Paracontrolled distributions and singular diffusions

NICOLAS PERKOWSKI

Abstract

We introduce paracontrolled distribution on the example of a stochastic differential equation with distributional drift.

These are the incomplete (in particular nearly no references!) and very unpolished notes for a mini course taught at the HCM winter school "Recent development in singular SPDEs" in February 2017 in Bonn. Participants may also want to consult the notes "Lectures on singular SPDEs" with M. Gubinelli which contain much more technical detail and also more references.

1 Motivation

1.1 Scaling limit for a random walk in a random potential

Consider a N-step simple random walk on $\mathbb{Z}^d$ that is weighted by a random potential. That is, we look at the random probability measure on $(\mathbb{Z}^d)^N$ given by

$$\frac{\mathrm{d}\mathbb{Q}_N^{\beta,\eta}}{\mathrm{d}\mathbb{P}_N} = \frac{1}{Z_N^{\beta,\eta}} \exp\left(\beta \sum_{k=0}^{N-1} \eta(X_k)\right),$$

where:

- $X_k(\omega) = \omega(k)$ is the coordinate process,
- $Z_N^{\beta,\eta} > 0$ is a normalization constant which sets the total mass of $\mathbb{Q}_N^{\beta,\eta}$ equal to 1,
- $(\eta(x))_{x \in \mathbb{Z}^d}$ are i.i.d. random variables,
- $\mathbb{P}_N$ is the distribution of a simple random walk on $\mathbb{Z}^d$ with N steps and that starts at $X_0 = 0$,
- $\beta > 0$ is a parameter called the *disorder strength.*

The potential has the effect of favoring paths that spend much time near points x where $\eta(x)$ is large. On the other hand the law of the iterated logarithm shows that $\mathbb{P}_N$ only gives very little mass to paths that leave the ball $(-N^\alpha, N^\alpha)^d$ with $\alpha > 1/2$. It is therefore interesting to understand the behavior of typical trajectories under $\mathbb{Q}_N^{\beta,\eta}$. If there exists a very favorable region where η is large, but it is outside of $(-N^\alpha, N^\alpha)^d$, will the trajectories under $\mathbb{Q}_N^{\beta,\eta}$ typically visit that region? Or is it too "expensive" and the trajectories will look more or less like those under $\mathbb{P}_N$? Intuitively we might picture that typical trajectories identify the most favorable region in $(-N, N)^d$, taking into account both the value of η and the probability of visiting that region under $\mathbb{P}_N$, and then go to that region as quickly as possible and spend most of their time there. Obviously the behavior will also depend on β: For $\beta = 0$ we simply obtain the law of the random walk, while for $\beta \gg 0$ the behavior will be radically different than that of the random walk.

This model should also be compared with models for random polymers, which are of the form

$$\frac{1}{Z_N}\exp(-\beta\mathcal{H}_N(\omega))\mathbb{P}_N(\mathrm{d}\omega),$$

where $\mathcal{H}_N$ is a so called *Hamiltonian*. An important example for the Hamiltonian is

$$\mathcal{H}_N(\omega)=\begin{cases} 0, & \omega\in\mathrm{SAW},\\ \infty, & \omega\notin\mathrm{SAW},\end{cases}$$

with $\mathrm{SAW}=\{\omega\in(\mathbb{Z}^d)^N\colon X_k(\omega)\neq X_\ell(\omega) \text{ for all } k\neq\ell\}$. In that case we obtain the self-avoiding random walk, independently of β.

For background material on random polymers see e.g. den Hollander or Caravenna, den Hollander, Petrelis. Due to the clustering that we expect under $\mathbb{Q}_N^{\beta,\eta}$, the trajectories may not really resemble a polymer (which is a long chain of molecules), so it is not obvious to me if $\mathbb{Q}_N^{\beta,\eta}$ is a good model for a polymer. But somtimes $\mathbb{Q}_N^{\beta,\eta}$ is called a *random polymer measure*.

We would like to understand how the trajectories under $\mathbb{Q}_N^{\beta,\eta}$ look on large temporal and spatial scales. That is, we would like to send $N\to\infty$ and rescale space as we pass to the limit to obtain a continuous-time limiting process, maybe a diffusion. However, given the clustering that we alluded to above we expect that on the time scale N the trajectories will identify the most favorable region in $(-N,N)^d$ and spend most of their time there. But on the time scale $2N$ the trajectories will spend most their time in the most favorable region in $(-2N,2N)^d$, which may be not at all compatible with the one in $(-N,N)^d$. So there does not seem to be a stabilization and the best we could hope for is that along subsequences the coordinate process converges to a jump process that instantly jumps to a fixed point and then just stays there. I am not sure if such a statement is exactly true and/or proven, but results in that direction are well established in the context of the *parabolic Anderson model* which roughly speaking descrives the transition function of $\mathbb{Q}_N$, see the recent survey König. So while the convergence statement itself is interesting, the limiting process will be quite boring. If we want to obtain a more interesting limit we should tune down the effect of the random potential as we pass to the limit, i.e. make the disorder strength $\beta=\beta_N$ depend on N and send it to zero as $N\to\infty$. We expect that:

- Too slow convergence $\beta_N\to 0$ gives the same qualitative behavior as for fixed $\beta_N\equiv\beta>0$.
- Too fast convergence $\beta_N\to 0$ gives the same qualitative behavior as for $\beta_N=0$, that is the trajectories converge to a d-dimensional Brownian motion.
- There exists (at least one) critical scaling of β_N where a phase transition occurs between the two regimes.

Let us try to understand what the critical scaling for β_N should be. Under $\mathbb{P}_{N^2}$ we have to rescale the coordinate process X to $X_t^N=N^{-1}X_{\lfloor N^2 t\rfloor}$ to get a process that converges to the Brownian motion. Let us also extend η piecewise constantly by setting $\eta|_{[k_1,k_1+1)\times\ldots\times[k_d,k_d+1)}=\eta((k_1,\ldots,k_d))$ and define $\eta^N(x)=N^{d/2}\eta(Nx)$, $x\in\mathbb{R}^d$.

Lemma 1. *Let us interpret the map* $\varphi \mapsto \int_{\mathbb{R}^d} \eta^N(x)\varphi(x)\mathrm{d}x =: \eta^N(\varphi)$ *as a distribution in* $\mathcal{D}'$*, the dual space of* $C_c^\infty = \mathcal{D}$*. Assume that* $\eta(0)$ *is a centered random variable with variance* 1*. Then* η^N *converges in distribution in* $\mathcal{D}'$ *to a centered Gaussian process* $(\xi(\varphi))_{\varphi \in C_c^\infty}$ *with covariance*

$$\mathbb{E}[\xi(\varphi)\xi(\psi)] = \langle \varphi, \psi \rangle_{L^2}. \tag{1}$$

This process is called the (space) white noise on $\mathbb{R}^d$

Proof. By general results ("Mitoma's criterion") it suffices to show that for $\varphi_1, \ldots, \varphi_n \in C_c^\infty$ the vector

$$\left(\int_{\mathbb{R}^d} \eta^N(x)\varphi_1(x)\mathrm{d}x, \ldots, \int_{\mathbb{R}^d} \eta^N(x)\varphi_n(x)\mathrm{d}x \right)$$

converges to a multivariate centered Gaussian random variable with the covariance (1). By taking linear combinations we can restrict our attention to the one-dimensional case, and for $\varphi \in C_c^\infty$ we have

$$\int_{\mathbb{R}^d} \eta^N(x)\varphi(x)\mathrm{d}x \simeq N^{-d} \sum_{x \in \mathbb{Z}^d} N^{d/2}\eta(x)\varphi(x/N) = N^{-d/2} \sum_{x \in \mathbb{Z}^d} \eta(x)\varphi(x/N).$$

There are $O(N^d)$ values of $x \in \mathbb{Z}^d$ for which x/N is in the support of $\varphi(x/N)$, so by the central limit theorem this integral converges to a centered Gaussian random variable $\xi(\varphi)$ with variance $\mathbb{E}[|\xi(\varphi)|^2] = \|\varphi\|_{L^2}^2$. □

So both η^N and X^N converge to nontrivial limits under the above scaling. Let us see formally how this translates for $\mathbb{Q}_N^{\beta_N, \eta}$. We have

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{Q}_{N^2}^{\beta_N,\eta}}{\mathrm{d}\mathbb{P}_{N^2}} &= \frac{1}{Z_N^{\beta,\eta}} \exp\left(\beta_N \sum_{k=0}^{N^2-1} \eta(X_k) \right) = \frac{1}{Z_N^{\beta,\eta}} \exp\left(\beta_N \sum_{k=0}^{N^2-1} \eta(N X_{N^{-2}k}^N) \right) \\
&= \frac{1}{Z_N^{\beta,\eta}} \exp\left(\beta_N \sum_{k=0}^{N^2-1} N^{-d/2} \eta^N(X_{N^{-2}k}^N) \right) \\
&\simeq \frac{1}{Z_N^{\beta,\eta}} \exp\left(\beta_N N^{2-d/2} \int_0^1 \eta^N(X_t^N)\mathrm{d}t \right).
\end{aligned}$$

So if $\beta_N = N^{d/2-2}$, then formally $\mathbb{Q}_{N^2}^{\beta_N,\eta}$ converges to the measure $\mathbb{Q}$ on $\Omega = C([0,1], \mathbb{R}^d)$ which is given by

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{1}{Z} \exp\left(\int_0^1 \xi(X_t)\mathrm{d}t \right), \tag{2}$$

where $\mathbb{P}$ is the Wiener measure on Ω and by a slight abuse of notation $X_t(\omega) = \omega(t)$ is again the coordinate process. But on the other side the factor

$$\beta_N = N^{d/2-2}$$

only converges to zero if $d<4$, so by our formal discussion above we cannot expect this convergence to hold if $d\geqslant 4$. And even for $d=1,2,3$ it is not at all obvious how to interpret the integral $\int_0^1 \xi(X_s)\mathrm{d}s$ if ξ is a white noise. After all the white noise is only a distribution. For example in $d=1$ it is the distributional derivative of the Brownian motion and since the Brownian motion has non-differentiable trajectories this means that ξ is not a function.

Our aim for the rest of the lecture is to understand how to construct the measure $\mathbb{Q}$ that we formally wrote down in (2). It turns out that depending on the dimension d the problem is more or less difficult:

- For $d=1$ the random variable $\int_0^1 \xi(X_t)\mathrm{d}t$ is almost surely well defined and $\mathbb{Q}$ is equivalent to the Wiener measure.
- For $d=2,3$ the random variable $\int_0^1 \xi(X_t)\mathrm{d}t$ is almost surely not well defined, but $\mathbb{Q}$ can still be constructed and is singular to the Wiener measure. Both cases $d=2,3$ are conceptually similar but $d=2$ is technically simpler.
- For $d\geqslant 4$ neither the random variable $\int_0^1 \xi(X_t)\mathrm{d}t$ nor the measure $\mathbb{Q}$ are well defined.

We will therefore focus on the case $d=2,3$. Once we understand this special case we will also understand how to deal with $d=1$ and why $d\geqslant 4$ is outside of our scope.

1.2 Brownian motion in a white noise potential

Let now $\mathbb{P}$ be the Wiener measure on $\Omega=C([0,1],\mathbb{R}^d)$ with $d\in\{1,2,3\}$, and let ξ be a white noise on $\mathbb{R}^d$. More precisely, let $(\xi(\varphi))_{\varphi\in C_c^\infty(\mathbb{R}^d)}$ be the centered Gaussian process with covariance

$$\mathbb{E}[\xi(\varphi)\xi(\psi)]=\langle\varphi,\psi\rangle_{L^2},\qquad \varphi,\psi\in C_c^\infty(\mathbb{R}^d).$$

We would like to construct the measure

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}=\frac{1}{Z}\exp\left(\int_0^1 \xi(X_t)\mathrm{d}t\right),$$

where $X_s(\omega)=\omega(s)$ is the coordinate process on Ω. For the moment we argue formally and pretend that ξ is a smooth function on $\mathbb{R}^d$, and we would like to understand the dynamics of X under $\mathbb{Q}$. For this purpose solve the backward Cauchy problem $h\colon[0,1]\times\mathbb{R}^d\to\mathbb{R}$,

$$\left(\partial_t+\frac{1}{2}\Delta\right)h(t,x)=-\frac{1}{2}|\nabla h(t,x)|^2-\xi(x),\qquad h(1,x)=0.$$

Then by Itô's formula we have $\mathbb{P}$-a.s.

$$\begin{aligned}0=h(1,X_1)&=h(0,0)+\int_0^1\left(\partial_t+\frac{1}{2}\Delta\right)h(t,X_t)\mathrm{d}t+\int_0^1\nabla h(t,X_t)\mathrm{d}X_t\\&=h(0,0)+\int_0^1\left(-\frac{1}{2}|\nabla h(t,X_t)|^2-\xi(X_t)\right)\mathrm{d}t+\int_0^1\nabla h(t,X_t)\mathrm{d}X_t,\end{aligned}$$

which is equivalent to

$$\frac{1}{Z}\exp\left(\int_0^1 \xi(X_t)\mathrm{d}t\right) = \frac{e^{h(0,0)}}{Z}\exp\left(\int_0^1 \nabla h(t,X_t)\mathrm{d}X_t - \frac{1}{2}\int_0^1 |\nabla h(t,X_t)|^2\mathrm{d}t\right).$$

The prefactor on the right hand side is deterministic, while for well behaved h the stochastic exponential has expectation 1 under $\mathbb{P}$. This means that the prefactor drops out and

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \exp\left(\int_0^1 \nabla h(t,X_t)\mathrm{d}X_t - \frac{1}{2}\int_0^1 |\nabla h(t,X_t)|^2\mathrm{d}t\right).$$

But now the Radon-Nikodym derivative is a stochastic exponential, so by Girsanov's theorem the coordinate process X solves the SDE

$$X_t = \int_0^t \nabla h(s,X_s)\mathrm{d}s + B_t \tag{3}$$

under $\mathbb{Q}$, where B is a $\mathbb{Q}$-Brownian motion. So under the assumption that (3) has a unique solution it is equivalent to construct $\mathbb{Q}$ or to solve (3) and in the following we concentrate on the solution theory for (3). This looks like a harmless SDE, but actually h is only a differentiable function if $d=1$ and in particular ∇h is only a distribution and not a function in dimensions $d=2,3$. So we have to understand SDEs with distributional drift. This will require tools from functional analysis and paracontrolled distributions and will be the main topic of these lectures.

1.3 Singular martingale problem

We now take a generic SDE $X\colon [0,1]\to\mathbb{R}^d$ with distributional drift $V\colon[0,1]\times\mathbb{R}^d\to\mathbb{R}^d$,

$$X_t = x + \int_0^t V(s,X_s)\mathrm{d}s + B_t \tag{4}$$

as our starting point, and try to develop a solution theory for such equations. Delarue and Diel realized that if $d=1$ then in certain situations rough path integration can be used to make sense of (4). This was extended to $d>1$ by Cannizzaro and Chouk who replaced rough paths by paracontrolled distributions. Since our main aim is to learn paracontrolled distributions, we follw Cannizzaro and Chouk.

The idea is elegant and in retrospective quite simple. If V is a nice function, then one way of describing the law of X is through the martingale problem: The law of X is the only probability measure on $\Omega = C([0,1],\mathbb{R}^d)$ such that for all $\varphi\in C_c^\infty([0,1]\times\mathbb{R}^d)$ the process

$$\varphi(t,X_t) - \varphi(0,x) - \int_0^t (\partial_s + \mathcal{G}_s)\varphi(X_s)\mathrm{d}s$$

is a martingale, where

$$\mathcal{G}_s\varphi(x) = \frac{1}{2}\Delta\varphi(x) + V(s,x)\nabla\varphi(x).$$

If $V(s)$ is not a nice function but only a distribution, then $\mathcal{G}_s\varphi$ is well defined (we can multiply the distribution $V(s)$ with the smooth function $\nabla\varphi$), but $\mathcal{G}_s\varphi$ is also just a distribution and not a function because multiplying a distribution with a smooth function leads to an object that is also just a distribution (think of multiplication with 1, the smoothest function there is). Now we have again the problem that the integral

$$\int_0^t (\partial_s + \mathcal{G}_s)\varphi(X_s)\mathrm{d}s$$

does not make any sense, so we cannot even write down the martingale problem. The idea is therefore to find a domain of <u>non-smooth</u> functions $\varphi: [0,1] \times \mathbb{R}^d \to \mathbb{R}$ such that

$$(\partial_s + \mathcal{G}_s)\varphi(s) \in C_b(\mathbb{R}^d)$$

for all $s \in [0,1]$. Indeed the multiplication with a non-smooth function may increase the regularity, think of multiplying a very irregular $f > 0$ with $1/f$. If we also have some continuity in s, then for such φ the integral $\int_0^t (\partial_s + \mathcal{G}_s)\varphi(s, X_s)\mathrm{d}s$ is well defined and we can formulate the martingale problem: $\mathbb{Q}$ solves the martingale problem associated to (4) if for all φ in our domain the process

$$\varphi(t, X_t) - \varphi(0, x) - \int_0^t (\partial_s + \mathcal{G}_s)\varphi(s, X_s)\mathrm{d}s$$

is a martingale. The next question is then how to obtain the existence and uniqueness of solutions to this martingale problem, but we will get back later to this point. For now let us just note that if $\varphi(s)$ is a non-smooth function, then it is not so obvious how to make sense of $V(s)\nabla\varphi$. So as a first step we need a way of multiplying distributions.

2 Products of Distributions

2.1 Distributions

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

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

where

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

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

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

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

Example 2. Clearly $L^p = L^p(\mathbb{R}^d) \subset \mathscr{S}'$ for all $p \in [1, \infty]$ if we identify $u \in L^p$ with the map $\varphi \mapsto \int_{\mathbb{R}^d} u(x)\varphi(x)\mathrm{d}x$, and more generally the space of finite signed measures on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ is contained in $\mathscr{S}'$. Another example of a tempered distribution is $\varphi \mapsto \partial^\mu \varphi(x)$ for $\mu \in \mathbb{N}_0^d$ and $x \in \mathbb{R}^d$. A continuous function u is in $\mathscr{S}'$ if and only if it has at most polynomial growth at infinity.

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

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

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

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

Example 3.

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

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

iii. For the Fourier transform

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

we have ${}^tA = A$.

iv. For the inverse Fourier transform

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

we have ${}^tA = A$.

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

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

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

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

Example 4. Let $u \in \mathscr{S}'$ and $\varphi, \psi \in \mathscr{S}$. The following relations will be used all the time in what follows:

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

- Parseval's identity:

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

  and by extension $u(\varphi^*) = \hat{u}(\hat{\varphi}^*)$;

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

Note that we can only define the product $u\varphi$ by duality if $\varphi \in C^\infty$ with partial derivatives of polynomial growth. If φ is a non-smooth function or even a distribution, then we need other arguments.

Example 5. In $d = 1$ we can turn $\frac{1}{x}$ into a tempered distribution via the so called *principal value*. The details of that construction are not important for us, but with it we obtain for the Dirac delta δ (i.e. $\delta(\varphi) = \varphi(0)$)

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

This example shows that a general extension of the product $u\varphi$ to distributions or non-smooth functions φ is not possible. Our way of overcoming this difficulty is to restrict both u and φ to suitable subspaces of $\mathscr{S}'$. The simplest solution is to require u and φ to have compatible regularity. For that purpose we need to introduce regularities on distribution spaces.

2.2 Besov spaces

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

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

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

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

Then $\Delta_j u$ is the projection of u onto its frequencies of order $\sim 2^j$. Since frequencies of order 2^j correspond to spatial scales of order 2^{-j}, the sum $\sum_{i\leqslant j}\Delta_i u$ provides a description of u up to the spatial scale 2^{-j}. For a smooth function u this should already give a very accurate picture of u, and therefore we expect $\Delta_j u$ to rapidly decay as $j\to\infty$. Measuring the strength of that decay will provide us with a notion of regularity. But there are two problems with the above formal decomposition: First of all it is not even well defined, because we are multiplying $\hat{u}\in\mathscr{S}'$ with non-smooth indicator functions. And even if we could make sense of this product, for example by setting

$$\mathscr{F}^{-1}(\mathbb{I}_{[2^j,2^{j+1})}(|\cdot|)\hat{u}) := \mathscr{F}^{-1}(\mathbb{I}_{[2^j,2^{j+1})}(|\cdot|)) * u,$$

then it still turns out that the operation $u\mapsto\Delta_j u$ is quite badly behaved. For example, we would like to estimate $\|\Delta_j u\|_{L^p}\leqslant\|\mathscr{F}^{-1}(\mathbb{I}_{[2^j,2^{j+1})}(|\cdot|))\|_{L^1}\|u\|_{L^p}$ via Young's inequality, but the L^1 norm on the right hand side is infinite because while $\mathscr{F}^{-1}(\mathbb{I}_{[2^j,2^{j+1})}(|\cdot|))\in C^\infty$, it is not in L^1.

Definition 6. *For $u\in\mathscr{S}'$ and $j\geqslant -1$ we define the <u>Littlewood-Paley blocks</u> of u as*

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

where $(\rho_j)\subset C_c^\infty$ is a <u>smooth partition of unity</u> with

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

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

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

The kernels $K_j, K_{<j}, K_{\leqslant j}$ are all bounded in L^1, uniformly in j.

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

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

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

Definition 7. *For $\alpha\in\mathbb{R}$ and $p,q\in[1,\infty]$ the Besov space $B^\alpha_{p,q}$ is defined as*

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

So the index p describes the integrability and the index α the "regularity" (i.e. the decay of the blocks). The index q provides some fine-tuning and is not very important. $B^{\alpha}_{p,q}$ is always a Banach space, for all α, p, q. We will mostly work with $p=q=\infty$ for which we introduce a special notation:

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

Exercise 1. Let δ denote the Dirac delta, $\delta(\varphi) = \varphi(0)$. Show that $\delta_0 \in \mathscr{C}^{-d}$.

If $\alpha \in (0,\infty) \setminus \mathbb{N}$, then $\mathscr{C}^{\alpha}$ is the space of $\lfloor \alpha \rfloor$ times differentiable functions whose partial derivatives of order $\lfloor \alpha \rfloor$ are $(\alpha - \lfloor \alpha \rfloor)$-Hölder continuous. But for $k \in \mathbb{N}$ the space $\mathscr{C}^k$ is strictly larger than C^k_b, the space of k times continuously differentiable functions with bounded partial derivatives of all order.

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

We will often use these inequalities without explicitly mentioning it.

The following Bernstein inequality is very useful when dealing with functions with compactly supported Fourier transform. We present the proof to show that it is not difficult and because many proofs that we will omit later are based on similar arguments.

Lemma 8. *(Bernstein inequality) Let $\mathscr{B}$ be a ball, $k \in \mathbb{N}_0$, and $1 \leqslant p \leqslant q \leqslant \infty$. There exists a constant $C > 0$ that depends only on k, $\mathscr{B}$, p and q such that for all $\lambda > 0$ and $u \in L^p$ with* $\operatorname{supp}(\mathscr{F}u) \subseteq \lambda \mathscr{B}$ *we have*

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

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

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

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

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

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

It follows immediately that for $\alpha \in \mathbb{R}$, $u \in \mathscr{C}^{\alpha}$, and $\mu \in \mathbb{N}^d_0$, we have $\|\partial^{\mu} u\|_{\alpha - |\mu|} \lesssim \|u\|_{\alpha}$. Another simple application is the Besov embedding theorem, whose proof we leave as an exercise.

Lemma 9. *(Besov embedding)*

Let $1 \leqslant p_1 \leqslant p_2 \leqslant \infty$ and $1 \leqslant q_1 \leqslant q_2 \leqslant \infty$, and let $\alpha \in \mathbb{R}$. Then for all $u \in \mathscr{S}'$

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

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

Lemma 10.

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

$$u = \sum_{j \geqslant -1} u_j \in \mathscr{C}^\alpha \qquad \text{and} \qquad \|u\|_\alpha \lesssim \sup_{j \geqslant -1} \{2^{j\alpha} \|u_j\|_{L^\infty}\}.$$

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

$$u = \sum_{j \geqslant -1} u_j \in \mathscr{C}^\alpha \qquad \text{and} \qquad \|u\|_\alpha \lesssim \sup_{j \geqslant -1} \{2^{j\alpha} \|u_j\|_{L^\infty}\}.$$

Proof. If $\mathscr{F} u_j$ is supported in $2^j \mathscr{A}$, then $\Delta_i u_j \neq 0$ only if $2^i \simeq 2^j$ and therefore

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

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

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

where in the last step we used that $\alpha > 0$. □

A similar result also holds for Besov spaces $B^\alpha_{p,q}$ with general $p, q \in [1, \infty]$, but we will not need this.

2.3 The paraproduct and the resonant term

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

Lemma 11. *There exists an annulus $\mathscr{A}$ such that for all $j \geqslant 1$ and all $i \leqslant j-2$*

$$\operatorname{supp}(\mathscr{F}(\Delta_i u \Delta_j v)) \subset 2^j \mathscr{A}, \qquad u, v \in \mathscr{S}'.$$

Moreover, there exists a ball $\mathscr{B}$ such that for all $i, j \geqslant -1$ with $|i-j| \leqslant 1$

$$\operatorname{supp}(\mathscr{F}(\Delta_i u \Delta_j v)) \subset 2^j \mathscr{B}.$$

Proof. This is quite simple:

$$\begin{aligned}\operatorname{supp}(\mathscr{F}(\Delta_i u \Delta_j v)) = \operatorname{supp}(\mathscr{F}\Delta_i u * \mathscr{F}\Delta_j u) &\subset \overline{\operatorname{supp}(\mathscr{F}\Delta_i u) + \operatorname{supp}(\mathscr{F}\Delta_j u)} \\ &\subset 2^i \tilde{\mathscr{A}} + 2^j \tilde{\mathscr{A}} = 2^j (2^{i-j} \tilde{\mathscr{A}} + \tilde{\mathscr{A}})\end{aligned}$$

for another annulus $\tilde{\mathscr{A}}$. By our assumptions on the dyadic partition of unity we can choose $\tilde{\mathscr{A}}$ such that $2^{i-j}\tilde{\mathscr{A}} + \tilde{\mathscr{A}} \subset \mathscr{A}$ for a new annulus $\mathscr{A}$ and all $i \leqslant j-2$.

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

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

Example 12. Below we see a slowly oscillating function u (red curve) and a fast sine curve v (blue curve). The product uv is shown under the two curves. We see that the local fluctuations of uv are due to v, and that uv is essentially oscillating with the same speed as v.

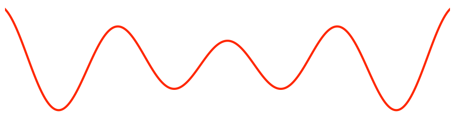

Figure 1. u oscillates slowly.

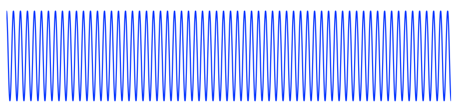

Figure 2. v is a fast sine curve.

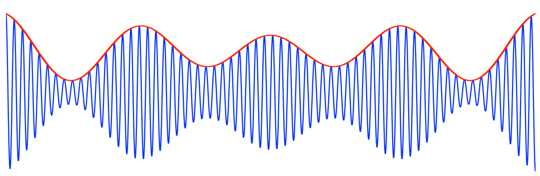

Figure 3. uv still lives on the same scale as v.

Formally we can decompose the product uv of two distributions as

$$uv = \sum_{i,j \geqslant -1} \Delta_i u \Delta_j v = u \prec v + u \succ v + u \circ v.$$

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

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

We call $u \prec v$ and $u \succ v$ *paraproducts*, and $u \circ v$ the *resonant term*.

Bony's crucial observation is that $u \prec v$ (and thus $u \succ v$) is always a well-defined distribution. Heuristically, $u \prec v$ behaves at large frequencies (i.e. small spatial scales) like v and thus retains the same regularity, and u provides only a frequency modulation of v. This can also be seen in Example 12 above, where the product uv is actually equal to the paraproduct $u \prec v$ because u has no rapidly oscillating components. The only difficulty in constructing uv for arbitrary distributions lies in handling the diagonal term $u \circ v$. The following key estimates provide the analytically precise formulation of the preceding heuristic discussion:

Theorem 13. *(Paraproduct estimates) For any $\beta \in \mathbb{R}$ and $u, v \in \mathscr{S}'$ we have*

$$\|u \prec v\|_\beta \lesssim \|u\|_{L^\infty} \|v\|_\beta, \tag{5}$$

and for $\alpha < 0$ furthermore

$$\|u \prec v\|_{\alpha+\beta} \lesssim \|u\|_\alpha \|v\|_\beta. \tag{6}$$

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

$$\|u \circ v\|_{\alpha+\beta} \lesssim \|u\|_\alpha \|v\|_\beta. \tag{7}$$

Proof. By Lemma 11 there exists an annulus $\mathscr{A}$ such that $\operatorname{supp}(\mathscr{F}(\Delta_{\leqslant j-2} u \Delta_j v)) \subset 2^j \mathscr{A}$, and for $u \in L^\infty$ we have

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

Inequality (5) now follows from Lemma 10. The proof of (6) and (7) works in the same way, except that for estimating $u \circ v$ we need $\alpha + \beta > 0$ because now the terms of the series are supported in balls and not in annuli. □

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

In combination with Exercise 2 above we deduce the following simple corollary:

Corollary 14. *Let $u \in \mathscr{C}^\alpha$ and $v \in \mathscr{C}^\beta$ with $\alpha + \beta > 0$. Then the product $(u, v) \mapsto uv$ of smooth functions can be extended to a bounded bilinear operator from $\mathscr{C}^\alpha \times \mathscr{C}^\beta$ to $\mathscr{C}^{\alpha \wedge \beta}$. While $u \prec v$, $u \succ v$, and $u \circ v$ depend on our specific dyadic partition of unity, the product uv does not.*

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

Example 15. Let α, $\beta \in \mathbb{R}$ and consider the functions $u_n(x) = n^{-\alpha} e^{inx}$ on $\mathbb{R}$, and $v_n(x) = n^{-\beta} e^{-inx}$. It is easy to see that $\|u_n\|_{\tilde{\alpha}} \to 0$ and $\|v_n\|_{\tilde{\beta}} \to 0$ for all $\tilde{\alpha} < \alpha$ and $\tilde{\beta} < \beta$. Nonetheless

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

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

3 Not-so-singular diffusions

3.1 Domain of the generator

Equipped with these tools, we now take a new shot at constructing the diffusion X: $[0, T] \to \mathbb{R}^d$

$$\mathrm{d}X_t = V(t, X_t)\mathrm{d}t + \mathrm{d}B_t, \qquad X_0 = x,$$

with a distributional drift $V \in C([0, T], \mathscr{C}^{-\beta}(\mathbb{R}^d, \mathbb{R}^d))$ for some $\beta > 0$. That is, $V(t)$ is a vector of d distributions in $\mathscr{C}^{-\beta}(\mathbb{R}^d)$. As discussed above, the idea is to understand the domain of the infinitesimal generator

$$\mathcal{G}_t u = \frac{1}{2}\Delta u + V(t)\nabla u.$$

For that purpose we have to find functions u with $\mathcal{G}_t u \in C_b(\mathbb{R}^d)$. The easiest way of guaranteeing this is to prescribe the right hand side f in $\mathcal{G}_t u = f$ and to solve for u. More precisely, we will study the time-dependent problem u: $[0, T] \times \mathbb{R}^d \to \mathbb{R}$

$$(\partial_t + \mathcal{G}_t) u = f, \qquad u(T) = \varphi,$$

for given $f \in C_b([0, T] \times \mathbb{R}^d)$ and $\varphi \in \mathscr{C}^\alpha(\mathbb{R}^d)$ with an $\alpha > 0$ to be determined. For simplicity we will slightly abuse notation and write for all $m \in \mathbb{N}$ and $\gamma \in \mathbb{R}$

$$C_T \mathscr{C}^\gamma = C([0, T], \mathscr{C}^\gamma(\mathbb{R}^d, \mathbb{R}^m)), \qquad \|v\|_{C_T \mathscr{C}^\gamma} = \sup_{t \in [0, T]} \|v(t)\|_\gamma.$$

Then the paraproduct estimates allow us to control $V\nabla u$ whenever $u \in C_T \mathscr{C}^\alpha$ with $\alpha > 1 + \beta$. Indeed, then Bernstein's inequality gives $\nabla u \in C_T \mathscr{C}^{\alpha-1}$ and therefore

$$\|V\nabla u\|_{C_T \mathscr{C}^{-\beta}} \lesssim \|V\|_{C_T \mathscr{C}^{-\beta}} \|\nabla u\|_{C_T \mathscr{C}^{\alpha-1}} \lesssim \|V\|_{C_T \mathscr{C}^{-\beta}} \|u\|_{C_T \mathscr{C}^\alpha}.$$

So we would like to set up a Picard iteration for u in the Banach space $C_T \mathscr{C}^\alpha$. Since $V\nabla u$ has only regularity $-\beta$, we need a mechanism to increase the regularity. This is provided by the Laplacian: We have

$$\left(\partial_t + \frac{1}{2}\Delta\right) u = -V\nabla u + f,$$

and the right hand side is in $C_T\mathscr{C}^{-\beta}$. Since the right hand side is a second order derivative of u, we might guess that $u\in C_T\mathscr{C}^{2-\beta}$. This is indeed justified by the Schauder (heat kernel) estimates for the semigroup generated by the Laplacian. The following statement is correct but not totally precise because we do not explain in which sense nor in which space we solve the equation.

Lemma 16. *Let $\alpha\in\mathbb{R}$ and let $(P_t)_{t\geqslant 0}$ be the semigroup generated by $\Delta/2$,*

$$P_t u = \mathscr{F}^{-1}(e^{-t2|\pi z|^2}\hat{u}) = \mathscr{F}^{-1}(e^{-t2|\pi z|^2}) * u = p_t * u,$$

where p_t is the Gaussian density with mean 0 and covariance $t\times \mathrm{Id}$. Given $f\in C_T\mathscr{C}^{\alpha-2}$ and $\varphi\in\mathscr{C}^\alpha$, the unique weak solution u to

$$\left(\partial_t+\frac{1}{2}\Delta\right)u = f, \qquad u(T)=\varphi,$$

is

$$u(t) = P_{T-t}\varphi - \int_t^T P_{s-t}f(s)\mathrm{d}s, \tag{8}$$

and we have

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

Proof. (Sketch of proof):

Let us write $\hat{\cdot}$ for the <u>space</u> Fourier transform. Then our equation is equivalent to

$$\partial_t\hat{u}(t,z) = -\frac{1}{2}\widehat{\Delta u}(t,z)+\hat{f}(t,z) = \frac{1}{2}|2\pi z|^2\hat{u}(t,z)+\hat{f}(t,z)$$

with terminal condition $\hat{u}(T,z)=\hat{\varphi}(z)$. This is an ordinary differential equation for $\hat{u}(z)$ that admits the explicit solution

$$\hat{u}(t,z) = e^{-(T-t)2|\pi z|^2}\hat{\varphi}(z) - \int_t^T e^{-(s-t)2|\pi z|^2}\hat{f}(s,z)\mathrm{d}s,$$

which is nothing but (8). To derive the estimate for $\|u\|_{C_T\mathscr{C}^\alpha}$ note that $\Delta_j\psi$ only contains frequencies of the order 2^j and therefore formally for all $\gamma\geqslant 0$ and $\beta\in\mathbb{R}$

$$\begin{aligned}\|\Delta_j P_t\psi\|_{L^\infty} &= \|P_t\Delta_j\psi\|_{L^\infty} \simeq e^{-ct2^{2j}}\|\Delta_j\psi\|_{L^\infty}\\ &= t^{-\gamma/2}2^{-j\gamma}\Big(\big(\sqrt{t}2^j\big)^\gamma e^{-c(\sqrt{t}2^j)^2}\Big)\|\Delta_j\psi\|_{L^\infty}\\ &\lesssim t^{-\gamma/2}2^{-j(\gamma+\beta)}\|\psi\|_\beta.\end{aligned}$$

This can be made rigorous by using similar arguments as in the proof of Bernstein's inequality. Thus, we have

$$\|P_t\psi\|_{\beta+\gamma}\lesssim t^{-\gamma/2}\|\psi\|_\beta$$

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

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

Also,

$$\left\| \int_t^T P_{s-t} f(s) \mathrm{d}s \right\|_\alpha \leqslant \int_t^T \|P_{s-t} f(s)\|_\alpha \mathrm{d}s \lesssim \int_t^T |s-t|^{-1} \|f\|_{C_T\mathscr{C}^{\alpha-2}} \mathrm{d}s.$$

Of course, the right hand side is infinite because $|s-t|^{-1}$ barely fails to be integrable. But we already see that if we wanted to regularize f by $2-\varepsilon$ degrees of regularity, then we would get an estimate for u. To gain two full derivatives we would have to be slightly more careful and use two different estimates, one for s close to t, and one for s close to T. □

We argued above that for $u \in C_T\mathscr{C}^\alpha$ with $\alpha > 1 + \beta$ the product $V\nabla u$ is well defined and in $C_T\mathscr{C}^{-\beta}$. In that case the Schauder estimates give us for the solution u to

$$\left(\partial_t + \frac{1}{2}\Delta\right) v = -V\nabla u + f, \qquad v(T) = \varphi,$$

the estimate

$$\|v\|_{C_T\mathscr{C}^{2-\beta}} \lesssim \|V\nabla u\|_{C_T\mathscr{C}^{-\beta}} + \|f\|_{C_T\mathscr{C}^{-\beta}} + \|\varphi\|_{2-\beta}.$$

So if $2 - \beta = \alpha$, then we can set up a Picard iteration in the Banach space $C_T\mathscr{C}^\alpha$ and hope to find a unique solution u to the equation. For that to work we need that $2 - \beta > 1 + \beta$, i.e. $\beta < 1/2$.

Proposition 17. *Let $\beta < 1/2$ and $V \in C_T\mathscr{C}^{-\beta}$ and set $\alpha = 2 - \beta$. Then for all $f \in C_T\mathscr{C}^{-\beta}$ and $\varphi \in \mathscr{C}^\alpha$ there exists a unique solution $u \in C_T\mathscr{C}^\alpha$ to the equation*

$$\left(\partial_t + \frac{1}{2}\Delta\right) u = -V\nabla u + f, \qquad u(T) = \varphi. \tag{9}$$

Moreover, u depends continuously on $(V, f, \varphi) \in C_T\mathscr{C}^{-\beta} \times C_T\mathscr{C}^{-\beta} \times \mathscr{C}^\alpha$.

Proof. (Sketch of proof):

We already know that

$$\Gamma(u) = P_{T-t}\varphi - \int_t^T P_{s-t}(-V\nabla u(s) + f(s)) \mathrm{d}s$$

is continuous from $C_T\mathscr{C}^\alpha$ to itself. One can also show that for $\alpha' < \alpha$ which still satisfies $\alpha' > 1 + \beta$ and for $t < T$ sufficiently close to T the map Γ is a contraction on $C([t, T], \mathscr{C}^{\alpha'})$. This gives us a unique solution u on $[t, T]$ and with regularity $\mathscr{C}^{\alpha'}$. A posteriori it is then easy to see that u is actually in $C([t, T], \mathscr{C}^\alpha)$ (but actually we don't care about this very much, $\mathscr{C}^{\alpha'}$ is good enough for our purposes). Because our equation is linear the length of the interval $\lambda = T - t$ can be chosen in a way that only depends on V but not on the terminal condition φ, and therefore we can repeat the Picard iteration on $[t - \lambda, t]$, etc., to construct the unique solution $u \in C_T\mathscr{C}^\alpha$.

The solution u depends continuously on the data because all the operations in the equation are continuous. □

3.2 Application to the Brownian motion in a white noise potential

We now know how to solve the generator equation

$$(\partial_t + \mathcal{G}_t) u = f, \qquad u(T) = \varphi,$$

for $\mathcal{G}_t = \frac{1}{2}\Delta + V(t)\nabla$ with $V, f \in C_T \mathscr{C}^{-\beta}$ and $\varphi \in \mathscr{C}^{2-\beta}$ whenever $\beta > 1/2$. Recall that we want to construct the probability measure $\mathbb{Q}$ on $C([0,T], \mathbb{R}^d)$ that is formally given by

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{1}{Z}\exp\left(\int_0^T \xi(X_t)\mathrm{d}t\right),$$

where $\mathbb{P}$ is the Wiener measure and ξ is a space white noise on $\mathbb{R}^d$, and that this is equivalent to solving the SDE

$$\mathrm{d}X_t = \nabla h(t, X_t)\mathrm{d}t + \mathrm{d}B_t$$

on $[0,T]$, where h solves

$$\left(\partial_t + \frac{1}{2}\Delta\right) h(t,x) = -\frac{1}{2}|\nabla h(t,x)|^2 - \xi(x), \qquad h(T,x) = 0.$$

So as a first step we should solve this equation which is of a similar form as (9) except that it is slightly more complicated because it is nonlinear. To see what regularity to expect of the solution h, let us first derive the regularity of the white noise. It turns out that the white noise itself does not lie in any Besov space at all because it is in a sense "unbounded" at infinity. This is a similar phenomenon as for the Brownian motion on $\mathbb{R}$, which is locally Hölder continuous of order $1/2 - \varepsilon$ but

$$\sup_n \frac{|B_{n+1} - B_n|}{|(n+1) - n|^\alpha} = \sup_n \frac{|B_{n+1} - B_n|}{1^\alpha} = \infty$$

for all $\alpha \in \mathbb{R}$ because the variables $(B_{n+1} - B_n)_{n \in \mathbb{N}}$ are independent standard Gaussians. This difficulty can be overcome by considering weighted Besov spaces, but to keep the presentation as simple as possible we will not do this but instead simply restrict ξ to a compact set:

Lemma 18. *Let ξ be a space white noise on $\mathbb{R}^d$. By the convergence result in Lemma 1 (which can be easily lifted from convergence in $\mathcal{D}'$ to convergence in $\mathscr{S}'$) we can assume that ξ almost surely takes values in $\mathscr{S}'$. Moreover, for all compactly supported bounded functions ψ, for all $\gamma < -d/2$, and $p \in [1, \infty)$ we have*

$$\mathbb{E}[\|\xi\psi\|_\gamma^p] < \infty.$$

Proof. Because we only know $\psi \in L^{\infty}$ it is not obvious that $\xi\psi$ is still an element of $\mathscr{S}'$. One way to see this would be to approximate $\xi\psi$ by $\eta^N\psi$ as in Lemma 1. Given $\xi\psi$ with values in $\mathscr{S}'$ we have for all $p\in[1,\infty)$ and all $\beta\in\mathbb{R}$

$$\mathbb{E}\Big[\|\xi\psi\|^p_{B^{\beta}_{p,p}}\Big] = \sum_{j\geqslant -1} 2^{j\beta p}\int_{\mathbb{R}^d}\mathbb{E}[|\Delta_j(\xi\psi)(x)|^p]\mathrm{d}x = \sum_{j\geqslant -1} 2^{j\beta p}\int_{\mathbb{R}^d}\mathbb{E}[|\xi(K_j(x-\cdot)\psi)|^p]\mathrm{d}x.$$

The random variable $\xi(K_j(x-\cdot)\psi)$ is Gaussian, and therefore its p-th moment is up to a constant simply equal to its second moment raised to the $p/2$, i.e.

$$\begin{aligned}\mathbb{E}[|\xi(K_j(x-\cdot)\psi)|^p] \simeq \mathbb{E}[|\xi(K_j(x-\cdot)\psi)|^2]^{p/2} &= \left(\int_{\mathbb{R}^d}|K_j(x-y)\psi(y)|^2\mathrm{d}y\right)^{p/2}\\ &= \left(\int_{\mathbb{R}^d}|2^{jd}K_0(2^jy)\psi(x-y)|^2\mathrm{d}y\right)^{p/2},\end{aligned}$$

where we used that the dyadic partition of unity K_j can be chosen such that $K_j(x) = 2^{jd}K_0(2^jx)$. The integral on the right hand side is only over $x-B$, where B is the compact set where ψ is supported. Using the faster-than-polynomial decay of K_0 we get the bound

$$\left(\int_{\mathbb{R}^d}|2^{jd}K_0(2^jy)\psi(x-y)|^2\mathrm{d}y\right)^{p/2} \lesssim \left(\int_{x-B}|2^{jd}K_0(2^jy)|^2\mathrm{d}y\right)^{p/2} \lesssim 2^{jdp/2}(1+|x|)^{-d-1},$$

which leads to

$$\sum_{j\geqslant -1} 2^{j\beta p}\int_{\mathbb{R}^d}\mathbb{E}[|\xi(K_j(x-\cdot)\psi)|^p]\mathrm{d}x \lesssim \sum_{j\geqslant -1} 2^{j\beta p}\int_{\mathbb{R}^d}2^{jdp/2}(1+|x|)^{-d-1}\mathrm{d}x \lesssim \sum_{j\geqslant -1} 2^{jp(\beta+d/2)}.$$

For $\beta<-d/2$ the right hand side is finite and we obtain $\mathbb{E}\Big[\|\xi\psi\|^p_{B^{\beta}_{p,p}}\Big]<\infty$. Now it suffices to apply the Besov embedding theorem: Given $\gamma<-d/2$ let $p\in[1,\infty)$ be such that $\gamma+d/p<-d/2$. Then

$$\mathbb{E}[\|\xi\psi\|^p_{\gamma}] \lesssim \mathbb{E}\Big[\|\xi\psi\|^p_{B^{\gamma+d/p}_{p,p}}\Big]<\infty.$$

□

Remark 19.

i. If we would not have localized by multiplying with ψ we would have only got the estimate

$$\int_{\mathbb{R}^d}\mathbb{E}[|\xi(K_j(x-\cdot))|^p]\mathrm{d}x \simeq \int_{\mathbb{R}^d}(\|K_j(x-\cdot)\|^2_{L^2})^{p/2}\mathrm{d}x = \int_{\mathbb{R}^d}(2^{jd}\|K\|^2_{L^2})^{p/2}\mathrm{d}x = \infty$$

simply because we are integrating over an infinite volume.

ii. A key tool in the proof was that for Gaussian random variables all moments are compatible, and therefore we were able to obtain an estimate for the p-th moment while only explicitly computing the second moment.

Actually it is more convenient to take a periodic version of ξ rather than to truncate it outside of a compact set. More precisely, let $\psi = \mathbb{I}_{[-M,M]^d}$, so $\psi\xi \in \mathscr{C}^{-d/2-}$ (we use the notation $\mathscr{C}^{-d/2-} = \bigcup_{\varepsilon>0} \mathscr{C}^{-d/2-\varepsilon}$), and set

$$\tilde{\xi}(\varphi) = \sum_{k\in\mathbb{Z}} \psi\xi(\varphi(\cdot - k2M)).$$

One can check that we still have $\tilde{\xi} \in \mathscr{C}^{-d/2-}$, and from now on we work with this $\tilde{\xi}$. To lighten the notation we still write ξ though.

Now let us get back to our equation

$$\left(\partial_t + \frac{1}{2}\Delta\right)h = -\frac{1}{2}|\nabla h|^2 - \xi.$$

We just saw that at best the right hand side is in $C_T\mathscr{C}^{-d/2-}$. Then by the Schauder estimates we expect $h \in C_T\mathscr{C}^{2-d/2-}$ and thus $\nabla h \in C_T\mathscr{C}^{1-d/2-}$. But this means that ∇h has negative regularity as soon as $d > 2$ and therefore the product $|\nabla h|^2$ is not well defined! But we simply ignore this problem and <u>assume</u> that we are given a solution h with the natural regularity $h \in C_T\mathscr{C}^{2-d/2-}$. We will discuss later how to solve the equation for h based on similar arguments that we used in the generator equation.

If $h \in C_T\mathscr{C}^{2-d/2-}$ is given, then $\nabla h \in C_T\mathscr{C}^{1-d/2-}$ and $d/2 - 1 < 1/2$ exactly if $d < 3$. So for $d = 1, 2$ we can set $\beta = d/2 - 1 - \varepsilon$ for some small $\varepsilon > 0$ and $\alpha = 2 - \beta$, and then the arguments from Section 3.1 allow us to solve

$$\left(\partial_t + \frac{1}{2}\Delta\right)u = -\nabla h \nabla u + f, \qquad u(T) = \varphi,$$

for all $f \in C_T\mathscr{C}^{-\beta}$ and $\varphi \in \mathscr{C}^{\alpha}$.

Theorem 20. *Let $d \in \{1, 2\}$, $\Omega = C([0,T], \mathbb{R}^d)$ and let X be the coordinate process on Ω. For any $x \in \mathbb{R}^d$ there exists a unique probability measure $\mathbb{Q}_x$ on Ω with $\mathbb{Q}(X_0 = x) = 1$ and such that for all $f \in C_b([0,T] \times \mathbb{R}^d)$ and $\varphi \in \mathscr{C}^{\alpha}$ the process*

$$u(t, X_t) - u(0, x) - \int_0^t f(s, X_s)\mathrm{d}s$$

is a martingale, where u is the solution to

$$\left(\partial_t + \frac{1}{2}\Delta\right)u = -\nabla h \nabla u + f, \qquad u(T) = \varphi.$$

Moreover, the coordinate process is Markovian under $\mathbb{Q}_x$ and if $\rho \in \mathscr{S}$ with $\int_{\mathbb{R}^d}\rho(x)\mathrm{d}x = 1$ and $\xi^n = \rho^n * \xi$ for $\rho^n(x) = n^d\rho(nx)$, then:

i. $\mathbb{Q}_x$ is the weak limit of

$$\frac{\mathrm{d}\mathbb{Q}_x^n}{\mathrm{d}\mathbb{P}_x} = \frac{1}{Z_x^n}\exp\left(\int_0^T \xi^n(X_s)\mathrm{d}s\right),$$

where $\mathbb{P}_x$ is the Wiener measure with $X_0 = x$;

ii. X is of the form

$$X_t = x + A_t + B_t, \tag{10}$$

where A is $\mathbb{Q}_x$-almost surely γ-Hölder continuous for all $\gamma < 1$ and satisfies

$$A_t = \lim_{n \to \infty} \int_0^t (\rho^n * \nabla h)(s, X_s) \mathrm{d}s.$$

Proof. (Sketch of proof):

The existence is shown by proving the relative compactness (tightness) of the measures $(\mathbb{Q}_x^n)_{n \in \mathbb{N}}$ on Ω. The initial condition is fixed as $X_0 = x$ under all the $(\mathbb{Q}_x^n)$, so by the Kolmogorov-Chentsov criterion it suffices to prove a bound of the form

$$\mathbb{E}_{\mathbb{Q}_x^n}[|X_t - X_s|^p] \lesssim |t - s|^{p/2}$$

uniformly in n, for some $p > 2$. For that purpose we fix $k \in \{1, ..., d\}$ and consider the solution $u^{n,k}$ to

$$\left(\partial_t + \frac{1}{2}\Delta\right) u^{n,k} = -\nabla h^n \nabla u^{n,k} + \nabla h^{n,k}, \qquad u^{n,k}(T) = 0,$$

where $\nabla h^{n,k}$ is the k-th coordinate of ∇h^n. With this $u^{n,k}$ we apply Itô's formula under $\mathbb{Q}_x^n$:

$$\begin{aligned} u^{n,k}(t, X_t) - u^{n,k}(s, X_s) &= \int_s^t \left(\partial_r + \frac{1}{2}\Delta + \nabla h^n \nabla\right) u^{n,k}(r, X_r) \mathrm{d}r + \int_s^t \nabla u^{n,k}(r, X_r) \mathrm{d}B_r \\ &= \int_s^t \nabla h^{n,k}(r, X_r) \mathrm{d}r + \int_s^t \nabla u^{n,k}(r, X_r) \mathrm{d}B_r \\ &= X_t^k - X_s^k - (B_t^k - B_s^k) + \int_s^t \nabla u^{n,k}(r, X_r) \mathrm{d}B_r. \end{aligned}$$

If we rearrange this equation to have $X_t^k - X_s^k$ on one side and all the other terms on the other side, then we can use the regularity of u^n (which can be controlled uniformly in n) to prove the desired uniform bound for $\mathbb{E}_{\mathbb{Q}_x^n}[|X_t^k - X_s^k|^p]$. Any limit point $\mathbb{Q}_x$ then satisfies the characterization i., and also it is not difficult to show that it solves the martingale problem. Moreover, actually we can show that

$$\mathbb{E}_{\mathbb{Q}_x^n}[|(X_t^k - X_s^k) - (B_t^k - B_s^k)|^p] \lesssim |t - s|^p,$$

whenever $p \geqslant 1$, which proves the decomposition (10) and the regularity of A under $\mathbb{Q}_x$.

As for the uniqueness, let $\mathbb{Q}_x$ be a solution to the martingale problem, let $f \in C_b([0, T] \times \mathbb{R}^d)$ and consider the solution u to

$$\left(\partial_t + \frac{1}{2}\Delta\right) u = -\nabla h \nabla u + f, \qquad u(T) = 0.$$

Then

$$0 = u(0, x) + \int_0^T f(s, X_s) \mathrm{d}s + M_T,$$

and taking the expectation we see that $\mathbb{E}_{\mathbb{Q}_x}\big[\int_0^T f(s, X_s)\mathrm{d}s\big]$ is uniquely determined by u. This shows that the finite-dimensional marginal distributions of X under $\mathbb{Q}_x$ are uniquely determined, and therefore $\mathbb{Q}_x$ is unique. Moreover, if a martingale problem gives rise to unique marginal distributions, then the coordinate process is Markovian under the solution to the martingale problem.

It remains to show that

$$A_t = \lim_{n\to\infty} \int_0^t \nabla h^n(s, X_s)\mathrm{d}s.$$

Since A_t^k is the limit of $u^{m,k}(t, X_t) - u^{m,k}(s, X_s) - \int_s^t \nabla u^{m,k}(r, X_r)\mathrm{d}B_r$, this will follow if we can show that uniformly in all large m

$$\mathbb{E}_{\mathbb{Q}_x^m}\left[\left|u^{m,k}(t, X_t) - u^{m,k}(s, X_s) - \int_s^t \nabla u^{m,k}(r, X_r)\mathrm{d}B_r - \int_0^t (\rho^n * \nabla h^k)(s, X_s)\mathrm{d}s\right|^2\right] \leqslant c(n)$$

with $c(n)\to 0$ for $n\to\infty$. So consider the equation

$$\left(\partial_t + \frac{1}{2}\Delta\right)u_n^{m,k} = -\nabla h^m \nabla u_n^{m,k} + \rho^n * \nabla h^k, \qquad u_n^{m,k}(T) = 0, \tag{11}$$

and note that

$$\int_0^t (\rho^n * \nabla h)(s, X_s)\mathrm{d}s = u_n^{m,k}(t, X_t) - u_n^{m,k}(s, X_s) - \int_s^t \nabla u_n^{m,k}(r, X_r)\mathrm{d}B_r.$$

Now the required bound follows from the continuous dependence of $u_n^{m,k}$ on the data: $u^{m,k} = u_m^{m,k}$ and the functions $u_m^{m,k}$ and $u_n^{m,k}$ are close to each other because of the continuous dependence of (11) on the data. □

Remark 21. Here we follow Cannizzaro and Chouk, but essentially the same result was previously shown by Flandoli, Issoglio, Russo. The difference is that Cannizzaro and Chouk are ultimately able to pass the regularity barrier $-1/2$ for the distributional drift.

Assuming that we can solve the equation for h, the above arguments allow us at least to construct the measure

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{1}{Z}\exp\left(\int_0^T \xi(X_s)\mathrm{d}s\right)$$

for the periodized version ξ of the white noise. By making the period $2M$ of ξ larger and larger it seems plausible that we can also take ξ as the white noise on all of $\mathbb{R}^d$, but we will not bother with this here. What is more interesting is that in dimension $d = 2$ the solution h^n to

$$\left(\partial_t + \frac{1}{2}\Delta\right)h^n = -\frac{1}{2}|\nabla h^n|^2 - \xi^n$$

actually does <u>not</u> converge for $n\to\infty$. Only after adding a large constant $c_n\to\infty$ and modifying the equation to

$$\Big(\partial_t+\frac{1}{2}\Delta\Big)h^n=-\frac{1}{2}(|\nabla h^n|^2-c_n)-\xi^n,\qquad h^n(T)=0,$$

we are able to pass to the limit. Picking up the argument from the introduction, we get $\mathbb{P}_x$-a.s.

$$\begin{aligned}0=&h^n(0,x)+\int_0^T\Big(\partial_t+\frac{1}{2}\Delta\Big)h^n(t,X_t)\mathrm{d}t+\int_0^T\nabla h^n(t,X_t)\mathrm{d}X_t\\=&h^n(0,x)+\int_0^T\Big(-\frac{1}{2}(|\nabla h^n(t,X_t)|^2-c_n)-\xi^n(X_t)\Big)\mathrm{d}t+\int_0^T\nabla h^n(t,X_t)\mathrm{d}X_t,\end{aligned}$$

which is equivalent to

$$\frac{1}{Z_x^n}\exp\Big(\int_0^T\xi^n(X_t)\mathrm{d}t\Big)=\frac{e^{h^n(0,x)+c^nT}}{Z_x^n}\exp\Big(\int_0^T\nabla h^n(t,X_t)\mathrm{d}X_t-\frac{1}{2}\int_0^T|\nabla h^n(t,X_t)|^2\mathrm{d}t\Big),$$

i.e. $Z_x^n=e^{h^n(0,x)+c^nT}$ diverges to $+\infty$ for $n\to\infty$. This indicates that $\mathbb{Q}_x^n$ becomes singular with respect to $\mathbb{P}_x$ in the limit $n\to\infty$. Indeed, one can show that when replacing ξ^n by $\beta\xi^n$ we should replace c_n by β^2c_n to obtain a nontrivial limit for h^n. But this means that for $p\in(0,1)$

$$\mathbb{E}_{\mathbb{P}_x}\Big[\Big(\frac{1}{Z_x^n}\exp\Big(\int_0^T\xi^n(X_t)\mathrm{d}t\Big)\Big)^p\Big]=\frac{\mathbb{E}_{\mathbb{P}_x}\big[\exp\big(\int_0^Tp\xi^n(X_t)\mathrm{d}t\big)\big]}{\exp(ph^n(0,x)+pc^nT)}=\frac{\exp(h^{n,p}(0,x)+p^2c_nT)}{\exp(ph^n(0,x)+pc_nT)}.$$

The first terms in the two exponentials converge to a finite limit as $n\to\infty$, but

$$\lim_{n\to\infty}\exp(c_nT(p^2-p))=0,$$

which proves that the Radon-Nikodym derivative $\mathrm{d}\mathbb{Q}_x^n/\mathrm{d}\mathbb{P}_x$ converges to zero in L^p and thus in probability. This easily implies that the limiting measure $\mathbb{Q}_x$ is singular with respect to $\mathbb{P}_x$.

4 Quite-singular diffusions

So far we are (modulo periodization of the white noise ξ) able to construct the measure $\mathbb{Q}_x$ formally given by

$$\frac{\mathrm{d}\mathbb{Q}_x}{\mathrm{d}\mathbb{P}_x}=\frac{1}{Z_x}\exp\Big(\int_0^T\xi(X_t)\mathrm{d}t\Big)$$

in $d=1,2$. The case $d=3$ is not contained in our arguments above because then the solution h to

$$\Big(\partial_t+\frac{1}{2}\Delta\Big)h=-\Big(\frac{1}{2}|\nabla h|^2-\infty\Big)-\xi,\qquad h(T)=0,$$

(it is also necessary to renormalize the equation in $d=3$) satisfies only $\nabla h\in C_T\mathscr{C}^{-1/2-}$. This means that we expect $u\in C_T\mathscr{C}^{1/2-}$ for the solution u to

$$\left(\partial_t+\frac{1}{2}\Delta\right)u=-\nabla h\nabla u+f,\qquad u(T)=\varphi,$$

and therefore the product $\nabla h\nabla u$ is ill-defined. To deal with this case we will need paracontrolled distributions.

4.1 Paracontrolled distributions

To simplify notation we write again V instead of ∇h, and we want to solve the PDE

$$\mathscr{L}u=-(\nabla u)V+f,\qquad u(T)=\varphi,$$

for $\mathscr{L}=\partial_t+\frac{1}{2}\Delta$ and $V\in C_T\mathscr{C}^{-1/2-}$.

First, let us try to see what is the worst regularity of V that we could possibly hope to treat, because this will allow us to understand the concept of *local subcriticality*. If $V\in C_T\mathscr{C}^{-\beta}$ and we are somehow able to make sense of the right hand side of the equation, then it is at best in $C_T\mathscr{C}^{-\beta}$ as well. Therefore, we expect $u\in C_T\mathscr{C}^{2-\beta}$ and then $\nabla u\in C_T\mathscr{C}^{1-\beta}$. Now if we naively apply the estimates for paraproduct and resonant term without making sure that the resonant term is even defined, then we get the following heuristic for the regularity of the product $(\nabla u)V$:

- If ∇u has positive regularity, then $(\nabla u)V\in C_T\mathscr{C}^{-\beta}$.

- But if ∇u has negative regularity, i.e. if $1-\beta<0$, then $(\nabla u)V\in C_T\mathscr{C}^{1-\beta-\beta}=C_T\mathscr{C}^{1-2\beta}$.

We see that something happens around $\beta=1$, and for concreteness let us assume that $\beta=1+\varepsilon$. Then our original hope to have regularity $2-\beta=1-\varepsilon$ for u was too optimistic, and we should at best expect regularity $-1-2\varepsilon$ for the right hand side of the equation, so $1-2\varepsilon$ for u. But then we should have $(\nabla u)V\in C_T\mathscr{C}^{-1-3\varepsilon}$, which leads us to update our guess for the regularity of u to $u\in C_T\mathscr{C}^{1-3\varepsilon}$. So we are caught in an infinite loop where we update our guesses for the regularity of u to worse and worse spaces, ultimately spiralling down to "$C_T\mathscr{C}^{-\infty}$". We say that the equation is locally supercritical for $\beta>1$, it is critical for $\beta=1$, and it is locally subcritical for $\beta<1$. Regularity structures, paracontrolled distributions, and all the other approaches only allow us to handle locally subcritical equations, so we definitely have to assume $\beta<1$. In the following we treat the case $\beta=1/2+$ which corresponds to $V=\nabla h$ in $d=3$. Our analysis can be easily seen to extend all the way to $\beta<2/3$. But to not get confused with regularity indices so much, the case $\beta=1/2+$ is more convenient.

For pedagogical reasons it is also more convenient to write the factor V on the right hand side in the following argumentation. The product $(\nabla u)V$ can be decomposed via the paraproduct and the resonant term as a sum of three terms

$$(\nabla u)V=(\nabla u)\prec V+(\nabla u)\succ V+(\nabla u)\circ V.$$

The first two terms are well defined without a problem and only the last term is difficult to make sense of. Let us be bold for the moment and assume that we have a way to make sense of $(\nabla u) \circ V$ and that this term has its natural regularity. Given $V \in C_T \mathscr{C}^{-1/2-}$ we expect $u \in C_T \mathscr{C}^{3/2-}$ and $\nabla u \in \mathscr{C}_T \mathscr{C}^{1/2-}$, and therefore we expect the following regularities for the paraproducts and the resonant term:

$$\underbrace{(\nabla u) \prec V}_{-1/2-} + \underbrace{(\nabla u) \succ V}_{0-} + \underbrace{(\nabla u) \circ V}_{0-}, \tag{12}$$

but of course the last term on the right hand side is not well defined. Nonetheless we see that the most irregular term in the decomposition of the product is the paraproduct $\nabla u \prec V$, the other terms should be more regular. Now we follow a hunch and guess that maybe also the solution u itself is given by a paraproduct plus a more regular term. More precisely we make the following *paracontrolled Ansatz:*

$$u = u' \prec H + u^\sharp$$

with $u' \in C_T \mathscr{C}^{1/2-}$, $H \in C_T \mathscr{C}^{3/2-}$, and $u^\sharp \in C_T \mathscr{C}^{2-}$ of better regularity. We hope to gain half a derivative, as much as in (12), but the regularity for u' admittedly falls from the sky at this point. Why would we hope to have a representation of u as a paraproduct plus smoother remainder? Well, we saw in Example 12 that the paraproduct $u' \prec H$ is a “frequency modulation” of g and looks very much like H on small scales (and thus $\nabla(u' \prec H)$ looks like ∇H on small scales). But the difficulty we have with defining $(\nabla u) \circ V$ is exactly coming from small scale contributions of ∇u and V which in the product create diverging resonances on large scales. So if we understand how the small scale contributions of ∇H interact with those of V and that no diverging resonances develop, then by the philosophy of controlled rough paths we might also hope that ∇u has no diverging resonances with V. More precisely, if u matches the paracontrolled Ansatz then we can make the decomposition of the product more precise:

$$(\nabla u) V = \underbrace{(\nabla u) \prec V}_{-1/2-} + \underbrace{(\nabla u) \succ V}_{0-} + \underbrace{(\nabla u^\sharp) \circ V}_{1/2-} + \underbrace{((\nabla u') \prec H) \circ V}_{1/2-} + \underbrace{(u' \prec \nabla H) \circ V}_{!!},$$

where with the notation “!!” we single out the only term which is still not well defined. To cure this term we will need the following commutator lemma, which is the main result in paracontrolled distributions:

Lemma 22. *Assume that $\alpha \in (0,1)$ and $\beta, \gamma \in \mathbb{R}$ are such that $\alpha + \beta + \gamma > 0$ and $\beta + \gamma < 0$. Then the trilinear operator on $\mathscr{S}^3$, defined by*

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

satisfies

$$\|C(f, g, h)\|_{\alpha + \beta + \gamma} \lesssim \|f\|_\alpha \|g\|_\beta \|h\|_\gamma, \tag{13}$$

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

Remark 23. It would be more aesthetically pleasing to have the same estimate for

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

but sadly this is not true.

We do not give the proof here (it can be found in Paraproduct paper), but equipped with the tools that we learned in this lecture it is not difficult at all.

Let us see what we can do with this estimate: Recall that the only term in the product $(\nabla u) \circ V$ that is still ill defined is $(u' \prec \nabla H) \circ V$, which we rewrite as

$$(u' \prec \nabla H) \circ V = \underbrace{C(u', \nabla H, V)}_{1/2-} + u'(\nabla H \circ V).$$

The last product on the right hand side is still not well defined because $\nabla H \circ V$ is not well defined. However, if we assume that $\nabla H \circ V$ is extrinsically given and has its natural regularity $\nabla H \circ V \in C_T \mathscr{C}^{0-}$, then $u'(\nabla X \circ V)$ is well defined and in $C_T \mathscr{C}^{0-}$. To summarize we then have the following decomposition of the product:

$$(\nabla u) V = \underbrace{(\nabla u) \prec V}_{-1/2-} + \underbrace{(\nabla u) \succ V}_{0-} + \underbrace{(\nabla u^\sharp) \circ V}_{1/2-} + \underbrace{((\nabla u') \prec H) \circ V}_{1/2-} + \underbrace{C(u', \nabla H, V)}_{1/2-} + \underbrace{u'(\nabla H \circ V)}_{0-}.$$

So far H and u' were completely arbitrary and the only ingredient we needed was sufficient regularity and the a priori knowledge of the product $\nabla H \circ V$. But for this to be useful we have to be able to set up a Picard iteration in the space of functions that oblige the paracontrolled Ansatz. That is, we have to show that the paracontrolled Ansatz is stable under the map $u \mapsto v$, where u solves

$$\mathscr{L} v = (\nabla u) V + f, \qquad v(T) = \varphi,$$

say for $\varphi \in \mathscr{C}^{2-}$ and $f \in C_b([0, T] \times \mathbb{R}^d)$. To see the stability of the paracontrolled Ansatz, we guess again a paracontrolled Ansatz for v, say

$$v = v' \prec H + v^\sharp,$$

with the same H as before and the same regularity requirements for v' and $v^\sharp$. Then we have

$$\mathscr{L} v^\sharp = \mathscr{L} v - \mathscr{L}(v' \prec H) = (\nabla u) \prec V + \mathcal{R} - \mathscr{L}(v' \prec H),$$

where $\mathcal{R} \in C_T \mathscr{C}^{0-}$ is a more regular remainder term. Now morally we have

$$\mathscr{L}(v' \prec H) - v' \prec \mathscr{L} H \in C_T \mathscr{C}^{0-}, \tag{14}$$

because for example by Leibniz's rule

$$\Delta(v' \prec H) - v' \prec \Delta H = \underbrace{\Delta v' \prec H}_{0-} + \underbrace{2 \nabla v' \prec \nabla H}_{0-}.$$

The problem is that $\mathscr{L}$ also contains the time derivative, and of course we cannot simply differentiate elements of $C_T\mathscr{C}^\alpha$ in time. This is a small technical problem which does not change the overall picture and which can be overcome by also requesting some time regularity of v' and H, and also we need to slightly adapt the definition of the paraproduct and introduce some smoothing in the time variable as well. For simplicity we brush this under the carpet and pretend that (14) holds. Then we get

$$\mathscr{L}v^\sharp = (\nabla u) \prec V - v' \prec \mathscr{L}H + \tilde{\mathcal{R}},$$

with a new $\tilde{\mathcal{R}} \in C_T\mathscr{C}^{0-}$. If we choose $v' = \nabla u$ and $\mathscr{L}H = V$, then the two irregular terms on the right hand side and we end up with a well defined equation for $v^\sharp$, from where we also learn that our guess $v^\sharp \in C_T\mathscr{C}^{2-}$ was justified because $\tilde{\mathcal{R}} \in C_T\mathscr{C}^{0-}$ and then by the Schauder estimate Lemma 16 $v^\sharp$ has the right regularity as long as $v^\sharp(T) \in \mathscr{C}^{2-}$. But

$$v^\sharp(T) = v(T) - ((\nabla u) \prec H)(T) = v(T)$$

if we assume that $H(T) = 0$. Therefore, it suffices to take $\varphi \in \mathscr{C}^{2-}$. From here the following result is shown by the same arguments as in Proposition 17:

Proposition 24. *Let $V \in C_T\mathscr{C}^{-1/2-}$, set $\mathscr{L}H = V$, $H(T) = 0$, and assume $H \circ V \in C_T\mathscr{C}^{0-}$ is an extrinsically given distribution of the required regularity. Then for all $f \in C_b([0,T] \times \mathbb{R}^d)$ and all $\varphi \in \mathscr{C}^{2-}$ there exists a unique paracontrolled solution $u = \nabla u \prec H + u^\sharp$ to*

$$\mathscr{L}u = (\nabla u)V + f, \qquad u(T) = \varphi,$$

where the product on the right hand side is interpreted as

$$(\nabla u)V = (\nabla u) \prec V + (\nabla u) \succ V + u^\sharp \circ V + C(\nabla u, H, V) + \nabla u(H \circ V).$$

Remark 25.

i. In the proof of Theorem 20 we needed to also take $f = V^k$ in the generator equation, for $k \in \{1, \ldots, d\}$. Now $V^k \in C_T\mathscr{C}^{-1/2-}$ and we cannot just add an arbitrary element of $C_T\mathscr{C}^{-1/2-}$. But for $f = V^k$ we can simply replace $(\nabla u)V$ by

$$(\nabla u + e_k)V,$$

where $(e_k)_{k=1,\ldots,d}$ are the k-th canonical basis vectors in $\mathbb{R}^d$.

ii. One can show that if $H \circ V = \lim_{n\to\infty} (\rho^n * H) \circ (\rho^n * V)$, then also

$$(\nabla u)V = \lim_{n\to\infty} (\rho^n * \nabla u)\rho^n * V.$$

The difference to the case $\beta < 1/2$ is that here this convergence may hold for one mollifier ρ but fail for another $\tilde{\rho}$ (or give another limit).

iii. It is not hard to see that if $V \in C_T\mathscr{C}^{-\beta}$ and $H \circ V \in C_T\mathscr{C}^{1-2\beta}$, then for $\beta < 2/3$ everything works as above. After we cross the threshold $\beta = 2/3$ we would have to go higher in the expansion of our solution, and it is not so obvious how to do this in the setting of paracontrolled distributions (see however Bailleul-Bernicot).

From here we can argue essentially as in the not-so-singular case before to obtain a version of Theorem 20 in the three-dimensional case. Only that now different approximations ξ^n to ξ can give rise to different limits h of the h^n, or for some approximations the convergence of h^n can even fail. So we have to assume that ξ^n converges to ξ such that h^n converges to h and then $H^n \circ \nabla h^n$ converges to $H \circ \nabla h$. In this case $\mathrm{d}\mathbb{Q}_x^n = (Z_x^n)^{-1}\exp\big(\int_0^T \xi^n(X_s)\big)\mathrm{d}\mathbb{P}_x^n$ converges to $\mathbb{Q}_x$ which is uniquely determined as the solution to the martingale problem. The point ii. in Theorem 20 is also much more subtle and to have the representation $A_t = \lim_{n\to\infty}\int_0^t \rho_n * \nabla h(s, X_s)\mathrm{d}s$ we need to assume that $H \circ \rho^n * \nabla h$ converges to $H \circ \nabla h$ in $C_T\mathscr{C}^{0-}$. Moreover, we only get $A \in C^{3/4-}([0,T], \mathbb{R}^d)$ almost surely.

4.2 A KPZ-like equation

To conclude our analysis it only remains to see that the equation for h actually has a meaning. In $d = 1$ this is quite easy because then ∇h is actually a function and not a distribution, so $|\nabla h|^2$ poses no problem.

In $d = 2$ we have the equation

$$\mathscr{L}h = -\frac{1}{2}|\nabla h|^2 - \xi, \qquad h(T) = 0.$$

We write $h = X + h_R$, where X solves $\mathscr{L}X = -\xi$, $X(T) = 0$, and get

$$\mathscr{L}h_R = -\frac{1}{2}|\nabla h|^2 = -\frac{1}{2}(|\nabla h_R|^2 + 2\nabla h_R \nabla X + |\nabla X|^2).$$

Now we saw that $\xi \in \mathscr{C}^{-1-}$, and therefore $X \in C_T\mathscr{C}^{1-}$ and $|\nabla X|^2$ is not well defined. But if it was defined, it should have regularity $|\nabla X|^2 \in C_T\mathscr{C}^{0-}$, as well as $\nabla X \in C_T\mathscr{C}^{0-}$. Then the worst term on the right hand side of the equation for h_R is in $C_T\mathscr{C}^{0-}$, so we guess $h_R \in C_T\mathscr{C}^{2-}$. Under this condition $|\nabla h_R|^2$ and $2\nabla h_R \nabla X$ are well defined, and it is no problem for us to solve the equation (at least on a small time interval $[t, T]$ because now the equation is quadratic and therefore global existence is not so trivial any more; but this can be dealt with and we simply ignore that problem here). So the only ingredient we need to make sense of the equation for h_R is the input $|\nabla X|^2$. It turns out that this term cannot be constructed, because when squaring the distribution ∇X we do indeed pick up diverging resonances in the resonant product. However, these divergences are of a particularly simple form and it turns out that for any given approximation of the identity (ρ^n) there exist constants $c_n \to \infty$ that depend on (ρ^n), but such that the limit

$$\lim_{n\to\infty} |\nabla X^n|^2 - c_n$$

exists in $C_T\mathscr{C}^{0-}$ and does not on the specific approximation procedure; here $\mathscr{L}X^n = \rho^n * \xi$, $X^n(T) = 0$. Moreover, $c_n = \lim_{n\to\infty}\mathbb{E}[|\nabla X^n(t,x)|^2]$ for all $t < T$ and $x \in \mathbb{R}^2$ which justifies that in Section 3.2 we replaced c_n by $\beta^2 c_n$ when we replaced ξ by $\beta\xi$. Then h is not the limit of the solutions to the naive equation $\mathscr{L}h^n = -\frac{1}{2}|\nabla h^n|^2 - \rho^n * \xi$, but instead we have to consider

$$\mathscr{L}h^n = -\frac{1}{2}(|\nabla h^n|^2 - c_n) - \rho^n * \xi$$

In $d = 3$ things are much more complicated and in fact exactly as complicated as for the usual KPZ equation in $d = 1$ where the space white noise is replaced by a space-time white noise. Here we will see the trees appear: Let again

$$\mathscr{L}h = -\frac{1}{2}|\nabla h|^2 - \xi, \qquad h(T) = 0$$

and $\mathscr{L}X = -\xi$, $X(T) = 0$. Now we write again $h = X + h_{\geqslant 1}$ and get

$$\mathscr{L}h_{\geqslant 1} = -\frac{1}{2}(|\nabla h_{\geqslant 1}|^2 + 2\nabla h_{\geqslant 1}\nabla X + |\nabla X|^2).$$

This time $X \in C_T\mathscr{C}^{1/2-}$ and therefore we expect $|\nabla X|^2 \in C_T\mathscr{C}^{-1-}$. This is still way too irregular, so we also take out this term: Let

$$\mathscr{L}X^{\vee} = -\frac{1}{2}|\nabla X|^2, \qquad X^{\vee}(T) = 0,$$

for which we expect $X^{\vee} \in C_T\mathscr{C}^{1-}$, and write $h = X + X^{\vee} + h_{\geqslant 2}$ with

$$\begin{aligned} \mathscr{L}h_{\geqslant 2} &= -\frac{1}{2}(|\nabla h_{\geqslant 1}|^2 + 2\nabla h_{\geqslant 1}\nabla X + |\nabla X|^2) \\ &= -\frac{1}{2}\big(|\nabla h_{\geqslant 2} + X^{\vee}|^2 + 2\nabla\big(h_{\geqslant 2} + X^{\vee}\big)\nabla X\big) \\ &= -\frac{1}{2}\big(|\nabla h_{\geqslant 2}|^2 + 2\nabla h_{\geqslant 2}\nabla X^{\vee} + |\nabla X^{\vee}|^2 + 2\nabla h_{\geqslant 2}\nabla X + 2\nabla X^{\vee}\nabla X\big). \end{aligned}$$

Now the most irregular term is expected to be $2\nabla X^{\vee}\nabla X \in C_T\mathscr{C}^{-1/2-}$, so we take it out as well by setting

$$\mathscr{L}X^{\vee\!\!\!\vee} = -\frac{1}{2}\nabla X^{\vee}\nabla X, \qquad X^{\vee\!\!\!\vee}(T) = 0,$$

and then $h_{\geqslant 3} = X + X^{\vee} + 2X^{\vee\!\!\!\vee}$. At this point the use of the tree notation is hopefully quite transparant: given two rooted binary trees τ_1 and τ_2 we write $(\tau_1\tau_2)$ for the tree that is obtained by joining the roots of τ_1 and τ_2 in a new root, and we set recursively

$$\mathscr{L}X^{(\tau_1\tau_2)} = -\frac{1}{2}\nabla X^{\tau_1}\nabla X^{\tau_2}$$

with $X^{\bullet} = X$. The tree notation is simply a very convenient tool to index all the objects that have to be recursively constructed from the noise ξ in order to solve the equation.

Now we could continue in the expansion of h and subtract for example $4X^{\vee\!\!\!\vee\!\!\!\vee}$ or $X^{\overset{\vee\vee}{\vee}}$. But it turns out that $h_{\geqslant 3} \in C_T\mathscr{C}^{3/2-}$ and the regularity cannot be improved by a further expansion, because in the equation for $h_{\geqslant n}$ there would always appear the term

$$\nabla h_{\geqslant n}\nabla X,$$

which at best can have the same regularity as X, so it should be in $C_T\mathscr{C}^{-1/2-}$. But then the best regularity we can hope for is $h_{\geqslant n} \in C_T\mathscr{C}^{3/2-}$ and then the product $\nabla h_{\geqslant n}\nabla X$ is ill-defined. So the idea is to abandon the expansion and rather make a paracontrolled ansatz for the remainder. More precisely, we assume

$$h = X + X^{\vee} + 2X^{\vee\!\!\!\vee} + h' \prec Q + h^{\sharp}$$

with $h' \in C_T \mathscr{C}^{1/2-}$, $h^\sharp \in C_T \mathscr{C}^{2-}$, and

$$\mathscr{L} Q = \nabla X, \qquad Q(T) = 0,$$

so $Q \in C_T \mathscr{C}^{3/2-}$. Then we can use the same arguments as in the generator equation to solve our KPZ type equation based on the paracontrolled Ansatz.

The next step is then to construct the tree data $(X, X^{\vee}, X^{\vee\!\!\vee}, X^{\vee\!\vee\!\vee}, X^{\vee\vee}, \nabla Q \circ X)$ from the white noise. The higher order trees do not appear in the expansion of h, but they are needed to make sense of the right hand side of the equation. The construction of the trees is already quite demanding and carried out in Cannizzaro-Chouk. Not surprisingly it also requires some renormalizations. How to do the renormalization and the stochastic estimates systematically in the (here equivalent) setting of regularity structures was explained in the lectures by Lorenzo Zambotti and Ajay Chandra.