equation on $\mathbb{T}^d$. For the existence of the area $\vartheta \circ \partial_x \vartheta$ we need the additional condition $2\sigma - d/2 - 1 > 0$; see [Per14], Lemma 5.4.3. But if $\sigma - d/2 > 1/3$, then this is always satisfied.

5 A generalized parabolic Anderson model

Consider now the following PDE on $[0,T] \times \mathbb{T}^2$ for some fixed $T > 0$:

$$Lu = F(u)\xi, \qquad u(0) = u_0, \tag{31}$$

where $L = \partial_t - \Delta$, the function F is continuous from $\mathbb{R}$ to $\mathbb{R}$, ξ is a spatial white noise, and $u_0 \in \mathscr{C}^\alpha$ for suitable $\alpha \in \mathbb{R}$. Formally, this equation is very similar to the rough differential equation (12).

The regularity of the spatial white noise η on $\mathbb{T}^d$ is $\eta \in \mathscr{C}^{-d/2-\varepsilon}$ for all $\varepsilon > 0$. Since we are in dimension $d = 2$, we have $\xi \in \mathscr{C}^{-1-\varepsilon}$. The Laplacian increases the regularity by 2, so we expect that for fixed $t > 0$ we have $u(t) \in \mathscr{C}^{1-\varepsilon}$, and therefore the product $F(u)\xi$ is ill-defined.

However, let us assume that $\xi \in \mathscr{C}^{\alpha-2}(\mathbb{T}^2)$ for some $2/3 < \alpha < 1$, and let $\vartheta \in \mathscr{C}^\alpha$ be such that $L\vartheta = \xi$. Consider the paracontrolled ansatz

$$u = F(u) \prec \vartheta + u^\sharp$$

with $u^\sharp \in C_T\mathscr{C}^{2\alpha}$, and where as in Section 4 the paraproduct $\prec$ is only acting on the spatial variables. If u is of this form, then Lemma 2.7 and Lemma 2.4 imply that

$$F(u)\xi = F(u) \prec \xi + F(u) \succ \xi + F'(u)F(u)(\vartheta \circ \xi) + F'(u)C(F(u), \vartheta, \xi) + F'(u)(u^\sharp \circ \xi) + \Pi_F(u, \xi)$$

is well defined provided that $(\vartheta \circ \xi) \in C_T\mathscr{C}^{2\alpha-2}$. Moreover, the algebraic rules for ∂_t and Δ acting on products imply that

$$Lu = (LF(u)) \prec \vartheta + F(u) \prec L\vartheta - 2\mathrm{D}_x F(u) \prec \mathrm{D}_x \vartheta + Lu^\sharp,$$

and thus we find the following equation for $u^\sharp$:

$$\begin{aligned} Lu^\sharp = {} & 2\mathrm{D}_x F(u) \prec \mathrm{D}_x\vartheta - (LF(u)) \prec \vartheta + F(u) \succ \xi + F'(u)F(u)(\vartheta \circ \xi) \\ & + F'(u)C(F(u), \vartheta, \xi) + F'(u)(u^\sharp \circ \xi) + \Pi_F(u, \xi). \end{aligned}$$

We would like all the terms on the right hand side to be in $C_T\mathscr{C}^{2\alpha-2}$. However, it is not easy to estimate the term $(LF(u)) \prec \vartheta$ in $C_T\mathscr{C}^\beta$ for any $\beta \in \mathbb{R}$: the term $\Delta F(u)$ can be controlled in $\mathscr{C}^{\alpha-2}$, but there are no straightforward estimates available for the time derivative $\partial_t F(u)$ appearing in $LF(u)$. Indeed, it would be more convenient to treat the generalized parabolic Anderson model in a space-time parabolic Besov space adapted to the operator L and to use the natural paraproduct associated to this space. An alternative strategy would be to stick with the simpler space $C_T\mathscr{C}^{\alpha-2}$ and to observe that

$$LF(u) = F'(u)Lu - F''(u)(\mathrm{D}_x u)^2 = F'(u)F(u)\xi - F''(u)(\mathrm{D}_x u)^2,$$

and that the terms on the right hand side can be analyzed using the paracontrolled ansatz. Since this strategy seems to require a lot of regularity from F, we do not pursue it further.

Instead, we keep working on $C_T\mathscr{C}^{\alpha-2}$, but we modify the paraproduct appearing in the paracontrolled ansatz. Let $\varphi\colon \mathbb{R} \to \mathbb{R}_+$ be a positive smooth function with compact support and total mass 1, and for all $i \geqslant -1$ define the operator $Q_i\colon C_T\mathscr{C}^\beta \to C_T\mathscr{C}^\beta$ by

$$Q_i f(t) = \int_{\mathbb{R}} 2^{2i}\varphi(2^{2i}(t-s)) f((s \wedge T) \vee 0)\mathrm{d}s.$$

For Q_i we have the following standard estimates, which we leave to the reader to prove:

$$\|Q_i f(t)\|_{L^\infty} \leqslant \|f\|_{C_T L^\infty}, \quad \|\partial_t Q_i f(t)\|_{L^\infty} \leqslant 2^{(2-2\gamma)i}\|f\|_{C_T^\gamma L^\infty}, \qquad (32)$$
$$\|(Q_i f - f)(t)\|_{L^\infty} \leqslant 2^{-2\gamma i}\|f\|_{C_T^\gamma L^\infty}$$

for all $t \in [0,T]$ and all $\gamma \in (0,1)$; for the second estimate we use that $\int \varphi'(t)\mathrm{d}t = 0$, and for the third estimate we use that φ has total mass 1. With the help of Q_i, let us define a modified paraproduct by setting

$$f \prec\!\!\prec g = \sum_i (S_{i-1}Q_i f)\Delta_i g$$

for $f, g \in C_T\mathscr{S}'$. It is easy to show that for this paraproduct we have essentially the same estimates as for the pointwise paraproduct $f \prec g$, only that we have to bound f uniformly in time; for example

$$\|(f \prec\!\!\prec g)(t)\|_\alpha \lesssim \|f\|_{C_T L^\infty}\|g(t)\|_\alpha.$$

for all t $[0,T]$. For us, the following two commutator estimates are the most useful properties of $\prec\!\!\prec$.

Lemma 5.1. *Let $T > 0$, $\alpha \in (0,1)$, $\beta \in \mathbb{R}$, and let $u \in C_T\mathscr{C}^\alpha \cap C_T^{\alpha/2}L^\infty$ and $v \in C_T\mathscr{C}^\beta$. Then*

$$\|L(u \prec\!\!\prec v) - u \prec\!\!\prec (Lv)\|_{C_T\mathscr{C}^{\alpha+\beta-2}} \lesssim (\|u\|_{C_T^{\alpha/2}L^\infty} + \|u\|_{C_T\mathscr{C}^\alpha})\|v\|_{C_T\mathscr{C}^\beta}, \qquad (33)$$

as well as

$$\|u \prec v - u \prec\!\!\prec v\|_{C_T\mathscr{C}^{\alpha+\beta}} \lesssim \|u\|_{C_T^{\alpha/2}L^\infty}\|v\|_{C_T\mathscr{C}^\beta}. \qquad (34)$$

Proof. For (33), observe that $L(u \prec\!\!\prec v) - u \prec\!\!\prec (Lv) = (Lu) \prec\!\!\prec v - 2\mathrm{D}_x u \prec\!\!\prec \mathrm{D}_x v$. The second term on the right hand side is easy to estimate. The first term is given by

$$(Lu) \prec\!\!\prec v = \sum_i (S_{i-1}Q_i Lu)\Delta_i v = \sum_i (LS_{i-1}Q_i u)\Delta_i v.$$

Observe that, as for the standard paraproduct, $(LS_{i-1}Q_i F(u))\Delta_i v$ has a spatial Fourier transform localized in an annulus $2^i\mathscr{A}$, so that according to Lemma A.3 it will be sufficient to control its $C_T L^\infty$ norm. But

$$\|LS_{i-1}Q_i u\|_{C_T L^\infty} \leqslant \|\partial_t Q_i S_{i-1} u\|_{C_T L^\infty} + \|Q_i \Delta S_{i-1} u\|_{C_T L^\infty}$$
$$\lesssim 2^{-(\alpha-2)i}\big(\|S_{i-1}u\|_{C_T^{\alpha/2}L^\infty} + \|u\|_{C_T\mathscr{C}^\alpha}\big),$$

where we used the bounds (32). It is easy to see that $\|S_{i-1}u\|_{C_T^{\alpha/2}} \lesssim \|u\|_{C_T^{\alpha/2}}$, and therefore we obtain (33).

As for (34), we have

$$u \prec v - u \prec\!\!\prec v = \sum_i (Q_i S_{i-1}u - S_{i-1}u)\Delta_i v,$$

and again it will be sufficient to control the $C_T L^\infty$ norm of each term of the series. But using once more (32), we obtain

$$\|(Q_i S_{i-1}u - S_{i-1}u)\Delta_i v\|_{C_T L^\infty} \lesssim 2^{-i\alpha}\|S_{i-1}u\|_{C_T^{\alpha/2}L^\infty}\|\Delta_i v\|_{C_T L^\infty} \lesssim 2^{-i(\alpha+\beta)}\|u\|_{C_T^{\alpha/2}L^\infty}\|v\|_{C_T\mathscr{C}^\beta},$$

and the result is proved. □

Letting

$$u = F(u) \prec\!\!\prec \vartheta + u^\sharp \tag{35}$$

and redoing the same computation as above gives

$$Lu^\sharp = \Phi^\sharp = -[L(F(u) \prec\!\!\prec \vartheta) - F(u) \prec\!\!\prec \xi] + [F(u) \prec \xi - F(u) \prec\!\!\prec \xi] + F(u) \succ \xi + F(u) \circ \xi. \tag{36}$$

Lemma 5.1 takes care of the first two terms on the right hand side. The term $F(u) \succ \xi$ can be controlled using the paraproduct estimates, so that it remains to control the resonant product $F(u) \circ \xi$. In principle, this can be achieved by combining the decomposition described above with (34), which enables us to switch between the two paraproducts $\prec\!\!\prec$ and $\prec$. However, in that way we would pick up a superlinear estimate from Lemma 2.7, and thus would get a problem when trying to construct global in time solutions. We therefore have to be slightly more careful.

Lemma 5.2. *Let $\alpha \in (2/3, 1)$ and $\beta \in (0, \alpha]$ be such that $2\alpha + \beta > 2$. Let $T > 0$, $\xi \in C(\mathbb{T}^2, \mathbb{R})$, $\vartheta \in C_T\mathscr{C}^\alpha$, $u \in C_T\mathscr{C}^\alpha$, and let $F \in C_b^{1+\beta/\alpha}$. Define $u^\sharp = u - F(u) \prec\!\!\prec \vartheta$. Then*

$$\|(F(u) \circ \xi)(t)\|_{\alpha+\beta-2} \lesssim C_F C_\xi \big(1 + \|u\|_{C_T\mathscr{C}^\alpha} + \|u\|_{C_T^{\alpha/2}L^\infty} + \|u^\sharp(t)\|_{\alpha+\beta}\big),$$

for all $t \in [0, T]$, where

$$C_\xi = (1 + \|\xi\|_{\alpha-2})(1 + \|\vartheta\|_{C_T\mathscr{C}^\alpha}^{1+\beta/\alpha}) + \|\vartheta \circ \xi\|_{C_T\mathscr{C}^{2\alpha-2}} \quad \text{and} \quad C_F = \|F\|_{C_b^{1+\beta/\alpha}} + \|F\|_{C_b^{1+\beta/\alpha}}^{2+\beta/\alpha}. \tag{37}$$

We pay attention to indicate that, for fixed $t \in [0, T]$, the estimate depends only on the $\mathscr{C}^{\alpha+\beta}$ norm of $u^\sharp(t)$. This will come useful below when introducing the right norm to control the contribution of the initial condition.

Proof. We decompose

$$\begin{aligned} F(u) \circ \xi = {} & (F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)) \circ \xi + C(F'(u), F(u) \prec\!\!\prec \vartheta, \xi) \\ & + F'(u)[(F(u) \prec\!\!\prec \vartheta - F(u) \prec \vartheta) \circ \xi] + F'(u)C(F(u), \vartheta, \xi) + F'(u)F(u)(\vartheta \circ \xi), \end{aligned} \tag{38}$$

from where we can use Lemma 5.1 and the commutator estimate Lemma 2.4 to see that

$$\|F(u) \circ \xi - (F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)) \circ \xi\|_{C_T\mathscr{C}^{\alpha+\beta-2}} \lesssim C_F C_\xi \big(1 + \|u\|_{C_T\mathscr{C}^\alpha} + \|u\|_{C_T^{\alpha/2}L^\infty}\big).$$

It remains to treat the first term on the right hand side of (38), which we split into two parts:

$$(F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)) \circ \xi = \sum_{\substack{i,j \leqslant n, \\ |i-j| \leqslant 1}} \Delta_i[F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)]\Delta_j \xi + \sum_{\substack{i,j > n, \\ |i-j| \leqslant 1}} \Delta_i[F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)]\Delta_j \xi \tag{39}$$

for $n \in \mathbb{N}$. For the first series, we use that $\alpha + \beta - 2 < 0$ and simply estimate

$$\Big\| \sum_{\substack{i,j \leqslant n, \\ |i-j| \leqslant 1}} \Delta_i[F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)]\Delta_j \xi \Big\|_{C_T \mathscr{C}^{\alpha+\beta-2}} \lesssim \sum_{i \leqslant n} 2^{i(2-\alpha)} C_F C_\xi \lesssim 2^{n(2-\alpha)} C_F C_\xi.$$

For the second series in (39) we get

$$\begin{aligned} &\Big\| \Delta_k \Big(\sum_{\substack{i,j > n, \\ |i-j| \leqslant 1}} \Delta_i[F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)]\Delta_j \xi \Big)(t) \Big\|_{L^\infty} \\ &\quad \lesssim (\mathbf{1}_{k \leqslant n} 2^{-n(2\alpha+\beta-2)} + \mathbf{1}_{k > n} 2^{-k(2\alpha+\beta-2)}) \|(F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta))(t)\|_{\alpha+\beta} \|\xi\|_{\alpha-2} \\ &\quad \lesssim 2^{-n\alpha} 2^{-k(\alpha+\beta-2)} \|(F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta))(t)\|_{\alpha+\beta} \|\xi\|_{\alpha-2} \end{aligned}$$

for all $t \in [0, T]$ and $k \geqslant -1$. Now a slight modification of the proof of Lemma 2.6 shows that for $f \in \mathscr{C}^\alpha$ and $g \in \mathscr{C}^{\alpha+\beta}$ we have

$$\|F(f+g) - F'(f+g) \prec f\|_{\alpha+\beta} \lesssim \|F\|_{C_b^{1+\beta/\alpha}} (1 + \|f\|_\alpha^{1+\beta/\alpha})(1 + \|g\|_{\alpha+\beta}^{\alpha+\beta}),$$

and applying this with $f = (F(u) \prec\!\!\prec \vartheta)(t)$ and $g = u^\sharp(t)$ we deduce that

$$\big\| \big((F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)) \circ \xi \big)(t) \big\|_{\alpha+\beta-2} \lesssim C_F C_\xi (2^{n(2-\alpha)} + 2^{-n\alpha} \|u^\sharp(t)\|_{\alpha+\beta}^{\alpha+\beta})$$

for all $n \in \mathbb{N}$. It remains to optimize over the parameter n. If $\|u^\sharp(t)\|_{\alpha+\beta} \leqslant 1$, we choose $n = 0$ and obtain

$$\big\| \big((F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)) \circ \xi \big)(t) \big\|_{\alpha+\beta-2} \lesssim C_F C_\xi (1 + \|u^\sharp(t)\|_{\alpha+\beta}).$$

If $\|u^\sharp(t)\|_{\alpha+\beta} > 1$, we choose n with $2^{n(2-\alpha)} \simeq \|u^\sharp(t)\|_{\alpha+\beta}$, so that

$$2^{n(2-\alpha)} + 2^{-n\alpha} \|u^\sharp(t)\|_{\alpha+\beta}^{\alpha+\beta} \simeq \|u^\sharp(t)\|_{\alpha+\beta} + \|u^\sharp(t)\|_{\alpha+\beta}^{-\frac{\alpha}{2-\alpha}+\alpha+\beta}.$$

Since $\beta \leqslant \alpha$, we can bound the exponent on the second term from above by $2\alpha - \alpha/(2-\alpha)$, and it is not hard to see that for $\alpha < 2$ this expression is smaller or equal to 1. Therefore, we obtain also in that case

$$\big\| \big((F(u) - F'(u) \prec (F(u) \prec\!\!\prec \vartheta)) \circ \xi \big)(t) \big\|_{\alpha+\beta-2} \lesssim C_F C_\xi (1 + \|u^\sharp(t)\|_{\alpha+\beta}),$$

which completes the proof. □

Let us summarize our observations so far.

Lemma 5.3. *Let $\alpha \in (2/3, 1)$, $\beta \in (2-2\alpha, \alpha]$, and $T > 0$. Let $u_0 \in \mathscr{C}^\alpha$, $\xi \in C(\mathbb{T}^2, \mathbb{R})$, let ϑ be such that $L\vartheta = \xi$, and let $F \in C_b^{1+\beta/\alpha}$. Then u solves the PDE*

$$Lu = F(u)\xi, \qquad u(0) = u_0 \in \mathscr{C}^\alpha$$

on $[0, T]$ if and only if $u = F(u) \prec\!\!\prec \vartheta + u^\sharp$, where $u^\sharp$ solves

$$Lu^\sharp = \Phi^\sharp, \qquad u^\sharp(0) = u_0 - (F(u) \prec\!\!\prec \vartheta)(0)$$

on $[0, T]$, for $\Phi^\sharp$ as defined in (36). Moreover, for all $t \in [0, T]$ we have the estimate

$$\|\Phi^\sharp(t)\|_{\alpha+\beta-2} \lesssim C_F C_\xi \big(1 + \|u\|_{C_T\mathscr{C}^\alpha} + \|u\|_{C_T^{\alpha/2}L^\infty} + \|u^\sharp(t)\|_{\alpha+\beta}\big), \tag{40}$$

where C_F and C_ξ are as defined in (37).

Next, we would like to close the estimate (40), so that the right hand side only depends on $\Phi^\sharp$. In order to estimate $\|u\|_{C_T\mathscr{C}^\alpha} + \|u\|_{C_T^{\alpha/2}L^\infty}$, we observe that $u = u^\sharp + F(u) \prec\!\!\prec \vartheta$ and that

$$LF(u) \prec\!\!\prec \vartheta = [L(F(u) \prec\!\!\prec \vartheta) - F(u) \prec\!\!\prec (L\vartheta)] + F(u) \prec\!\!\prec \xi.$$

Now it is easy to see that

$$\|L(F(u) \prec\!\!\prec \vartheta) - F(u) \prec\!\!\prec (L\vartheta)\|_{C_T\mathscr{C}^{\alpha-2}} \lesssim \|F(u)\|_{C_T L^\infty} \|\vartheta\|_{C_T\mathscr{C}^\alpha} \lesssim \|F\|_{L^\infty} \|\vartheta\|_{C_T\mathscr{C}^\alpha}$$

(compare also the proof of Lemma 5.1). Thus, we can apply the heat flow estimates Lemma A.7, Lemma A.8, and Lemma A.9, to deduce

$$\begin{aligned} \|u\|_{C_T\mathscr{C}^\alpha} + \|u\|_{C_T^{\alpha/2}L^\infty} &\lesssim \|u^\sharp\|_{C_T\mathscr{C}^\alpha} + \|u^\sharp\|_{C_T^{\alpha/2}L^\infty} + \|F(u) \prec\!\!\prec \vartheta(0)\|_\alpha + \|L(F(u) \prec\!\!\prec \vartheta)\|_{C_T\mathscr{C}^{\alpha-2}} \\ &\lesssim \|u^\sharp\|_{C_T\mathscr{C}^\alpha} + \|u^\sharp\|_{C_T^{\alpha/2}L^\infty} + \|u_0\|_\alpha + \|F\|_{L^\infty}(\|\vartheta\|_{C_T\mathscr{C}^\alpha} + \|\xi\|_{\alpha-2}). \end{aligned}$$

Plugging this into (40), we get

$$\|\Phi^\sharp(t)\|_{\alpha+\beta-2} \lesssim C_F C_\xi \big(1 + C_F C_\xi + \|u_0\|_\alpha + \|u^\sharp\|_{C_T\mathscr{C}^\alpha} + \|u^\sharp\|_{C_T^{\alpha/2}L^\infty} + \|u^\sharp(t)\|_{\alpha+\beta}\big).$$

Moreover, since $u^\sharp(0) = u_0 - (F(u) \prec\!\!\prec \vartheta)(0)$ and $Lu^\sharp = \Phi^\sharp$, we can apply Lemma A.7 and Lemma A.9 to obtain

$$t^{\beta/2}\|u^\sharp(t)\|_{\alpha+\beta} \lesssim \|u_0\|_\alpha + C_F C_\xi + \sup_{s\in[0,t]} (s^{\beta/2}\|\Phi^\sharp(s)\|_{\alpha+\beta-2}),$$

so that our new estimate for $\Phi^\sharp$ reads

$$t^{\beta/2}\|\Phi^\sharp(t)\|_{\alpha+\beta-2} \lesssim C_F C_\xi \big(1 + C_F C_\xi + \|u_0\|_\alpha + \|u^\sharp\|_{C_T\mathscr{C}^\alpha} + \|u^\sharp\|_{C_T^{\alpha/2}L^\infty} + \sup_{s\in[0,t]} (s^{\beta/2}\|\Phi^\sharp(s)\|_{\alpha+\beta-2})\big),$$

uniformly in $t \in [0, T]$. It remains to control $u^\sharp$ in $C_T^{\alpha/2}L^\infty \cap C_T\mathscr{C}^\alpha$. For $0 \leqslant s < t \leqslant T$, we have

$$\begin{aligned} \|u^\sharp(t) - u^\sharp(s)\|_{L^\infty} \leqslant \|(P_{t-s} - \mathrm{id})P_s(u^\sharp(0))\|_{L^\infty} &+ \left\| \int_s^t P_{t-s}\Phi^\sharp(r)\mathrm{d}r \right\|_{L^\infty} \\ &+ \left\| \int_0^s (P_{t-s} - \mathrm{id})P_{s-r}\Phi^\sharp(r)\mathrm{d}r \right\|_{L^\infty}. \end{aligned}$$

Applying Lemma A.8 to the first and third term and Lemma A.7 to the second term, we obtain

$$
\begin{aligned}
\|u^\sharp(t) - u^\sharp(s)\|_{L^\infty} &\lesssim (t-s)^{\alpha/2}\|u^\sharp(0)\|_\alpha + \int_s^t (t-s)^{-1+\alpha/2+\beta/2}\|\Phi^\sharp(r)\|_{\alpha+\beta-2}\mathrm{d}r \\
&\quad + (t-s)^{\alpha/2}\int_0^s \|P_{s-r}\Phi^\sharp(r)\|_\alpha \mathrm{d}r \\
&\lesssim (t-s)^{\alpha/2}(C_F C_\xi + \|u_0\|_\alpha) \\
&\quad + (t-s)^{\alpha/2}\int_0^t (t-r)^{-1+\beta/2} r^{-\beta/2}\mathrm{d}r \sup_{r\in[0,t]}(r^{\beta/2}\|\Phi^\sharp(r)\|_{\alpha+\beta-2}) \\
&\quad + (t-s)^{\alpha/2}\int_0^s (s-r)^{-1+\beta/2} r^{-\beta/2}\mathrm{d}r \sup_{r\in[0,s]}(r^{\beta/2}\|\Phi^\sharp(r)\|_{\alpha+\beta-2}).
\end{aligned}
$$

For the time integrals we have $\int_0^t (t-r)^{-1+\beta/2} r^{-\beta/2}\mathrm{d}r = \int_0^1 (1-r)^{1-\beta/2} r^{-\beta/2}\mathrm{d}r \lesssim 1$, so that

$$
\|u^\sharp\|_{C_T^{\alpha/2}L^\infty} \lesssim C_F C_\xi + \|u_0\|_\alpha + \sup_{s\in[0,T]}(s^{\beta/2}\|\Phi^\sharp(s)\|_{\alpha+\beta-2}).
$$

Similar (but easier) arguments can be used to bound the $C_T\mathscr{C}^\alpha$ norm of $u^\sharp$, and thus we obtain our final estimate for $\Phi^\sharp$:

$$
\sup_{t\in[0,T]}(t^{\beta/2}\|\Phi^\sharp(t)\|_{\alpha+\beta-2}) \lesssim C_F C_\xi (1 + C_F C_\xi)\Big(1 + \|u_0\|_\alpha + \sup_{t\in[0,T]}(t^{\beta/2}\|\Phi^\sharp(t)\|_{\alpha+\beta-2})\Big). \tag{41}
$$

In order to use this estimate to bound $\Phi^\sharp$, we will apply the usual scaling argument. More precisely, we set $\Lambda_\lambda f(t,x) = f(\lambda^2 t, \lambda x)$, so that $L\Lambda_\lambda = \lambda^2\Lambda_\lambda L$. Now let $u^\lambda = \Lambda_\lambda u$, $u_0^\lambda = \Lambda_\lambda u_0$, $\xi^\lambda = \lambda^{2-\alpha}\Lambda_\lambda \xi$, and $\vartheta^\lambda = \lambda^\alpha \Lambda_\lambda \vartheta$. Note that $u^\lambda\colon [0,T/\lambda^2]\times\mathbb{T}^2_\lambda \to \mathbb{R}$, where $\mathbb{T}^2_\lambda = (\mathbb{R}/(2\pi\lambda^{-1}\mathbb{Z}))^2$ is a rescaled torus, and that u^λ solves the equation

$$
Lu^\lambda = \lambda^2 F(u^\lambda)\Lambda_\lambda \xi = \lambda^\alpha F(u^\lambda)\xi^\lambda, \qquad u^\lambda(0) = u_0^\lambda.
$$

The scaling is chosen in such a way that $\|u_0^\lambda\|_\alpha \lesssim \|u_0\|_\alpha$, $\|\xi^\lambda\|_{\mathscr{C}^{\alpha-2}} \lesssim \|\xi\|_{\mathscr{C}^{\alpha-2}}$, $\|\vartheta^\lambda\|_{C_T\mathscr{C}^\alpha} \lesssim \|\vartheta\|_{C_T\mathscr{C}^\alpha}$, and according to Lemma B.1 also $\|\vartheta^\lambda \circ \xi^\lambda\|_{C_T\mathscr{C}^{2\alpha-2}} \lesssim \|\vartheta\circ\xi\|_{C_T\mathscr{C}^{2\alpha-2}} + \|\vartheta\|_{C_T\mathscr{C}^\alpha}\|\xi\|_{\alpha-2}$, all uniformly in $\lambda\in(0,1]$. In particular, $C_{\xi^\lambda} \lesssim C_\xi$ and $C_{\lambda^\alpha F} \leqslant \lambda^\alpha C_F$ for all $\lambda\in(0,1]$. Injecting these estimates into (41), we obtain

$$
\sup_{t\in[0,T]}(t^{\beta/2}\|\Phi^{\sharp,\lambda}(t)\|_{\alpha+\beta-2}) \lesssim 1 + \|u_0^\lambda\|_\alpha
$$

for all sufficiently small $\lambda > 0$ (depending only on C_ξ and C_F, but not on u_0), where $\Phi^{\sharp,\lambda}$ is defined analogously to $\Phi^\sharp$. From here we easily get the existence of paracontrolled solutions to (31). Similar arguments show that if $F \in C_b^{2+\beta/\alpha}$, then the map $(u_0, \xi, \vartheta, \xi\circ\vartheta) \mapsto u \in C_T\mathscr{C}^\alpha$ is locally Lipschitz continuous, and in particular there is a unique paracontrolled solution.

5.1 Renormalization

So far, we argued under the assumption that there exist continuous functions (ξ^ε) such that $(\xi^\varepsilon, \vartheta^\varepsilon, \vartheta^\varepsilon\circ\xi^\varepsilon)$ converges to $(\xi, \vartheta, \vartheta\circ\xi)$ in $\mathscr{C}^{\alpha-2}\times C_T\mathscr{C}^{2\alpha-2}\times C_T\mathscr{C}^{2\alpha-2}$ as $\varepsilon\to 0$. Here the superscript ε refers to a smooth regularization of the noise, whereas in the previous section the superscript λ referred to a scaling transform. From now on we will no longer consider scaling transforms, so that no confusion should arise.

One further difficulty is that the resonant product $(\vartheta^\varepsilon\circ\xi^\varepsilon)$ does not converge in some relevant cases; in particular, if ξ is a spatial white noise. However, what we will show below is that for the white noise there exist constants $c_\varepsilon \in \mathbb{R}$ such that $((\vartheta^\varepsilon \circ \xi^\varepsilon) - c_\varepsilon)$ converges in probability in $C_T\mathscr{C}^{2\alpha-2}$. In order to make the term c_ε appear in the equation, we can introduce a suitable correction term in the regularized problems and consider the renormalized PDE

$$Lu^\varepsilon = F(u^\varepsilon)\xi^\varepsilon - c_\varepsilon F'(u^\varepsilon)F(u^\varepsilon). \tag{42}$$

For this equation we use again the paracontrolled ansatz (35). The same derivation as for (36) yields

$$Lu^{\sharp,\varepsilon} = G(u^\varepsilon, \vartheta^\varepsilon, \xi^\varepsilon) + F(u^\varepsilon) \circ \xi^\varepsilon - c_\varepsilon F'(u^\varepsilon)F(u^\varepsilon)$$

for some bounded functional G, and as in Lemma 5.2 we decompose

$$F(u^\varepsilon) \circ \xi^\varepsilon - c_\varepsilon F'(u^\varepsilon)F(u^\varepsilon) = H(u^\varepsilon, u^{\sharp,\varepsilon}, \vartheta^\varepsilon, \xi^\varepsilon) + F'(u^\varepsilon)F(u^\varepsilon)(\vartheta^\varepsilon \circ \xi^\varepsilon - c_\varepsilon)$$

for another bounded functional H. We see that $Lu^{\sharp,\varepsilon}$ only depends on ξ^ε, ϑ^ε, and $(\vartheta^\varepsilon \circ \xi^\varepsilon) - c_\varepsilon$. Thus, the convergence of $(\xi^\varepsilon, \vartheta^\varepsilon, \vartheta^\varepsilon \circ \xi^\varepsilon - c_\varepsilon)$ to (ξ, ϑ, η) in $\mathscr{C}^{\alpha-2} \times C_T\mathscr{C}^{2\alpha-2} \times C_T\mathscr{C}^{2\alpha-2}$ implies that the solutions (u^ε) to (42) converge to a limit which only depends on ξ, ϑ, and η, but not on the approximating family.

Theorem 5.4. *Let $\alpha \in (2/3, 1)$, $\beta \in (2 - 2\alpha, \alpha]$ and assume that $(\xi^\varepsilon)_{\varepsilon>0} \subset C(\mathbb{T}^2, \mathbb{R})$ and $F \in C_b^{2+\beta/\alpha}$. Suppose that there exist $\xi \in \mathscr{C}^{\alpha-2}$, $\vartheta \in C_T\mathscr{C}^\alpha$, and $\eta \in C_T\mathscr{C}^{2\alpha-2}$ such that $(\xi^\varepsilon, \vartheta^\varepsilon, (\vartheta^\varepsilon \circ \xi^\varepsilon) - c_\varepsilon)$ converges to (ξ, ϑ, η) in $\mathscr{C}^{\alpha-2} \times C_T\mathscr{C}^\alpha \times C_T\mathscr{C}^{2\alpha-2}$, where ϑ^ε are solutions to $L\vartheta^\varepsilon = \xi^\varepsilon$, and where $c_\varepsilon \in \mathbb{R}$ for all $\varepsilon > 0$. Let for $\varepsilon > 0$ the function u^ε be the unique solution to the Cauchy problem*

$$Lu^\varepsilon = F(u^\varepsilon)\xi^\varepsilon - c_\varepsilon F'(u^\varepsilon)F(u^\varepsilon), \qquad u^\varepsilon(0) = u_0,$$

where $u_0 \in \mathscr{C}^\alpha$. Then there exists $u \in C_T\mathscr{C}^\alpha$ such that $u^\varepsilon \to u$ in $C_T\mathscr{C}^\alpha$. The limit u depends only on $(u_0, \xi, \vartheta, \eta)$, and not on the approximating family $(\xi^\varepsilon, \vartheta^\varepsilon, (\vartheta^\varepsilon \circ \xi^\varepsilon) - c_\varepsilon)$.

As for the previous equations, u is the unique paracontrolled weak solution to $Lu = F(u) \diamond \xi$ with $u(0) = u_0$ if we interpret the renormalized product $F(u) \diamond \xi$ in the right way, the initial condition $\vartheta(0)$ is of no importance, and u depends continuously on u_0.

5.2 Regularity of the area and renormalized products

It remains to study the regularity of the area $\vartheta \circ \xi$. As already indicated, we will have to renormalize the product by "subtracting an infinite constant" in order to obtain a well-defined object.

To simplify the arguments below, we assume that ξ is given by $\tilde{\xi} - (2\pi)^{-2}\mathscr{F}\tilde{\xi}(0)$, where $\tilde{\xi}$ is a spatial white noise on $\mathbb{T}^2$. Then $(\mathscr{F}\xi(k))_{k\in\mathbb{Z}^2}$ is a complex valued, centered Gaussian process with covariance

$$\mathbb{E}[\mathscr{F}\xi(k)\overline{\mathscr{F}\xi(k')}] = (2\pi)^2 \mathbf{1}_{k=k'}\mathbf{1}_{k\neq 0}.$$

Since ξ is a smooth additive perturbation of $\tilde{\xi}$, this simplification will pose no problems, and we indicate below how to handle $\tilde{\xi}$ once we are able to handle ξ.

Since ξ is a mean zero distribution, there exists a stationary solution ϑ to $L\vartheta = \xi$, given by

$$\vartheta(x) = \int_0^\infty (P_t\xi)(x)\mathrm{d}t.$$

Then $L\vartheta = \xi$ by definition, and it is easily verified that $(\mathscr{F}\vartheta(k))$ is a centered, complex valued Gaussian process with covariance

$$\mathbb{E}[\mathscr{F}\vartheta(k)\mathscr{F}\vartheta(k')] = (2\pi)^2 \frac{1}{|k|^4}\mathbf{1}_{k=-k'}\mathbf{1}_{k\neq 0}$$

and such that $\overline{\mathscr{F}\vartheta(k)} = \mathscr{F}\vartheta(-k)$. This yields, using Gaussian hypercontractivity and Besov embedding, that $\mathbb{E}[\|\vartheta\|^p_{\mathscr{C}^\alpha(\mathbb{T}^2)}] < \infty$ for all $\alpha < 1$ and $p \geqslant 1$. Since $P_t\xi$ is a smooth function for $t > 0$, the resonant term $P_t\xi \circ \xi$ is a smooth function, and therefore we could formally set $\vartheta \circ \xi = \int_0^\infty (P_t\xi \circ \xi)\mathrm{d}t$. However, this expression is not well defined:

Lemma 5.5. *For any $x \in \mathbb{T}^2$ and $t > 0$ we have*

$$g_t = \mathbb{E}[(P_t\xi \circ \xi)(x)] = \mathbb{E}[\Delta_{-1}(P_t\xi \circ \xi)(x)] = (2\pi)^{-2} \sum_{k\in\mathbb{Z}^2\setminus\{0\}} e^{-t|k|^2}.$$

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 $\ell \geqslant -1$. 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 $p \geqslant 1$. Now $P_t = e^{-t|\mathrm{D}|^2}$, and therefore

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

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

$$g_t = \mathbb{E}[(P_t\xi \circ \xi)(x)] = \sum_{k\in\mathbb{Z}^2\setminus\{0\}} \sum_{i,j} \rho_i(k)\rho_j(k)e^{-t|k|^2} = (2\pi)^{-2} \sum_{k\in\mathbb{Z}^2\setminus\{0\}} e^{-t|k|^2},$$

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

Remark 5.6. The same calculation shows that if $\psi \in \mathscr{S}$, and if $\xi^\varepsilon = \varepsilon^{-2}\psi(\varepsilon^{-1}\cdot) * \xi$, then

$$\mathbb{E}[(P_t\xi^\varepsilon \circ \xi^\varepsilon)(x)] = \mathbb{E}[\Delta_{-1}(P_t\xi^\varepsilon \circ \xi^\varepsilon)(x)] = (2\pi)^{-2} \sum_{k\in\mathbb{Z}^2\setminus\{0\}} e^{-t|k|^2}|\mathscr{F}\psi(\varepsilon k)|^2.$$

The diverging time integral motivates us to study the renormalized product $\vartheta \circ \xi - \int_0^\infty g_t\mathrm{d}t$, where $\int_0^\infty g_t\mathrm{d}t$ is an infinite constant:

Lemma 5.7. *Set*

$$(\vartheta \diamond \xi) = \int_0^\infty (P_t\xi \circ \xi - g_t)\mathrm{d}t.$$

Then $\mathbb{E}[\|\vartheta \diamond \xi\|^p_{2\alpha-2}] < \infty$ for all $\alpha < 1$, $p \geqslant 1$. Moreover, if $\psi \in \mathscr{S}$ satisfies $\int \psi(x)\mathrm{d}x = 1$, and if $\xi^\varepsilon = \varepsilon^{-2}\psi(\varepsilon\cdot) * \xi$ with $\psi^\varepsilon = \varepsilon^{-2}\psi(\varepsilon\cdot)$ for $\varepsilon > 0$, and $\vartheta^\varepsilon = \int_0^\infty P_t\xi^\varepsilon\mathrm{d}t$, then

$$\lim_{\varepsilon\to 0} \mathbb{E}[\|\vartheta \diamond \xi - (\vartheta^\varepsilon \circ \xi^\varepsilon - c_\varepsilon)\|^p_{2\alpha-2}] = 0$$

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

$$c_\varepsilon = \mathbb{E}[\vartheta^\varepsilon(x)\xi^\varepsilon(x)] = \mathbb{E}[\vartheta^\varepsilon \circ \xi^\varepsilon(x)] = \int_0^\infty \mathbb{E}[P_t\xi^\varepsilon \circ \xi^\varepsilon(x)]\mathrm{d}t = (2\pi)^{-2} \sum_{k\in\mathbb{Z}^2\setminus\{0\}} \frac{|\mathscr{F}\psi(\varepsilon k)|^2}{|k|^2}.$$

Proof. We split the time integral into two components, $\int_0^1 \dots \mathrm{d}t$ and $\int_1^\infty \dots \mathrm{d}t$. The second integral can be treated without relying on probabilistic estimates: Given $x \in \mathbb{T}^2$, we have

$$\begin{aligned}
&\Big\| \int_1^\infty (P_t\xi \circ \xi - g_t)\mathrm{d}t - \int_1^\infty (P_t\xi^\varepsilon \circ \xi^\varepsilon - \mathbb{E}[P_t\xi^\varepsilon \circ \xi^\varepsilon(x)])\mathrm{d}t \Big\|_{2\alpha-2} \\
&\lesssim \int_1^\infty \|P_t\xi \circ \xi - P_t\xi^\varepsilon \circ \xi^\varepsilon\|_{2\alpha}\mathrm{d}t + \int_1^\infty \sum_{k\in\mathbb{Z}^2\setminus\{0\}} e^{-t|k|^2}|1 - |\mathscr{F}\psi(\varepsilon k)|^2|\mathrm{d}t \\
&\lesssim \int_1^\infty (\|P_t(\xi - \xi^\varepsilon)\|_{\alpha+2}\|\xi\|_{\alpha-2} + \|P_t\xi^\varepsilon\|_{\alpha+2}\|\xi - \xi^\varepsilon\|_{\alpha-2})\mathrm{d}t + \sum_{k\in\mathbb{Z}^2\setminus\{0\}} \frac{e^{-|k|^2}}{|k|^2}|1 - |\mathscr{F}\psi(\varepsilon k)|^2|,
\end{aligned}$$

Since $\mathscr{F}\xi(0) = 0$, the estimate $\|P_t\xi^\varepsilon\|_{\alpha+2} \lesssim t^{-2}\|\xi^\varepsilon\|_{\alpha-2}$ of Lemma A.7 holds uniformly over $t > 0$, and thus the time integral is finite. The convergence in $L^p(\mathbb{P})$ now easily follows from the dominated convergence theorem.

We will treat the integral from 0 to 1 using similar arguments as in the proof of Lemma 4.6. To lighten the notation, we will only show that $\mathbb{E}[\| \int_0^1 (P_t\xi \circ \xi - g_t)\mathrm{d}t\|_{2\alpha-2}^p] < \infty$. The difference

$$\mathbb{E}\Big[\Big\| \int_0^1 (P_t\xi \circ \xi - g_t)\mathrm{d}t - \int_0^1 (P_t\xi^\varepsilon \circ \xi^\varepsilon - \mathbb{E}[P_t\xi^\varepsilon \circ \xi^\varepsilon(x)])\mathrm{d}t \Big\|_{2\alpha-2}^p\Big]$$

can be treated with the same arguments, we only have to include some additional factors of the form $|1 - \mathscr{F}\psi(\varepsilon k)|^2$ in the sums below. The convergence of the expectation can then be shown using dominated convergence.

Let $t \in (0, 1]$ and define $\Xi_t = P_t\xi \circ \xi - g_t$. By Gaussian hypercontractivity we obtain for $p \geqslant 1$ and $m \geqslant -1$ that

$$\mathbb{E}[\|\Delta_m\Xi_t\|_{L^{2p}(\mathbb{T}^2)}^{2p}] \lesssim_p \|\mathbb{E}[|\Delta_m\Xi_t(x)|^2]\|_{L_x^p(\mathbb{T}^2)}^p. \tag{43}$$

By Lemma 5.5 we have

$$\mathbb{E}[|\Delta_m\Xi_t(x)|^2] = \mathrm{Var}(\Delta_m(P_t\xi \circ \xi)(x)), \tag{44}$$

for all $m \geq -1$, where $\mathrm{Var}(\cdot)$ denotes the variance. Now

$$\Delta_m(P_t\xi \circ \xi)(x) = (2\pi)^{-4} \sum_{k_1,k_2\in\mathbb{Z}^2} \sum_{|i-j|\leqslant 1} e^{\imath\langle k_1+k_2,x\rangle}\rho_m(k_1 + k_2)\rho_i(k_1)e^{-t|k_1|^2}\mathscr{F}\xi(k_1)\rho_j(k_2)\mathscr{F}\xi(k_2),$$

and therefore

$$\begin{aligned}
&\mathrm{Var}(\Delta_m(P_t\xi \circ \xi)(x)) \\
&\quad= (2\pi)^{-8} \sum_{k_1,k_2\in\mathbb{Z}^2} \sum_{k_1',k_2'\in\mathbb{Z}^2} \sum_{|i-j|\leqslant 1} \sum_{|i'-j'|\leqslant 1} e^{\imath\langle k_1+k_2,x\rangle}\rho_m(k_1 + k_2)\rho_i(k_1)e^{-t|k_1|^2}\rho_j(k_2) \\
&\qquad\times e^{\imath\langle k_1'+k_2',x\rangle}\rho_m(k_1' + k_2')\rho_i(k_1')e^{-t|k_1'|^2}\rho_j(k_2') \operatorname{cov}(\mathscr{F}\xi(k_1)\mathscr{F}\xi(k_2), \mathscr{F}\xi(k_1')\mathscr{F}\xi(k_2')),
\end{aligned}$$

where exchanging summation and expectation can again be justified a posteriori by the uniform L^p–boundedness of the partial sums, and where cov denotes the covariance. Since $(\widehat{\xi}(k))_{k\in\mathbb{Z}^2}$ is a centered Gaussian process, we can apply Wick's theorem ([Jan97], Theorem 1.28) to deduce

$$\operatorname{cov}(\widehat{\xi}(k_1)\widehat{\xi}(k_2), \widehat{\xi}(k_1')\widehat{\xi}(k_2')) = (2\pi)^4(\mathbf{1}_{k_1=-k_1'}\mathbf{1}_{k_2=-k_2'} + \mathbf{1}_{k_1=-k_2'}\mathbf{1}_{k_2=-k_1'})\mathbf{1}_{k_1\neq 0}\mathbf{1}_{k_2\neq 0},$$

and therefore

$$\begin{aligned}
&\mathrm{Var}(\Delta_m(P_t\xi \circ \xi)(x)) \\
&\quad= (2\pi)^{-4} \sum_{k_1,k_2\in\mathbb{Z}^2\setminus\{0\}} \sum_{|i-j|\leqslant 1} \sum_{|i'-j'|\leqslant 1} \Big[\mathbf{1}_{m\lesssim i}\mathbf{1}_{m\lesssim i'}\rho_m^2(k_1 + k_2)\rho_i(k_1)\rho_j(k_2)\rho_{i'}(k_1)\rho_{j'}(k_2)e^{-2t|k_1|^2} \\
&\qquad\qquad + \mathbf{1}_{m\lesssim i}\mathbf{1}_{m\lesssim i'}\rho_m^2(k_1 + k_2)\rho_i(k_1)\rho_j(k_2)\rho_{i'}(k_2)\rho_{j'}(k_1)e^{-t|k_1|^2-t|k_2|^2}\Big].
\end{aligned}$$

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 $i \geqslant -1$. 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

$$
\begin{aligned}
&\mathrm{Var}(\Delta_m(P_t\xi \circ \xi))(x)) \\
&\qquad \lesssim \sum_{i,j,i',j'} \mathbf{1}_{m \lesssim i} \mathbf{1}_{i\sim j \sim i' \sim j'} \sum_{k_1,k_2 \in \mathbb{Z}^2 \setminus \{0\}} \mathbf{1}_{\mathrm{supp}(\rho_m)}(k_1+k_2) \mathbf{1}_{\mathrm{supp}(\rho_i)}(k_1) \mathbf{1}_{\mathrm{supp}(\rho_j)}(k_2) e^{-2tc2^{2i}} \\
&\qquad \lesssim \sum_{i: i \gtrsim m} 2^{2i} 2^{2m} e^{-tc2^{2i}} \lesssim \frac{2^{2m}}{t} \sum_{i:i\gtrsim m} e^{-tc2^{2i}} \lesssim \frac{2^{2m}}{t} e^{-tc2^{2m}},
\end{aligned}
\tag{45}
$$

where we used that $t2^{2i} \lesssim e^{t(c-c')2^{2i}}$ for any $c' < c$.

Now let $\alpha < 1$. We apply Jensen's inequality and combine (43), (44), and (45) to obtain

$$
\mathbb{E}[\|\Xi_t\|_{B^{2\alpha-2}_{2p,2p}}] \lesssim \Big(\sum_{m \geqslant -1} 2^{(2\alpha-2)m2p} \mathbb{E}[\|\Delta_m \Xi_t\|^{2p}_{L^{2p}(\mathbb{T}^2)}]\Big)^{\frac{1}{2p}}
$$

$$
\lesssim t^{-1/2} \Big(\sum_{m \geqslant -1} 2^{(2\alpha-2)m2p} 2^{2mp} e^{-tcp2^{2m}} \Big)^{\frac{1}{2p}} \lesssim t^{-1/2} \Big(\int_{-1}^{\infty} (2^x)^{2p(2\alpha-1)} e^{-ctp(2^x)^2} \mathrm{d}x \Big)^{\frac{1}{2p}}.
$$

The change of variables $y = \sqrt{t}2^x$ then yields

$$
\mathbb{E}[\|\Xi_t\|_{B^{2\alpha-2}_{2p,2p}}] \lesssim t^{-1/2} \Big(t^{-p(2\alpha-1)} \int_0^\infty y^{2p(2\alpha-1)-1} e^{-cpy^2} \mathrm{d}y \Big)^{\frac{1}{2p}}.
$$

If $\alpha > 1/2$, the integral is finite for all sufficiently large p, and therefore $\mathbb{E}[\|\Xi_t\|_{B^{2\alpha-2}_{2p,2p}}] \lesssim_p t^{-\alpha}$, so that $\int_0^1 \mathbb{E}[\|\Xi_t\|_{B^{2\alpha-2}_{2p,2p}}]\mathrm{d}t < \infty$ for all $\alpha < 1$. Gaussian hypercontractivity allows us to conclude that also

$$
\mathbb{E}\Big[\Big\| \int_0^1 \Xi_t \mathrm{d}t \Big\|^p_{B^{2\alpha-2}_{2p,2p}}\Big] < \infty
$$

for all $p \geqslant 1$. The result now follows from the Besov embedding theorem, Lemma A.2. □

Corollary 5.8. *Let $\tilde\xi$ be a spatial white noise, and let ξ, ϑ, and $\vartheta \diamond \xi$ be as defined above. Set $\tilde\vartheta(t,x) = \vartheta(x) + t(2\pi)^{-2}\mathscr{F}\tilde\xi(0)$ (so that $L\tilde\vartheta = \tilde\xi$) and*

$$
(\tilde\vartheta \diamond \tilde\xi)(t,x) = (\vartheta \diamond \xi)(t,x) + (2\pi)^{-2}(\vartheta \circ \mathscr{F}\tilde\xi(0))(x) + t(2\pi)^{-2}(\mathscr{F}\tilde\xi(0) \circ \xi)(x) + t(2\pi)^{-2}(\mathscr{F}\tilde\xi(0))^2.
$$

If $\psi \in \mathscr{S}$ satisfies $\int \psi(x)\mathrm{d}x = 1$, and if $\tilde\xi^\varepsilon = \psi^\varepsilon * \tilde\xi$ and $\tilde\vartheta^\varepsilon = \psi^\varepsilon * \tilde\vartheta$, where $\psi^\varepsilon = \varepsilon^{-2}\psi(\varepsilon\cdot)$ for $\varepsilon > 0$, then $\lim_{\varepsilon\to 0} \mathbb{E}[\|\tilde\vartheta \diamond \tilde\xi - (\tilde\vartheta^\varepsilon \circ \tilde\xi^\varepsilon - c_\varepsilon)\|^p_{C_T\mathscr{C}^{2\alpha-2}}] = 0$ for all $T > 0$ and $p \geqslant 1$, where c_ε is as defined in Lemma 5.7.

Combining the existence of the renormalized product $\vartheta \diamond \xi$ with Theorem 5.4, we obtain the existence and uniqueness of solutions to the generalized parabolic Anderson model:

Corollary 5.9. *Let $\alpha \in (2/3, 1)$, $\beta \in (2-2\alpha, \alpha]$, $F \in C_b^{2+\beta/\alpha}$, $u_0 \in \mathscr{C}^\alpha$, $L = \partial_t - \Delta$, and let ξ be a spatial white noise on $\mathbb{T}^2$. Then there exists a unique solution u to*

$$
Lu = F(u) \diamond \xi, \qquad u(0) = u_0,
$$

in the following sense: For $\psi \in \mathscr{S}$ with $\int \psi \mathrm{d}t = 1$ and for $\varepsilon > 0$ consider the solution u^ε to

$$
Lu^\varepsilon = F(u^\varepsilon)\xi^\varepsilon - c_\varepsilon F'(u^\varepsilon)F(u^\varepsilon), \qquad u^\varepsilon(0) = u_0,
$$

on $[0,\infty) \times \mathbb{T}^2$, where $\xi^\varepsilon = \varepsilon^{-1}\psi(\varepsilon\cdot) * \xi$, and where c_ε is as defined in Lemma 5.7. Then for all $T > 0$, (u^ε) converges in probability in $C_T\mathscr{C}^\alpha$ to u.

Remark 5.10. Concerning the convergence of $(\vartheta^\varepsilon \circ \xi^\varepsilon)$, let us make the following remark: Since $L\vartheta^\varepsilon = \xi^\varepsilon$, we have

$$(\vartheta^\varepsilon \circ \xi^\varepsilon) = (\vartheta^\varepsilon \circ L\vartheta^\varepsilon) = \frac{1}{2}L(\vartheta^\varepsilon \circ \vartheta^\varepsilon) - (\mathrm{D}_x\vartheta^\varepsilon \circ \mathrm{D}_x\vartheta^\varepsilon),$$

from which we see that the only problem in passing to the limit is given by the second term on the right hand side. This integration by parts formula is the crucial difference with what happens in the RDE case, which otherwise shares many structural properties with the PAM model. The fact that L is a second order operator generates the term $(\mathrm{D}_x\vartheta^\varepsilon \circ \mathrm{D}_x\vartheta^\varepsilon)$ in the above computation, which is absent in case of the operator ∂_t. This term, whose convergence is equivalent to the convergence of the positive term $|\mathrm{D}_x\vartheta^\varepsilon|^2$, cannot have simple cancellation properties and it is the origin for the need of introducing an additive renomalization when considering the PAM model.

Our previous analysis easily implies that the solutions to the modified problem

$$Lu^\varepsilon = F(u^\varepsilon)\xi^\varepsilon + F'(u^\varepsilon)F(u^\varepsilon)|\mathrm{D}_x\vartheta^\varepsilon|^2$$

will converge as soon as $\xi^\varepsilon \to \xi$ in $\mathscr{C}^{\alpha-2}$, without any requirements on the bilinear term $\vartheta^\varepsilon \circ \xi^\varepsilon$.

6 Relation with regularity structures

In [Hai14] Hairer introduces a general setup suitable to describe distributions which locally behave like a linear combination of a set of basic distributions. He calls this set a *model.* A *modelled distribution* is the result of patching up in a coherent fashion the local models according to a set of coefficients. At the core of his theory of regularity structures is the *reconstruction map* $\mathcal{R}$ which, for a given set of coefficients, delivers a modelled distribution which has the required local behavior up to small errors. In this section we review the concepts of model and modelled distribution and we use paracontrolled techniques to explicitly identify the modelled distributions as distributions paracontrolled by a given model, thus partially bridging the gap between the two theories. We conjecture that there is a complete correspondence between paracontrolled and modelled distributions however at this point this remains an open problem.

We denote by $(K_i)_{i\geqslant -1}$ the convolution kernels corresponding to the family of Littlewood–Paley projectors $(\Delta_i)_{i\geqslant -1}$, and we write $K_{<i} = \sum_{j<i} K_j$ and $K_{\leqslant i} = \sum_{j\leqslant i} K_j$. For any integral kernel V denote $V_x(y) = V(x-y)$ so for example $K_{i,x}(y) = K_i(x-y)$.

Let us briefly recall the basic setup for regularity structures. For more details the reader is referred to Hairer's original paper [Hai14].

Definition 6.1. *Let $A \subset \mathbb{R}$ be bounded from below and without accumulation points and let $T = \oplus_{\alpha\in A} T_\alpha$ be a vector space graded by A and such that T_α is a Banach space for all $\alpha \in A$. Let G be a group of continuous operators on T such that for all $\tau \in T_\alpha$ and $\Gamma \in G$ we have $\Gamma\tau - \tau \in \oplus_{\beta<\alpha} T_\beta$. The triple $\mathcal{T} = (A, T, G)$ is called a* regularity structure *with* model space T *and* structure group G.

For $\tau \in T$ we write $\|\tau\|_\alpha$ for the norm of the component of τ in T_α. We assume also that $0 \in A$ and $T_0 \simeq \mathbb{R}$ and that T_0 is invariant under G. Write also $\varphi_x^\lambda(y) = \lambda^{-d}\varphi((y-x)/\lambda)$.

Definition 6.2. *Given a regularity structure $\mathcal{T}$ and an integer $d \geqslant 1$, a model for $\mathcal{T}$ on $\mathbb{R}^d$ consists of maps*

$$\begin{array}{ll} \Pi : \mathbb{R}^d \to \mathcal{L}(T, \mathscr{S}'(\mathbb{R}^d)) & \qquad \Gamma : \mathbb{R}^d \times \mathbb{R}^d \to G \\ \qquad x \mapsto \Pi_x & \qquad\qquad (x,y) \mapsto \Gamma_{x,y} \end{array}$$

such that $\Gamma_{x,y}\Gamma_{y,z} = \Gamma_{x,z}$ and $\Pi_x\Gamma_{x,y} = \Pi_y$. Furthermore, given $r > |\min A|$, $\gamma > 0$, there exists a constant C such that the bounds

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

hold uniformly over $\varphi \in C_b^r(\mathbb{R}^d)$ with $\|\varphi\|_{C_b^r} \leqslant 1$ and with support in the unit ball of $\mathbb{R}^d$, $x, y \in \mathbb{R}^d$, $0 < \lambda \leqslant 1$ and $\tau \in T_\alpha$ with $\alpha \leqslant \gamma$ and $\beta < \alpha$.

In [Hai14], these conditions are only required to hold locally uniformly, that is for x, y contained in a compact subset of $\mathbb{R}^d$. To simplify the presentation and to facilitate the comparison with the paracontrolled approach, we will work here in the global framework.

Lemma 6.3. *Let φ be a Schwartz function, let $\gamma > 0$, and $r > |\min A|$. Then there exists $C(\varphi) > 0$ such that*

$$|(\Pi_x\tau)(\varphi_x^\lambda)| \leqslant C(\varphi)\lambda^\alpha\|\tau\|_\alpha$$

holds uniformly over $0 < \lambda \leqslant 1$ and $\tau \in T_\alpha$ with $\alpha \leqslant \gamma$. The constant $C(\varphi)$ can be chosen proportional to

$$\sup_{|\mu|\leqslant\lceil r\rceil}\sup_{x\in\mathbb{R}^d}(1+|x|)^{d+r+\gamma}|\partial^\mu\varphi(x)|.$$

Proof. We can decompose $\varphi = \sum_{k\in\mathbb{Z}^d}\varphi_k$, where every $\varphi_k \in C_c^\infty$ is supported in the ball with radius $\sqrt{d}$, centered at $k \in \mathbb{Z}^d$. Then $\psi = \sum_{|k|\leqslant\sqrt{d}+1}\varphi_k$ is a compactly supported smooth function, and therefore

$$|(\Pi_x\tau)(\psi_x^\lambda)| \lesssim_\varphi \lambda^\alpha\|\tau\|_\alpha.$$

For $|k| > \sqrt{d}+1$ we have $(\varphi_k)_x^\lambda = (\widetilde{\varphi}_k)_{x-k}^\lambda$ for $\widetilde{\varphi}_k$ supported in a ball centered at 0. Using that φ is a Schwartz function, we can estimate $\|(\widetilde{\varphi}_k)^\lambda\|_{C_b^r} \lesssim_\varphi \lambda^{-r-d}(|k|/\lambda)^{-(d+r+\alpha)}$. Therefore,

$$\begin{aligned}\sum_{|k|>\sqrt{d}+1}|(\Pi_x\tau)((\varphi_k)_x^\lambda)| &\lesssim \sum_{|k|>\sqrt{d}+1}|(\Pi_{x-k}\Gamma_{x-k,x}\tau)((\widetilde{\varphi}_k)_{x-k}^\lambda)| \\ &\lesssim_{\varphi,m} \sum_{|k|>\sqrt{d}+1}\sum_{\beta\leqslant\alpha}|k|^{\alpha-\beta}\|\tau\|_\alpha|k|^{-(d+r+\alpha)}\lambda^{-r-d+(d+r+\alpha)} \lesssim \|\tau\|_\alpha\lambda^\alpha.\end{aligned}$$

$\square$

Definition 6.4. *For $\gamma \in \mathbb{R}$, the set $\mathcal{D}^\gamma(\mathcal{T},\Gamma)$ consists of all functions $f^\pi\colon \mathbb{R}^d \to \oplus_{\alpha<\gamma}T_\alpha$ such that for every $\alpha < \gamma$ there exists a constant C with*

$$\|f_x^\pi - \Gamma_{x,y}f_y^\pi\|_\alpha \leqslant C|x-y|^{\gamma-\alpha}, \qquad \|f_x^\pi\|_\alpha \leqslant C,$$

uniformly over $x, y \in \mathbb{R}^d$.

6.1 The reconstruction operator

Definition 6.5. *Let $\gamma \in \mathbb{R}$ and $r > |\min A|$. A reconstruction $\mathcal{R}f^\pi$ of $f^\pi \in \mathcal{D}^\gamma(\mathcal{T},\Gamma)$ is a distribution such that*

$$|\mathcal{R}f^\pi(\varphi_x^\lambda) - \Pi_x f_x^\pi(\varphi_x^\lambda)| \lesssim \lambda^\gamma \tag{46}$$

for all $0 < \lambda \leqslant 1$, uniformly in $x \in \mathbb{R}^d$ and uniformly over $\varphi \in C_b^{r+\gamma}(\mathbb{R}^d)$ with $\|\varphi\|_{C_b^{r+\gamma}} \leqslant 1$ and with support in the unit ball of $\mathbb{R}^d$.

In [Hai14] inequality (46) is assumed to hold for all $\varphi \in C_b^r(\mathbb{R}^d)$ with $\|\varphi\|_{C_b^r} \leqslant 1$ and with support in the unit ball of $\mathbb{R}^d$. It should be possible to show that this follows from (46) and the definition of Π and $\mathcal{D}^\gamma(\mathcal{T},\Gamma)$. But for our purposes Definition 6.5 will be sufficient.

Lemma 6.6. *Property* (46) *is equivalent to*

$$|\mathcal{R}f^\pi(K_{<i,x}) - \Pi_x f^\pi_x(K_{<i,x})| \lesssim 2^{-i\gamma} \tag{47}$$

for all $i \geqslant 0$ *and* $x \in \mathbb{R}^d$.

Proof. Start by assuming (47). Lemma 6.3 yields $|\Pi_x f^\pi_x(K_{<i,x})| \lesssim 2^{-i\alpha_0}$, where $\alpha_0 = \min A$, and therefore $|\mathcal{R}f^\pi(K_{<i,x})| \lesssim 2^{-i\alpha_0}$. In particular, $\mathcal{R}f^\pi \in \mathscr{C}^{\alpha_0}$ and $|\mathcal{R}f^\pi(\psi)| \lesssim \|\psi\|_{C^r_b}$ for all $\psi \in C^r_b$. If now $\varphi \in C^{\gamma+r}_b$ is supported in the unit ball and if $i \geqslant 0$ is such that $2^{-i} \simeq \lambda$, then Lemma 6.3 yields

$$|(\mathcal{R}f^\pi - \Pi_x f^\pi_x)(\varphi^\lambda_x - S_i\varphi^\lambda_x)| \lesssim 2^{-i\gamma}\|\varphi\|_{C^{\gamma+r}_b} \lesssim \lambda^\gamma \|\varphi\|_{C^{\gamma+r}_b}.$$

Next, observe that

$$\begin{aligned}
(\mathcal{R}f^\pi - \Pi_x f^\pi_x)(S_i\varphi^\lambda_x) &= \int \mathrm{d}z (\mathcal{R}f^\pi - \Pi_x f^\pi_x)(K_{<i,z})\lambda^{-d}\varphi(\lambda^{-1}(x-z)) \\
&= \int \mathrm{d}z (\mathcal{R}f^\pi - \Pi_z f^\pi_z)(K_{<i,z})\lambda^{-d}\varphi(\lambda^{-1}(x-z)) \\
&\quad + \int \mathrm{d}z \Pi_z(f^\pi_z - \Gamma_{z,x} f^\pi_x)(K_{<i,z})\lambda^{-d}\varphi(\lambda^{-1}(x-z)).
\end{aligned}$$

In the second term of this sum we can estimate $|\Pi_z(f^\pi_z - \Gamma_{z,x}f^\pi_x)(K_{<i,z})| \lesssim \sum_{\beta<\gamma} 2^{-i\beta}|x-z|^{\gamma-\beta}$, where we used that $f^\pi \in \mathcal{D}^\gamma$. The first term in the sum is estimated using (47), giving

$$|(\mathcal{R}f^\pi - \Pi_x f^\pi_x)(S_i\varphi^\lambda_x)| \lesssim 2^{-i\gamma} + \sum_{\beta<\gamma} 2^{-i\beta} \int \mathrm{d}z |x-z|^{\gamma-\beta}\lambda^{-d}\varphi(\lambda^{-1}(z-x)) \lesssim 2^{-i\gamma}.$$

So requiring (47) is sufficient to have the general bound (46). To see that (46) implies (47) we can use similar arguments as in the proof of Lemma 6.3. □

The characterization of the reconstruction given by (47) is better suited for us, so we will stick with it in the following.

Lemma 6.7. *If* $\gamma > 0$, *the reconstruction operator is unique.*

Proof. Indeed, for the difference of two reconstructions $\mathcal{R}f^\pi$ and $\tilde{\mathcal{R}}f^\pi$ we have

$$\|S_i(\mathcal{R}f^\pi - \tilde{\mathcal{R}}f^\pi)\|_{L^\infty} \lesssim 2^{-i\gamma},$$

and therefore $0 = \lim_{i\to\infty} S_i(\mathcal{R}f^\pi - \tilde{\mathcal{R}}f^\pi) = \mathcal{R}f^\pi - \tilde{\mathcal{R}}f^\pi$. □

6.2 Paraproducts and modelled distributions

We are now going to generalize the paraproduct defined previously in order to apply it to a given model. Fix a model Π and for every $i \geqslant 0$ and $\gamma \in \mathbb{R}$ define the operator $P_i : \mathcal{D}^\gamma(\mathcal{T}, \Gamma) \to \mathcal{S}'(\mathbb{R}^d)$ by

$$P_i f^\pi(x) = \int \mathrm{d}z K_{<i-1,x}(z)\Pi_z f^\pi_z(K_{i,x}).$$

Note that

$$P_i f^\pi(x) = \int \mathrm{d}z K_{<i-1,x}(z)\Pi_x f^\pi_x(K_{i,x}) + \int \mathrm{d}z K_{<i-1,x}(z)\Pi_x(\Gamma_{x,z}f^\pi_z - f^\pi_x)(K_{i,x}) = \Pi_x f^\pi_x(K_{i,x}) + O(2^{-i\gamma})$$

for all $i \geqslant 1$, where we used that $\int \mathrm{d}z K_{<i-1,x}(z) = 1$, and where the estimate for the second integral follows from arguments similar to those used in Lemma 6.6. Now define the operator

$$Pf^\pi = P(f^\pi, \Pi) = \sum_{i \geqslant 0} P_i f^\pi$$

and note that this always gives a well defined distribution since every $P_i f^\pi$ is spectrally supported in an annulus $2^i \mathscr{A}$. In the particular case where $\Pi_z f_z^\pi(z') = a_z b(z')$, we get $P_i(f^\pi) = S_{i-1} a \Delta_i b$ and $Pf^\pi = a \prec b$, which justifies the claim that P is a generalization of the usual paraproduct.

The following lemma links Pf^π with the local behavior of the distribution $\Pi_x f_x^\pi$ around the point x.

Lemma 6.8. *Let $\gamma \in \mathbb{R}$ and $f^\pi \in \mathcal{D}^\gamma(\mathcal{T}, \Gamma)$ and set*

$$T_i f^\pi(x) = Pf^\pi(K_{i,x}) - \Pi_x f_x^\pi(K_{i,x})$$

for all $i \geqslant 0$. Then $\|T_i f^\pi\|_{L^\infty} \lesssim 2^{-i\gamma}$.

Proof. Observe that

$$Pf^\pi(K_{i,x}) = \sum_j (P_j f^\pi)(K_{i,x}) = \sum_{j:j\sim i} \int \mathrm{d}y \mathrm{d}z K_{i,x}(y) K_{<j-1,y}(z) \Pi_z f_z^\pi(K_{j,y})$$

and also that, since $\sum_{j:j\sim i} K_i * K_j = K_i$,

$$\Pi_x f_x^\pi(K_{i,x}) = \sum_{j:j\sim i} \int \mathrm{d}y K_{i,x}(y) \Pi_x f_x^\pi(K_{j,y}).$$

Using the decomposition $\Pi_z f_z^\pi(K_{j,y}) - \Pi_x f_x^\pi(K_{j,y}) = \Pi_y \Gamma_{y,z}(f_z^\pi - \Gamma_{z,x} f_x^\pi)(K_{j,y})$, we further have

$$T_i f^\pi(x) = Pf^\pi(K_{i,x}) - \Pi_x f_x^\pi(K_{i,x}) = \sum_{j:j\sim i} \int \mathrm{d}y \mathrm{d}z K_{i,x}(y) K_{<j-1,y}(z) \Pi_y \Gamma_{y,z}(f_z^\pi - \Gamma_{z,x} f_x^\pi)(K_{j,y})$$

from which the claimed bound can be shown to hold. Indeed, using the fact that $f^\pi \in \mathcal{D}^\gamma(\mathcal{T}, \Gamma)$ we obtain

$$\begin{aligned}
&\sum_{j:j\sim i} \left| \int \mathrm{d}y \mathrm{d}z K_{i,x}(y) K_{<j-1,y}(z) \Pi_y \Gamma_{y,z}(f_z^\pi - \Gamma_{z,x} f_x^\pi)(K_{j,y}) \right| \\
&\qquad \lesssim \sum_{j:j\sim i} \sum_{\beta<\gamma} \int \mathrm{d}y \mathrm{d}z \, |K_{i,x}(y) K_{<j-1,y}(z)| \, \|\Gamma_{y,z}(f_z^\pi - \Gamma_{z,x} f_x^\pi)\|_\beta 2^{-j\beta} \\
&\qquad \lesssim \sum_{j:j\sim i} \sum_{\beta<\gamma} \sum_{\alpha:\beta<\alpha<\gamma} \int \mathrm{d}y \mathrm{d}z \, |K_{i,x}(y) K_{<j-1,y}(z)| \, |y-z|^{\alpha-\beta} |z-x|^{\gamma-\alpha} 2^{-j\beta}.
\end{aligned}$$

Now it suffices to note that $|z-x|^{\gamma-\alpha} \lesssim |z-y|^{\gamma-\alpha} + |y-x|^{\gamma-\alpha}$, and the proof is complete. $\square$

Lemma 6.9. *Let $\gamma > 0$ and $f^\pi \in \mathcal{D}^\gamma(\mathcal{T}, \Gamma)$ and define*

$$Tf^\pi(x) = \sum_i T_i f^\pi(x) = \sum_i [Pf^\pi(K_{i,x}) - \Pi_x f_x^\pi(K_{i,x})].$$

Then $Tf^\pi \in \mathscr{C}^\gamma$.

Proof. According to Lemma 6.8, the series converges in L^∞. Let us analyze its regularity. Consider $\Delta_j T f^\pi = \sum_i \Delta_j T_i f^\pi$ and split the sum into two contributions, $\Delta_j T f^\pi = \Delta_j T_{\leqslant j+1} f^\pi + \Delta_j T_{>j+1} f^\pi$, where $T_{\leqslant j+1} f^\pi = \sum_{i\leqslant j+1} T_i f^\pi$ and $T_{>j+1} f^\pi = T f^\pi - T_{\leqslant j+1} f^\pi$. For the second term we have

$$\|\Delta_j T_{>j+1} f^\pi\|_{L^\infty} \leqslant \sum_{i>j+1} \|\Delta_j T_i f^\pi\|_{L^\infty} \lesssim \sum_{i>j+1} \|T_i f^\pi\|_{L^\infty} \lesssim 2^{-j\gamma}.$$

For the first one we proceed as follows. Note that $T_{\leqslant j+1} f^\pi(x) = P f^\pi(K_{\leqslant j+1,x}) - \Pi_x f^\pi_x(K_{\leqslant j+1,x})$, so that using $K_j * K_{\leqslant j+1} = K_j$ we get

$$\begin{aligned}
\Delta_j T_{\leqslant j+1} f^\pi(x) &= P f^\pi(K_{j,x}) - \int \mathrm{d}y K_{j,x}(y) \Pi_y f^\pi_y(K_{\leqslant j+1,y}) \\
&= P f^\pi(K_{j,x}) - \Pi_x f^\pi_x(K_{j,x}) - \int \mathrm{d}y K_{j,x}(y) \Pi_y (f^\pi_y - \Gamma_{y,x} f^\pi_x)(K_{\leqslant j+1,y}) \\
&= T_j f^\pi(x) - \int \mathrm{d}y K_{j,x}(y) \Pi_y (f^\pi_y - \Gamma_{y,x} f^\pi_x)(K_{\leqslant j+1,y}),
\end{aligned}$$

where in the last line we have used the definition of $T_j f^\pi$. Now

$$|\Pi_y (f^\pi_y - \Gamma_{y,x} f^\pi_x)(K_{\leqslant j+1,y})| \lesssim \sum_{\beta<\gamma} |y-x|^{\gamma-\beta} 2^{-j\beta},$$

so that $\|\Delta_j T f^\pi - T_j f^\pi\|_{L^\infty} \lesssim 2^{-j\gamma}$. This implies that $\|\Delta_j T f^\pi\|_{L^\infty} \lesssim 2^{-j\gamma}$ and thus concludes the proof. □

Theorem 6.10. *The reconstruction operator $\mathcal{R}$ exists for all $\gamma \in \mathbb{R} \setminus \{0\}$. If $\gamma > 0$ we have $\mathcal{R} = P - T$ while if $\gamma < 0$ we can take $\mathcal{R} = P$.*

Proof. In case $\gamma > 0$ set $\mathcal{R} f^\pi = P f^\pi - T f^\pi$ and observe that

$$\begin{aligned}
\mathcal{R} f^\pi(K_{<i,x}) - \Pi_x f^\pi_x(K_{<i,x}) &= P f^\pi(K_{<i,x}) - \Pi_x f^\pi_x(K_{<i,x}) - T f^\pi(K_{<i,x}) \\
&= T f^\pi(x) - (P f^\pi(K_{\geqslant i,x}) - \Pi_x f^\pi_x(K_{\geqslant i,x})) - T f^\pi(K_{<i,x}) \\
&= -P f^\pi(K_{\geqslant i,x}) + \Pi_x f^\pi_x(K_{\geqslant i,x}) + T f^\pi(K_{\geqslant i,x}) = \sum_{j\geqslant i} (\Delta_j T f^\pi(x) - T_j f^\pi(x)).
\end{aligned}$$

With the bounds of Lemma 6.8 and Lemma 6.9 we can conclude that

$$|\mathcal{R} f^\pi(K_{<i,x}) - \Pi_x f^\pi_x(K_{<i,x})| \lesssim 2^{-i\gamma},$$

which implies that $\mathcal{R}$ is the reconstruction operator. If $\gamma < 0$, just set $\mathcal{R} = P$ and observe that

$$|\mathcal{R} f^\pi(K_{<i,x}) - \Pi_x f^\pi_x(K_{<i,x})| \lesssim \sum_{j<i} |T_j f^\pi(x)| \lesssim \sum_{j<i} 2^{-j\gamma} \lesssim 2^{-i\gamma},$$

which shows that also in this case $\mathcal{R}$ is an admissible reconstruction operator. □

For $\gamma > 0$, we could say that a distribution f is *paracontrolled* by Π if there exist $f^\pi \in \mathcal{D}^\gamma(\mathcal{T}, \Gamma)$ and $f^\sharp \in \mathcal{C}^\gamma$ such that

$$f = P(f^\pi, \Pi) + f^\sharp;$$

in that case we write $f \in \mathcal{Q}^\gamma$. In particular, every modelled distribution is a paracontrolled distribution since the reconstruction map $\mathcal{R}$ delivers an injection

$$f^\pi \in \mathcal{D}^\gamma(\mathcal{T}, \Gamma) \longmapsto \mathcal{R} f^\pi = P(f^\pi, \Pi) - T f^\pi \in \mathcal{Q}^\gamma.$$

Moreover, every paracontrolled distribution can be decomposed into "slices", each of which has its natural regularity. More precisely, let us write τ^α for the component of $\tau \in T$ in T_α, for $\alpha < \gamma$. Then the distribution $P(f^\pi, \Pi)$ is given as

$$P(f^\pi, \Pi) = \sum_{i \geqslant 0} P_i f^\pi = \sum_{i \geqslant 0} \int \mathrm{d}z K_{<i-1,x}(z) \Pi_z f_z^\pi(K_{i,x}) = \sum_{\alpha<\gamma} \left(\sum_{i \geqslant 0} \int \mathrm{d}z K_{<i-1,x}(z) \Pi_x (\Gamma_{x,z} f_z^\pi)^\alpha (K_{i,x}) \right).$$

Now

$$\|\Gamma_{x,z} f_z^\pi\|_\alpha \lesssim \sum_{\beta:\alpha \leqslant \beta < \gamma} |x-z|^{\beta-\alpha} \|f_x^\pi\|_\beta \lesssim 1 + |x-z|^{\gamma-\alpha},$$

and similar arguments as in Lemma 6.6 show that $|\Pi_x \tau^\alpha(K_{i,x})| \lesssim 2^{-i\alpha}\|\tau\|_\alpha$ for all $\tau \in T$, $i \geqslant -1$. Combining these estimates with the fact that $\int \mathrm{d}z K_{<i-1,x}(z)\Pi_z f_z^{\pi,\alpha}(K_{i,x})$ is spectrally supported in an annulus $2^i \mathscr{A}$, we deduce that

$$\sum_{i \geqslant 0} \int \mathrm{d}z K_{<i-1,x}(z) \Pi_x (\Gamma_{x,z} f_z^\pi)^\alpha (K_{i,x}) \in \mathscr{C}^\alpha.$$

In particular, if $r = |\inf A|$, then every paracontrolled distribution is in $\mathscr{C}^{-r}$.

Note also that the paraproduct vanishes on constant and polynomial components of the model. Indeed, if τ is such that $\Pi_x \tau(y) = (y-x)^\mu$ for some $\mu \in \mathbb{N}^d$, then $P(\cdot, \tau) = 0$ since $(\Pi_x \tau)(K_{i,x}) = 0$ for any $i \geqslant 0$.

A Besov spaces and paraproducts

A.1 Littlewood-Paley theory and Besov spaces

In the following, we describe the concepts from Littlewood–Paley theory which are necessary for our analysis, and we recall the definition and some properties of Besov spaces. For a general introduction to Littlewood–Paley theory, Besov spaces, and paraproducts, we refer to the nice book of Bahouri, Chemin, and Danchin [BCD11].

Littlewood–Paley theory allows for an efficient way of characterizing the regularity of functions and distributions. It relies on the decomposition of an arbitrary distribution into a series of smooth functions whose Fourier transforms have localized support.

Let $\chi, \rho \in \mathscr{D}$ be nonnegative radial functions on $\mathbb{R}^d$, such that

i. the support of χ is contained in a ball and the support of ρ is contained in an annulus;

ii. $\chi(z) + \sum_{j \geqslant 0} \rho(2^{-j}z) = 1$ for all $z \in \mathbb{R}^d$;

iii. $\mathrm{supp}(\chi) \cap \mathrm{supp}(\rho(2^{-j}\cdot)) = \emptyset$ for $j \geqslant 1$ and $\mathrm{supp}(\rho(2^{-i}\cdot)) \cap \mathrm{supp}(\rho(2^{-j}\cdot)) = \emptyset$ for $|i-j| > 1$.

We call such (χ, ρ) *dyadic partition of unity*, and we frequently employ the notation

$$\rho_{-1} = \chi \qquad \text{and} \qquad \rho_j = \rho(2^{-j}\cdot) \text{ for } j \geqslant 0.$$

For the existence of dyadic partitions of unity see [BCD11], Proposition 2.10. The Littlewood–Paley blocks are now defined as

$$\Delta_{-1} u = \mathscr{F}^{-1}(\chi \mathscr{F} u) = \mathscr{F}^{-1}(\rho_{-1} \mathscr{F} u) \qquad \text{and} \qquad \Delta_j u = \mathscr{F}^{-1}(\rho_j \mathscr{F} u) \text{ for } j \geqslant 0.$$

Then $\Delta_j u = K_j * u$, where $K_j = \mathscr{F}^{-1}\rho_j$ for all $j \geqslant -1$. In particular, $\Delta_j u$ is a smooth function for every $j \geqslant -1$. We also use the notation

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

It is easy to see that $u = \sum_{j \geqslant -1} \Delta_j u = \lim_{j\to\infty} S_j u$ for every $u \in \mathscr{S}'$.

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

$$B^\alpha_{p,q}(\mathbb{R}^d, \mathbb{R}^n) = \left\{ u \in \mathscr{S}'(\mathbb{R}^d, \mathbb{R}^n) : \|u\|_{B^\alpha_{p,q}} = \left(\sum_{j \geqslant -1} (2^{j\alpha} \|\Delta_j u\|_{L^p})^q \right)^{1/q} < \infty \right\},$$

with the usual interpretation as ℓ^∞ norm in case $q = \infty$. The $\|\cdot\|_{L^p}$ norm is taken with respect to Lebesgue measure on $\mathbb{R}^d$. While the norm $\|\cdot\|_{B^\alpha_{p,q}}$ depends on the dyadic partition of unity (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 $\alpha \in (0, \infty) \backslash \mathbb{N}$, then $\mathscr{C}^\alpha$ is the space of $\lfloor \alpha \rfloor$ times differentiable functions, whose partial derivatives up to order $\lfloor \alpha \rfloor$ are bounded, and whose partial derivatives of order $\lfloor \alpha \rfloor$ are $(\alpha - \lfloor \alpha \rfloor)$-Hölder continuous (see p. 99 of [BCD11]). Note however, that for $k \in \mathbb{N}$ the Hölder-Besov space $\mathscr{C}^k$ is strictly larger than C^k_b.

We will use without comment that $\|\cdot\|_\alpha \leqslant \|\cdot\|_\beta$ for $\alpha \leqslant \beta$, that $\|\cdot\|_{L^\infty} \lesssim \|\cdot\|_\alpha$ for $\alpha > 0$, and that $\|\cdot\|_\alpha \lesssim \|\cdot\|_{L^\infty}$ for $\alpha \leqslant 0$. We will also use that $\|S_j u\|_{L^\infty} \lesssim 2^{j\alpha} \|u\|_\alpha$ for $\alpha < 0$ and $u \in \mathscr{C}^\alpha$.

We denote by $\mathscr{C}^\alpha_{\mathrm{loc}}$ the set of all distributions u such that $\varphi u \in \mathscr{C}^\alpha$ for all $\varphi \in \mathscr{D}$. If the difference $\varphi(u_n - u)$ converges to 0 in $\mathscr{C}^\alpha$ for all $\varphi \in \mathscr{D}$, then we say that (u_n) converges to u in $\mathscr{C}^\alpha_{\mathrm{loc}}$.

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

Lemma A.1 (Lemma 2.1 of [BCD11]). *Let $\mathscr{A}$ be an annulus and let $\mathscr{B}$ be a ball. For any $k \in \mathbb{N}$, $\lambda > 0$, and $1 \leqslant p \leqslant q \leqslant \infty$ we have that*

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

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

For example, it is a simple consequence of the Bernstein inequalities that $\|\mathrm{D}^k u\|_{\alpha - k} \lesssim \|u\|_\alpha$ for all $\alpha \in \mathbb{R}$ and $k \in \mathbb{N}$.

We point out that everything above and everything that follows can (and will) be applied to distributions on the torus. More precisely, let $\mathscr{D}'(\mathbb{T}^d)$ be the space of distributions on $\mathbb{T}^d$. Any $u \in \mathscr{D}'(\mathbb{T}^d)$ can be interpreted as a periodic tempered distribution on $\mathbb{R}^d$, with frequency spectrum contained in $\mathbb{Z}^d$ – and vice versa. For details see [ST87], Chapter 3.2. In particular, $\Delta_j u$ is a periodic smooth function, and therefore $\|\Delta_j u\|_{L^\infty} = \|\Delta_j u\|_{L^\infty(\mathbb{T}^d)}$. In other words, we can define

$$\mathscr{C}^\alpha(\mathbb{T}^d) = \{u \in \mathscr{C}^\alpha : u \text{ is } (2\pi) - \text{periodic}\}$$

for $\alpha \in \mathbb{R}$. For $p \neq \infty$ however, this definition is not very useful, because no nontrivial periodic function is in L^p for $p < \infty$. Therefore, general Besov spaces on the torus are defined as

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

where we set

$$\Delta_j u = (2\pi)^{-d} \sum_{k \in \mathbb{Z}^d} e^{\imath \langle k, x\rangle} \rho_j(k) (\mathscr{F}_{\mathbb{T}^d} u)(k) = \mathscr{F}_{\mathbb{T}^d}^{-1}(\rho_j \mathscr{F}_{\mathbb{T}^d} u),$$

and where $\mathscr{F}_{\mathbb{T}^d}$ and $\mathscr{F}_{\mathbb{T}^d}^{-1}$ denote Fourier transform and inverse Fourier transform on the torus. The two definitions are compatible: we have $\mathscr{C}^\alpha(\mathbb{T}^d) = B^\alpha_{\infty,\infty}(\mathbb{T}^d)$. Strictly speaking we will not work with $B^\alpha_{p,q}(\mathbb{T}^d)$ for $(p,q) \neq (\infty,\infty)$. But we will need the Besov embedding theorem on the torus.

Lemma A.2. *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 $B^\alpha_{p_1,q_1}(\mathbb{T}^d)$ is continuously embedded in $B^{\alpha - d(1/p_1 - 1/p_2)}_{p_2,q_2}(\mathbb{T}^d)$, and $B^\alpha_{p_1,q_1}(\mathbb{R}^d)$ is continuously embedded in $B^{\alpha - d(1/p_1 - 1/p_2)}_{p_2,q_2}(\mathbb{R}^d)$.*

For the embedding theorem on $\mathbb{R}^d$ see [BCD11], Proposition 2.71. The result on the torus can be shown using the same arguments, see for example [CG06]. In both cases, the proof is based on the Bernstein inequalities, Lemma A.1.

The following characterization of Besov regularity for functions which can be decomposed into pieces that are well localized in Fourier space will be useful below.

Lemma A.3. *(Lemmas 2.69 and 2.84 of [BCD11])*

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. It $\mathscr{F} u_j$ 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. □

A.2 Linear operators acting on Besov spaces

Here we discuss the action of some important linear operators on Besov spaces. We start with the rescaling of the spatial variable:

Lemma A.4. *For $\lambda > 0$ and $u \in \mathscr{S}'$ we define the scaling transformation $\Lambda_\lambda u(\cdot) = u(\lambda \cdot)$. Then*

$$\|\Lambda_\lambda u\|_\alpha \lesssim \max\{1, \lambda^\alpha\} \|u\|_\alpha$$

for all $\alpha \in \mathbb{R} \setminus \{0\}$ and all $u \in \mathscr{C}^\alpha$.

Proof. Let $u \in \mathscr{C}^\alpha$ and let $\Lambda_\lambda u(x) = u(\lambda x)$ for some $\lambda > 0$. Note that $\Lambda_\lambda \mathrm{D} = \lambda^{-1}\mathrm{D}\Lambda_\lambda$, and therefore $\Lambda_\lambda \Delta_j u = \Lambda_\lambda \rho(2^{-j}\mathrm{D})u = \rho(2^{-j}\lambda^{-1}\mathrm{D})\Lambda_\lambda u$, which implies that the Fourier transform of $\Lambda_\lambda \Delta_j u$ is supported in the annulus $\lambda 2^j \mathscr{A}$ (where $\mathscr{A}$ is the annulus in which ρ is supported). In particular, if $k \geqslant 0$, we have $\Delta_k \Lambda_\lambda \Delta_j u \neq 0$ only if $2^k \sim \lambda 2^j$. Thus, there exist $a, b > 0$ such that

$$\begin{aligned}\|\Delta_k \Lambda_\lambda u\|_{L^\infty} &\lesssim \sum_{j: a2^k \leqslant \lambda 2^j \leqslant b2^k} \|\Delta_k \Lambda_\lambda \Delta_j u\|_{L^\infty} \lesssim \sum_{j: a2^k \leqslant \lambda 2^j \leqslant b2^k} \|\Delta_j u\|_{L^\infty} \\ &\lesssim \|u\|_\alpha \sum_{j: a2^k \leqslant \lambda 2^j \leqslant b2^k} 2^{-\alpha j} \lesssim \|u\|_\alpha \lambda^\alpha 2^{-\alpha k}\end{aligned}$$

for all $k \geqslant 0$. For $k = -1$ we can simply bound

$$\|\Delta_{-1}\Lambda_\lambda u\|_{L^\infty} \lesssim \sum_{j:\lambda 2^j \lesssim 1} \|\Delta_k \Lambda_\lambda \Delta_j u\|_{L^\infty} \lesssim \|u\|_\alpha \sum_{j:\lambda 2^j \lesssim 1} 2^{-\alpha j} \lesssim \|u\|_\alpha \max\{1, \lambda^\alpha\}.$$

□

Next, we are concerned with the action of Fourier multipliers on Besov spaces.

Lemma A.5. *Let $\alpha \in \mathbb{R}$. Let φ be a continuous function, such that φ is infinitely differentiable everywhere except possibly at 0, and such that φ and all its partial derivatives decay faster than any rational function at infinity. Assume also that $\mathscr{F}\varphi \in L^1$. Then*

$$\|\varphi(\varepsilon \mathrm{D})u\|_{\alpha+\delta} \lesssim_{\delta,\varphi} \varepsilon^{-\delta}\|u\|_\alpha \qquad \text{and} \qquad \|\varphi(\varepsilon \mathrm{D})u\|_\delta \lesssim_{\delta,\varphi} \varepsilon^{-\delta}\|u\|_{L^\infty}.$$

for all $\varepsilon \in (0,1]$, $\delta \geqslant 0$, and $u \in \mathscr{S}'$.

Proof. Let $\psi \in \mathscr{D}$ with support in an annulus be such that $\psi\rho = \rho$, where (χ, ρ) is our dyadic partition of unity. Then we have for $j \geqslant 0$ that

$$\varphi(\varepsilon \mathrm{D})\Delta_j u = \left[\mathscr{F}^{-1}(\varphi(\varepsilon\cdot)\psi(2^{-j}\cdot))\right] * \Delta_j u,$$

and therefore Young's inequality implies

$$\|\varphi(\varepsilon \mathrm{D})\Delta_j u\|_{L^\infty} \lesssim \left\|\mathscr{F}^{-1}(\varphi(\varepsilon\cdot)\psi(2^{-j}\cdot))\right\|_{L^1} 2^{-j\alpha}\|u\|_\alpha = \left\|\mathscr{F}^{-1}(\varphi(2^j\varepsilon\cdot)\psi)\right\|_{L^1} 2^{-j\alpha}\|u\|_\alpha.$$

Hence, it suffices to prove $\left\|\mathscr{F}^{-1}(\varphi(2^j\varepsilon\cdot)\psi)\right\|_{L^1} \lesssim \varepsilon^{-\delta}2^{-j\delta}$. But

$$\begin{aligned}\left\|\mathscr{F}^{-1}(\varphi(2^j\varepsilon\cdot)\psi)\right\|_{L^1} &\lesssim \left\|(1+|\cdot|^2)^d \mathscr{F}^{-1}(\varphi(2^j\varepsilon\cdot)\psi)\right\|_{L^\infty} \lesssim \left\|\mathscr{F}^{-1}((1+\Delta)^d(\varphi(2^j\varepsilon\cdot)\psi))\right\|_{L^\infty} \\ &\lesssim \|(1+\Delta)^d(\varphi(2^j\varepsilon\cdot)\psi)\|_{L^1} \lesssim (1+2^j\varepsilon)^{2d} \max_{\mu\in\mathbb{N}^d:|\mu|\leqslant 2d} \|\partial^\mu\varphi(2^j\varepsilon\cdot)\|_{L^\infty(\mathrm{supp}(\psi))}.\end{aligned}$$

By assumption, φ is smooth away from 0, and φ and all its partial derivatives decay faster than any rational function at infinity. Thus, there exists $C = C(\varphi, \delta) > 0$ such that

$$\sup_{x\geqslant 1}(1+|x|)^{\delta+2d}|\partial^\mu\varphi(x)| \leqslant C$$

for all $\mu \in \mathbb{N}^d$ with $|\mu| \leqslant 2d$. Since $\mathrm{supp}(\psi)$ is bounded away from 0, there exists a minimal $j_0 \in \mathbb{N}$, such that $2^{j_0}\varepsilon|x| \geqslant 1$ for all $x \in \mathrm{supp}(\psi)$. Thus, we obtain

$$\left\|\mathscr{F}^{-1}(\varphi(2^j\varepsilon\cdot)\psi)\right\|_{L^1} \lesssim C(1+2^j\varepsilon)^{2d}(1+2^j\varepsilon)^{-\delta-2d} \lesssim_{\delta,\varphi} (1+2^j\varepsilon)^{-\delta} \lesssim 2^{-j\delta}\varepsilon^{-\delta}$$

for all $j \geqslant j_0$. On the other side, we get for $j \leqslant j_0$

$$\begin{aligned}\|\varphi(\varepsilon \mathrm{D})\Delta_j u\|_{L^\infty} &\lesssim \|\mathscr{F}^{-1}(\varphi(\varepsilon\cdot))\|_{L^1}\|\Delta_j u\|_{L^\infty} \lesssim_\varphi 2^{-j\alpha}\|u\|_\alpha = (\varepsilon 2^j)^\delta \varepsilon^{-\delta} 2^{-j(\alpha+\delta)}\|u\|_\alpha \\ &\leqslant (\varepsilon 2^{j_0})^\delta \varepsilon^{-\delta}2^{-j(\alpha+\delta)}\|u\|_\alpha \lesssim_\delta \varepsilon^{-\delta}2^{-j(\alpha+\delta)}\|u\|_\alpha,\end{aligned}$$

where we used that $\delta \geqslant 0$.

The estimate for $v \in L^\infty$ follows from the same arguments. □

Remark A.6. If the support of $\mathscr{F}u$ has a "hole" at 0, that is if there exists a ball $\mathscr{B}$ centered at 0 such that $\mathscr{F}u$ is supported outside of $\mathscr{B}$, then the estimates of Lemma A.5 hold uniformly in $\varepsilon > 0$ and not just for $\varepsilon \in (0, 1]$. This is an immediate consequence of the proof above.

As an application, we derive the smoothing properties of the heat kernel generated by the fractional Laplacian.

Lemma A.7. *Let $\sigma \in (0, 1]$, let $-(-\Delta)^\sigma$ be the fractional Laplacian with periodic boundary conditions on $\mathbb{T}^d$, and let $(P_t)_{t\geqslant 0}$ be the semigroup generated by $-(-\Delta)^\sigma$. Then for all $T > 0$, $t \in (0, T]$, $\alpha \in \mathbb{R}$, $\delta \geqslant 0$, and $u \in \mathscr{S}'$ we have*

$$\|P_t u\|_{\alpha+\delta} \lesssim_T t^{-\delta/(2\sigma)}\|u\|_\alpha \qquad \text{and} \qquad \|P_t v\|_\delta \lesssim_T t^{-\delta/(2\sigma)}\|v\|_{L^\infty}.$$

If $\mathscr{F}u$ is supported outside of a ball centered at 0, then these estimates are uniform in $t > 0$ and not just in $t \in (0, T]$.

Proof. The semigroup is given by $P_t = \varphi(t^{1/(2\sigma)}\mathrm{D})$ with $\varphi(z) = e^{-|z|^{2\sigma}}$. Now φ and its derivatives decay faster than any rational function at ∞. For $\sigma \leqslant 1$, $\mathscr{F}\varphi$ is the density of a symmetric 2σ-stable random variable, and therefore in L^1. For $\sigma > 1$ it is easily shown that $(1+|\cdot|^{d+1})\mathscr{F}\varphi$ is bounded, and therefore in L^1. Thus, the estimates follow from Lemma A.5. □

Lemma A.8. *Let σ and $(P_t)_{t\geqslant 0}$ be as in Lemma A.7. Let $\alpha \in \mathbb{R}$, $\beta \in [0, 1]$, and let $u \in \mathscr{C}^\alpha$. Then we have for all $t \geqslant 0$*

$$\|(P_t - \mathrm{Id})u\|_{L^\infty} \lesssim t^{\beta/(2\sigma)}\|u\|_\beta.$$

Proof. For the uniform estimate of $(P_t - \mathrm{Id})u$, we write $P_t - \mathrm{Id}$ as convolution operator: if $\varphi(z) = e^{-|z|^{2\sigma}}$ and $K(x) = \mathcal{F}^{-1}\varphi$, then

$$\begin{aligned}|(P_t - \mathrm{Id})u(x)| &= \left|t^{-d/(2\sigma)}\int K\left(\frac{x-y}{t^{1/(2\sigma)}}\right)(u(y) - u(x))\mathrm{d}y\right| \\ &\lesssim t^{-d/(2\sigma)}\int K\left(\frac{x-y}{t^{1/(2\sigma)}}\right)|y-x|^\beta\|u\|_\beta \mathrm{d}y \lesssim t^{\beta/(2\sigma)}\|u\|_\beta,\end{aligned}$$

where we identified $\mathscr{C}^\beta$ with the space of Hölder continuous functions. □

Based on Lemma A.7 and Lemma A.8, we derive the following Schauder estimates:

Lemma A.9. *Let σ and $(P_t)_{t\geqslant 0}$ be as in Lemma A.7. Assume that $v \in C_T\mathscr{C}^\beta$ for some $\beta \in \mathbb{R}$ and $T > 0$. Letting $V(t) = \int_0^t P_{t-s}v(s)\mathrm{d}s$, we have*

$$t^\gamma\|V(t)\|_{\beta+2\sigma} \lesssim \sup_{s\in[0,t]}(s^\gamma\|v(s)\|_\beta) \tag{48}$$

for all $\gamma \in [0, 1)$ and all $t \in [0, T]$. If $0 < \beta + 2\sigma < 1$, then we also have

$$\|V\|_{C_T^{(\beta+2)/(2\sigma)}L^\infty} \lesssim \sup_{s\in[0,t]}\|v(s)\|_\beta. \tag{49}$$

Proof. Consider $\Delta_q V$ for some $q \geqslant 0$ and let $\delta \in [0, t/2]$. We decompose the integral into two parts:

$$\Delta_q V(t) = \int_0^t P_{t-s}(\Delta_q v)(s)\mathrm{d}s = \int_0^\delta P_s(\Delta_q v)(t-s)\mathrm{d}s + \int_\delta^t P_s(\Delta_q v)(t-s)\mathrm{d}s.$$

Letting $M = \sup_{s\in[0,t]}(s^{\gamma}\|v(s)\|_{\beta})$, we estimate the first term by

$$\left\|\int_0^{\delta} P_s(\Delta_q v)(t-s)\mathrm{d}s\right\|_{L^\infty} \leqslant \int_0^{\delta} 2^{-q\beta}\|v(t-s)\|_{\beta}\mathrm{d}s \leqslant 2^{-q\beta}M\int_0^{\delta}(t-s)^{-\gamma}\mathrm{d}s$$
$$= M2^{-q\beta}t^{1-\gamma}\int_0^{\delta/t}\frac{\mathrm{d}s}{(1-s)^{\gamma}} \lesssim M2^{-q\beta}t^{-\gamma}\delta,$$

using $|1-(1-\delta/t)^{1-\gamma}| \lesssim \delta/t$ in the last step. On the other side, we can use Lemma A.7 to estimate the second term for $\varepsilon > 0$ by

$$\left\|\int_{\delta}^{t} P_s(\Delta_q v)(t-s)\mathrm{d}s\right\|_{L^\infty} \lesssim \int_{\delta}^{t} s^{-1-\varepsilon}2^{-q(\beta+2\sigma(1+\varepsilon))}\|v(t-s)\|_{\beta}\mathrm{d}s$$
$$\lesssim M2^{-q(\beta+2\sigma(1+\varepsilon))}\int_{\delta}^{t}\frac{\mathrm{d}s}{s^{1+\varepsilon}(t-s)^{\gamma}} = M2^{-q(\beta+2\sigma(1+\varepsilon))}t^{-\varepsilon-\gamma}\int_{\delta/t}^{1}\frac{\mathrm{d}s}{s^{1+\varepsilon}(1-s)^{\gamma}}$$
$$\lesssim M2^{-q(\beta+2\sigma(1+\varepsilon))}t^{-\gamma}\delta^{-\varepsilon} = M2^{-q(\beta+2\sigma)}(2^{q2\sigma}\delta)^{-\varepsilon}t^{-\gamma}.$$

If $2^{-q2\sigma} \leqslant t/2$, we can take $\delta = 2^{-q2\sigma}$ to obtain $\|\Delta_q V(t)\|_{L^\infty} \lesssim Mt^{-\gamma}2^{-q(\beta+2\sigma)}$. If $2^{-q2\sigma} > t/2$, we have $\|\Delta_q V(t)\|_{L^\infty} \lesssim M2^{-q\beta}t^{1-\gamma} \lesssim Mt^{-\gamma}2^{-q(\beta+2\sigma)}$, and the first claim follows.

As for the second claim, note that for $0 \leqslant s < t \leqslant T$ we have

$$V(t) - V(s) = (P_{t-s} - \mathrm{Id})V(s) + \int_s^t P_{t-r}v(r)\mathrm{d}r,$$

and therefore we can apply Lemma A.8 to obtain

$$\|V(t)-V(s)\|_{L^\infty} \lesssim \|(P_{t-s}-\mathrm{Id})V(s)\|_{L^\infty} + \int_s^t \|P_{t-r}v(r)\|_{L^\infty}\mathrm{d}r$$
$$\lesssim |t-s|^{(\beta+2)/(2\sigma)}\|V(s)\|_{\beta+2} + \int_s^t \|v(r)\|_{\beta}\mathrm{d}r \lesssim_T |t-s|^{(\beta+2)/(2\sigma)}\sup_{r\in[0,t]}\|v(r)\|_{\beta},$$

where we used that $(\beta+2)\in(0,1)$ and that $|t-s| \leqslant T$. This yields the second claim. $\square$

When dealing with RDEs, the convolution with the (fractional) heat kernel has a natural correspondence in the integral map.

Lemma A.10. *Let $u \in \mathscr{C}^{\alpha-1}(\mathbb{R})$ for some $\alpha\in(0,1)$. Then there exists a unique $U \in \mathscr{C}^{\alpha}_{\mathrm{loc}}(\mathbb{R})$ such that $\mathrm{D}U = u$ and $U(0) = 0$. This antiderivative U satisfies*

$$|U(t)-U(s)| \lesssim |t-s|^{\alpha}\|u\|_{\alpha-1} \tag{50}$$

for all $s,t\in\mathbb{R}$ with $|s-t| \leqslant 1$. We will use the notation $U(t) = \int_0^t u(s)\mathrm{d}s$ to denote this map, which is an extension of the usual definite integral. If the support of u is contained in $[-T,T]$ for some $T>0$, then $U \in \mathscr{C}^{\alpha}$ and

$$\|U\|_{\alpha} \lesssim T\|u\|_{\alpha-1}.$$

Proof. The second statement about compactly supported u follows from the first statement by identifying $\mathscr{C}^{\alpha}$ with the space of bounded Hölder continuous functions.

As for the first statement, we define

$$U(t) = \sum_{j\geqslant -1}\int_0^t \Delta_j u(s)\mathrm{d}s.$$

If we can show (50), then U is indeed in $\mathscr{C}^\alpha_{\mathrm{loc}}$ and therefore in particular in $\mathscr{S}'$. Since the derivative D is a continuous operator on $\mathscr{S}'$, we then conclude that $\mathrm{D}U = \sum_j \Delta_j u = u$. Let therefore $s, t \in \mathbb{R}$ with $|s-t| \leqslant 1$. We have

$$\left| \int_s^t \Delta_j u(r) \mathrm{d}r \right| \leqslant 2^{j(1-\alpha)} \|u\|_{\alpha-1} |t-s|.$$

If $j \geqslant 0$, then $\Delta_j u = \mathrm{D}\mathrm{D}^{-1}(\Delta_j u)$, where D^{-1} is the Fourier multiplier with symbol $1/(\iota z)$, and therefore

$$\left| \int_s^t \Delta_j u(r) \mathrm{d}r \right| = |\mathrm{D}^{-1}\Delta_j u(t) - \mathrm{D}^{-1}\Delta_j u(s)| \lesssim 2^{-j} \|\Delta_j u\|_{L^\infty} \lesssim 2^{-j\alpha} \|u\|_{\alpha-1},$$

where we used the Bernstein inequality, Lemma A.1. If j_0 is such that $2^{-j_0} \leqslant |t-s| < 2^{-j_0+1}$, then we use the first estimate for $j \leqslant j_0$ and the second estimate for $j > j_0$, and obtain

$$\begin{aligned} |U(t) - U(s)| &\leqslant \sum_{j \geqslant -1} \left| \int_s^t \Delta_j u(r) \mathrm{d}r \right| \lesssim \sum_{j \leqslant j_0} 2^{j(1-\alpha)} \|u\|_{\alpha-1} |t-s| + \sum_{j > j_0} 2^{-j\alpha} \|u\|_{\alpha-1} \\ &\lesssim (2^{j_0(1-\alpha)} |t-s| + 2^{-j_0\alpha}) \|u\|_{\alpha-1} \simeq |t-s|^\alpha \|u\|_{\alpha-1}. \end{aligned}$$

Uniqueness is easy since every distribution with zero derivative is a constant function. □

B Some more commutator estimates

When applying the scaling argument to solve equations, we need to control the resonant product of the rescaled data. This can be done by relying on the following commutator estimate.

Lemma B.1. *Let $\alpha, \beta \in \mathbb{R}$ and $f, g \in \mathscr{S}$. Then we have uniformly in $\lambda \in (0,1]$*

$$\|\Lambda_\lambda(f \circ g) - (\Lambda_\lambda f) \circ (\Lambda_\lambda g)\|_{\alpha+\beta} \lesssim \max\{\lambda^{\alpha+\beta}, 1\} \|f\|_\alpha \|g\|_\beta,$$

and thus $\Lambda_\lambda(\cdot \circ \cdot) - (\Lambda_\lambda \cdot) \circ (\Lambda_\lambda \cdot)$ extends to a bounded bilinear operator from $\mathscr{C}^\alpha \times \mathscr{C}^\beta$ to $\mathscr{C}^{\alpha+\beta}$.

Proof. We have $\Lambda_\lambda \Delta_j = \Lambda_\lambda \rho_j(\mathrm{D}) = \rho_j(\lambda^{-1}\mathrm{D})\Lambda_\lambda$ for all $j \geqslant -1$. Let $k \in \mathbb{N}$ and $\lambda' \in (1/2, 1]$ be such that $\lambda = \lambda' 2^{-k}$. Then

$$\Lambda_\lambda(f \circ g) = \sum_{\substack{|i-j| \leqslant 1 \\ i,j \leqslant k}} \Lambda_\lambda(\Delta_i f \Delta_j g) + \sum_{\substack{|i-j| \leqslant 1 \\ i,j > k}} \rho(2^{-i+k}\lambda'^{-1}\mathrm{D})\Lambda_\lambda f \rho(2^{-j+k}\lambda'^{-1}\mathrm{D})\Lambda_\lambda g. \tag{51}$$

The first sum is spectrally supported in a ball centered at zero (which is independent of k), and therefore

$$\Bigg\| \sum_{\substack{|i-j| \leqslant 1 \\ i,j \leqslant k}} \Lambda_\lambda(\Delta_i f \Delta_j g) \Bigg\|_{\alpha+\beta} \lesssim \sum_{\substack{|i-j| \leqslant 1 \\ i,j \leqslant k}} 2^{-i\alpha - j\beta} \|f\|_\alpha \|g\|_\beta \lesssim \max\{\lambda^{\alpha+\beta}, 1\} \|f\|_\alpha \|g\|_\beta.$$

The second sum is the resonant paraproduct $(\Lambda_\lambda f \,\tilde{\circ}\, \Lambda_\lambda g)$ with respect to the dyadic partition of unity $(\chi(\lambda'^{-1}\cdot), \rho(\lambda'^{-1}\cdot))$, except that the sum only starts in $i, j = 1$. By Lemma 3.5 we can therefore bound

$$\Bigg\| \sum_{\substack{|i-j| \leqslant 1 \\ i,j > k}} \rho(2^{-i+k}\lambda'^{-1}\mathrm{D})\Lambda_\lambda f \rho(2^{-j+k}\lambda'^{-1}\mathrm{D})\Lambda_\lambda g - (\Lambda_\lambda f) \circ (\Lambda_\lambda g) \Bigg\|_{\alpha+\beta} \lesssim \|f\|_\alpha \|g\|_\beta.$$

□

Next, we prove that it is possible to "pull the time integral inside the paraproduct":

Lemma B.2. *Let $\alpha, \beta \in (0,1)$ with $\alpha+\beta<1$. Let $u \in \mathscr{C}^\alpha(\mathbb{R}, \mathbb{R}^{d\times n})$ and $v \in \mathscr{C}^\beta(\mathbb{R}, \mathbb{R}^n)$. Then*

$$\left|\int_s^t (u \prec \partial_t v)(r)\mathrm{d}r - u(s)(v(t)-v(s))\right| \lesssim |t-s|^{\alpha+\beta}\|u\|_\alpha \|v\|_\beta,$$

for all $s,t \in \mathbb{R}$ with $|t-s| \leqslant 1$, where we write $\int_s^t f(r)\mathrm{d}r = \int_0^t f(r)\mathrm{d}r - \int_0^s f(r)\mathrm{d}r$.

Proof. Fix $s,t \in \mathbb{R}$ with $|s-t| \leqslant 1$. We can rewrite

$$\int_s^t (u \prec \partial_t v)(r)\mathrm{d}r - u(s)(v(t)-v(s)) = \sum_j \int_s^t [S_{j-1}u(r) - u(s)]\partial_r \Delta_j v(r)\mathrm{d}r.$$

We will use two different estimates, one for large j and one for small j. First note that

$$\begin{aligned}\left|\int_s^t [S_{j-1}u(r) - u(s)]\partial_r \Delta_j v(r)\mathrm{d}r\right| &\leqslant \left|\int_s^t [S_{j-1}u(r) - S_{j-1}u(s)]\partial_r \Delta_j v(r)\mathrm{d}r\right| \\ &\quad + \left|\int_s^t [S_{j-1}u(s) - u(s)]\partial_r \Delta_j v(r)\mathrm{d}r\right|.\end{aligned}$$

Now $|S_{j-1}u(r) - S_{j-1}u(s)| \lesssim |r-s|^\alpha \|u\|_\alpha$, and therefore

$$\begin{aligned}\left|\int_s^t [S_{j-1}u(r) - u(s)]\partial_r \Delta_j v(r)\mathrm{d}r\right| &\lesssim \left(\int_s^t |r-s|^\alpha 2^{j(1-\beta)}\mathrm{d}r + \int_s^t 2^{-j\alpha}2^{j(1-\beta)}\mathrm{d}r\right)\|u\|_\alpha\|v\|_\beta \\ &\lesssim (2^{j(1-\beta)}|t-s|^{1+\alpha} + 2^{j(1-\alpha-\beta)}|t-s|)\|u\|_\alpha\|v\|_\beta. \qquad (52)\end{aligned}$$

On the other side, it follows from integration by parts that

$$\begin{aligned}\left|\int_s^t [S_{j-1}u(r) - u(s)]\partial_r \Delta_j v(r)\mathrm{d}r\right| &\leqslant \left|\int_s^t [S_{j-1}u(r) - S_{j-1}u(s)]\partial_r \Delta_j v(r)\mathrm{d}r\right| \qquad (53) \\ &\quad + \left|\int_s^t [S_{j-1}u(s) - u(s)]\partial_r \Delta_j v(r)\mathrm{d}r\right| \\ &\leqslant |(S_{j-1}u(t) - S_{j-1}u(s))\Delta_j v(t)| + \left|\int_s^t \partial_r S_{j-1}u(r)\Delta_j v(r)\mathrm{d}r\right| \\ &\quad + |(S_{j-1}u(s) - u(s))(\Delta_j v(t) - \Delta_j v(s))| \\ &\lesssim \left(|t-s|^\alpha 2^{-j\beta} + |t-s|^{\alpha+\beta-\varepsilon}2^{-j\varepsilon} + 2^{-j(\alpha+\beta)}\right)\|u\|_\alpha\|v\|_\beta,\end{aligned}$$

for all $\varepsilon \in [0, \alpha+\beta)$, where for the middle term we applied Lemma A.10, which gives us

$$\begin{aligned}\left|\int_s^t \partial_r S_{j-1}u(r)\Delta_j v(r)\mathrm{d}r\right| &\lesssim |t-s|^{\alpha+\beta-\varepsilon}\|\partial_r S_{j-1}u(r)\Delta_j v(r)\|_{\alpha+\beta-\varepsilon-1} \\ &\lesssim |t-s|^{\alpha+\beta-\varepsilon}2^{j(\alpha+\beta-\varepsilon-1)}\|\partial_r S_{j-1}u(r)\Delta_j v(r)\|_{L^\infty} \\ &\lesssim |t-s|^{\alpha+\beta-\varepsilon}2^{-j\varepsilon}\|u\|_\alpha\|v\|_\beta.\end{aligned}$$

Let now $j_0 \in \mathbb{N}$ be such that $2^{-j_0} \leqslant |t-s| < 2^{-j_0+1}$. We use estimate (52) for $j \leqslant j_0$ and (53) for $j > j_0$ to obtain

$$\begin{aligned}\left|\int_s^t (u \prec \partial_t v)(r)\mathrm{d}r - u(s)(v(t)-v(s))\right| &\lesssim \sum_{j\leqslant j_0}(2^{j(1-\beta)}|t-s|^{1+\alpha} + 2^{j(1-\alpha-\beta)}|t-s|)\|u\|_\alpha\|v\|_\beta \\ &\quad + \sum_{j>j_0}(|t-s|^\alpha 2^{-j\beta} + |t-s|^{\alpha+\beta-\varepsilon}2^{-j\varepsilon} + 2^{-j(\alpha+\beta)})\|u\|_\alpha\|v\|_\beta \\ &\simeq \|u\|_\alpha\|v\|_\beta|t-s|^{\alpha+\beta},\end{aligned}$$

where we used that $\alpha+\beta<1$. □

References

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

[BGN13] Zdzisław Brzeźniak, Massimiliano Gubinelli, and Misha Neklyudov. Global evolution of random vortex filament equation. *Nonlinearity*, 26(9):2499, 2013.

[BGR05] Hakima Bessaih, Massimiliano Gubinelli, and Francesco Russo. The evolution of a random vortex filament. *Ann. Probab.*, 33(5):1825–1855, 2005.

[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.

[CC13] Rémi Catellier and Khalil Chouk. Paracontrolled distributions and the 3-dimensional stochastic quantization equation. *arXiv preprint arXiv:1310.6869*, 2013.

[CF09] Michael Caruana and Peter Friz. Partial differential equations driven by rough paths. *J. Differential Equations*, 247(1):140–173, 2009.

[CFO11] Michael Caruana, Peter K. Friz, and Harald Oberhauser. A (rough) pathwise approach to a class of non-linear stochastic partial differential equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(1):27–46, 2011.

[CG06] Jean-Yves Chemin and Isabelle Gallagher. On the global wellposedness of the 3-D Navier-Stokes equations with large initial data. *Ann. Sci. École Norm. Sup. (4)*, 39(4):679–698, 2006.

[CG13] Khalil Chouk and Massimiliano Gubinelli. Rough sheets. preprint, 2013.

[DF12] Joscha Diehl and Peter Friz. Backward stochastic differential equations with rough drivers. *Ann. Probab.*, 40(4):1715–1758, 2012.

[DGT12] Aurélien Deya, Massimiliano Gubinelli, and Samy Tindel. Non-linear rough heat equations. *Probab. Theory Related Fields*, 153(1-2):97–147, 2012.

[FGGR12] Peter K Friz, Benjamin Gess, Archil Gulisashvili, and Sebastian Riedel. Spatial rough path lifts of stochastic convolutions. *arXiv preprint arXiv:1211.0046*, 2012.

[FO11] Peter Friz and Harald Oberhauser. On the splitting-up method for rough (partial) differential equations. *J. Differential Equations*, 251(2):316–338, 2011.

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

[GIP14] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. A Fourier approach to pathwise integration. in preparation, 2014.

[GLT06] Massimiliano Gubinelli, Antoine Lejay, and Samy Tindel. Young integrals and SPDEs. *Potential Anal.*, 25(4):307–326, 2006.

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

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

[Gub12] Massimiliano Gubinelli. Rough solutions for the periodic Korteweg–de Vries equation. *Commun. Pure Appl. Anal.*, 11(2):709–733, 2012.

[Hai11] Martin Hairer. Rough stochastic PDEs. *Comm. Pure Appl. Math.*, 64(11):1547–1585, 2011.

[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. to appear.

[HMW12] Martin Hairer, Jan Maas, and Hendrik Weber. Approximating rough stochastic PDEs. *arXiv preprint arXiv:1202.3094*, 2012.

[HW13] Martin Hairer and Hendrik Weber. Rough Burgers-like equations with multiplicative noise. *Probab. Theory Related Fields*, 155(1-2):71–126, 2013.

[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.

[NT11] David Nualart and Samy Tindel. A construction of the rough path above fractional Brownian motion using Volterra's representation. *Ann. Probab.*, 39(3):1061–1096, 2011.

[Per14] Nicolas Perkowski. *Studies of Robustness in Stochastic Analysis and Mathematical Finance*. PhD thesis, Humboldt-Universität zu Berlin, 2014.

[ST87] Hans-Jürgen Schmeisser and Hans Triebel. *Topics in Fourier analysis and function spaces*, volume 42. Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig, 1987.

[Tei11] Josef Teichmann. Another approach to some rough and stochastic partial differential equations. *Stoch. Dyn.*, 11(2-3):535–550, 2011.

[Tri06] Hans Triebel. *Theory of function spaces. III*, volume 100 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 2006.

[Unt10a] Jérémie Unterberger. A rough path over multidimensional fractional Brownian motion with arbitrary Hurst index by Fourier normal ordering. *Stochastic Processes and their Applications*, 120(8):1444–1472, 2010.

[Unt10b] Jérémie Unterberger. Hölder-Continuous Rough Paths by Fourier Normal Ordering. *Comm. Math. Phys.*, 298(1):1–36, 2010.