Math 279 Lecture Notes
Topics in Stochastic Partial Differential Equations

Professor: Fraydoun Rezakhanlou
Scribe: Daniel Raban

Contents

1 A Motivating Example for Studying Stochastic PDEs

1.1 Fluids: an example of what a stochastic PDE looks like

Here is an analogy. When you see a line of ants, you may think that the line is relatively straight, so you write down an equation that describes the motion. If you increase the precision of your model, you may see that the ants actually move with some random fluctuations, so you add some randomness to your model. The more precision you require, the more you realize that the ants are not moving in a straight line at all and are instead constantly bumping into each other, exchanging information. This is how stochastic PDEs are.

Imagine that we have a fluid for which the velocity of fluid particles are known, say $u(x,t)$. As a simple model for the fluid particle, we write

$$\frac{dx}{dt} = u(x,t).$$

This is an ODE which, as a first approximation, gives us a good idea of a model for what is happening. To take into account the thermal fluctuation of the fluid, we may write

$$\frac{dx}{dt} = \underbrace{u(x,t)}_{\text{vector field}} + \underbrace{\sigma(x,t)}_{\text{matrix}} \eta(t), \tag{1}$$

with η representing "white noise" (to be formally defined later) and $\sigma(x,t)$ measuring the strength of the fluctuation at (x,t). Here, η is a Gaussian process with $\mathbb{E}[\eta(t)] = 0$ and $\mathbb{E}[\eta(t)\eta(s)] = \delta_0(t-s)$, where δ_0 is the Dirac delta "function" at 0.

In reality, u itself solves some PDE, and in the case of a (viscous) incompressible fluid, we have

$$u_t + (u \cdot \nabla)u + \underbrace{\nabla P}_{\text{pressure}} = \nu \Delta u + \underbrace{f}_{\text{force}}$$
$$\nabla \cdot u = 0,$$

where for simplicity we assume $\sigma = \sqrt{2\nu}I$. We have a system of 4 equations with 3 unknowns (the function u), so we need to solve this equation for the pair (u,P). A natural model example for f is that $f = f(x,t)$ is "white noise" in (x,t) (sometimes, we assume f is white in t and "colored" in x).

1.2 Regularity issues with white noise

Going back to the previous equation (1), how can we make sense of this equation? The problem is that "white noise" cannot be realized as a function. A solution to (1) is an example of a diffusion.[1] Observe that if $u = 0$ and $\sigma = 1$, then $\frac{dx}{dt} = \eta$. As it turns

[1] Diffusions were first described by Kolmogorov in the early 30s and later described by Paul Lévy and Itô.

out, $x(t) = x(0) + B(t)$, where B is a standard Brownian motion.[2] It is well-known that Brownian motion can be realized as a continuous function, in fact $B \in \mathcal{C}^{1/2-}$. Here, we write $\mathcal{C}^\alpha$ as the space of Hölder continuous functions of exponent α and $\mathcal{C}^{\alpha-} = \bigcap_{\beta<\alpha} \mathcal{C}^\beta$. In fact, $\eta = \dot{B} \in \mathcal{C}^{-1/2-}$. By $f \in \mathcal{C}^\beta$ for $\beta < 0$, we mean $f = \dot{g}$ with $g \in \mathcal{C}^{\beta+1}$ (we will give a more robust definition of this later).

Going back to $\dot{x} = u(x,t) + \sigma(x,t)\eta(t)$, we expect this to have a solution $x(\cdot) \in \mathcal{C}^{1/2-}$. To make sense of this, we write

$$x(t) = x(0) + \int_0^t u(x(s), s)\, ds + \int_0^t \sigma(x(s), s)\, \underbrace{\eta(s)\, ds}_{dB(s)}.$$

We face the following difficulty:

$$\eta(\varphi) = \int \eta(s)\varphi(s)\, ds = \int \dot{B}(s)\varphi(s)\, ds \overset{\text{IBP}}{=} -\int B(s)\dot{\varphi}(s)\, ds,$$

where φ is smooth with compact support. The problem is that f is not $\mathcal{C}^1$, only $\mathcal{C}^{1/2-}$. This calls for studying $\int_0^t g\, df$ with f, g continuous functions. This problem has a rich history that we now review:

1. In fact, Riemann and Steiltjes defined the integral $\int_0^t g\, df$ as

$$\int_0^t g\, df = \lim_{n\to\infty} \sum_{i=0}^{2^n} g(s_i)(f(t_{i+1}) - f(t_i)) \tag{2}$$

   with $s_i \in [t_i, t_{i+1}]$, where the t_i form a mesh with 2^n points. It turns out that this equation converges (no matter what we choose for s_i) if g is continuous ($g \in \mathcal{C}^0$) and $f \in \mathrm{BV}$ is of bounded variation. Recall that $f \in \mathrm{BV}$ means $\|f\|_{\mathrm{BV}} < \infty$, where $\|f\|_{\mathrm{BV}} = \sup_{0<t_1<\cdots<t_k<t} \sum_{i=1}^{k-1} |f(t_{i+1}) - f(t_i)|$. In particular, if $g \in \mathcal{C}^0$ and $f \in \mathcal{C}^1$, then $\int_0^t f\, dg$ can be defined.

2. Lebesgue theory allows us to interpret $\int_0^t f\, dg$ as $\int_0^t f\, d\mu$, where $\mu = g'$ in a weak sense:

$$\int \varphi\, d\mu = g'(\varphi) = -\int \varphi' g\, dt$$

   for all smooth φ. In this picture, $f \in \mathrm{BV} \iff f'$ can be realized as a measure.

3. So far, we know how to define $\int g\, df$ with $g \in \mathcal{C}^0, f \in \mathrm{BV}$. But we can also make sense of it if $g \in \mathrm{BV}, f \in \mathcal{C}^0$ by declaring $\int_0^t g\, df = g(t)f(t) - g(0)f(0) - \int_0^t f\, dg$.

[2] The moral here is to still differentiate things that are not differentiable. Don't let that stop you.

4. Young observed that equation (2) still works if $f \in \mathcal{C}^\alpha, g \in \mathcal{C}^\beta$ with $\alpha + \beta > 1$. In fact, (2) works even when $f \in \mathrm{BV}_{1/\alpha}, g \in \mathrm{BV}_{1/\beta}$, where

$$\|f\|_{\mathrm{BV}_p} = \sup_{0 \le t_1 < \cdots < t_h \le t} \sum_i |f(t_{i+1}) - f(t_i)|^p$$

for $p \ge 1$. Observe that $\mathrm{BV}_{1/\alpha} \supsetneq \mathcal{C}^\alpha$. Moreover, Young proved that $h(t) = \int_0^t g\, df$ satisfies the following bound:

$$|h(t) - h(s) - g(s)(f(t) - f(s))| \le c|t-s|^{\alpha+\beta} \tag{3}$$

where c is a constant depending on $\|f\|_{\mathcal{C}^\alpha}$ and $\|g\|_{\mathcal{C}^\alpha}$. In fact, h can be uniquely specified as the only function for which $h(0) = 0$, and h satisfies (for some constant c) (3). If $h, \widetilde{h}$ are two solutions, then $k = h - \widetilde{h}$ satisfies $|k(t) - k(s)| \le c|t-s|^{\alpha+\beta}$.

1.3 Ways of defining the stochastic integral with irregular functions

Going back to our integral $\int_0^t \sigma(x(s), s)\, dB(s)$, Young's theory does not apply because both $\sigma(x(s), s)$ and $B(s)$ are both in $\mathcal{C}^{1/2-}$. As an example, consider $\int_0^t F(B(s))\, dB(s)$ for $F \in \mathcal{C}^1$. In fact, the approximation in (2) may fail in two ways. Either the limit does not exist or the limit exists but depends on the choice of s_i! Some popular choices of limits in probability theory are:

Example 1.1. Itô defined the integral

$$M(t) = \int_0^t F(B(s))\, dB(s) = \lim_{n\to\infty} \sum_{i=0}^{2^n - 1} F(B(t_i))(B(t_{i+1}) - B(t_i)).$$

The advantage is that the outcome $M(t)$ is a martingale.

Here is another choice:

Example 1.2. Statonovich defined the approximation by replacing $F(B(t_i))$ with

$$\frac{F(B(t_i)) + F(B(t_{i+1}))}{2}.$$

There is also a "backward" way, where we choose $F(B(t_{i+1}))$ instead. Next time, we will discuss the drawbacks of Itô calculus and introduce rough path theory.

2 Stochastic Integration With Irregular Functions

2.1 Integration of rough deterministic functions

Our ultimate goal is to study stochastic PDEs, but before that, we need to study certain developments in studying stochastic ODEs from the 90s. For now, we are reviewing the stochastic differential equation

$$\frac{dx}{dt} = u(x,t) + \sigma(x,t)\xi(t),$$

where $\xi(t)$ is "white noise." As we discussed last time, we may make sense of this equation if we have a good candidate for

$$\int_0^t f(s)\, dg(s)$$

if f and g are as bad as Brownian motion. That is, we need to be able to deal with $f, g \in \mathcal{C}^\alpha$ for $\alpha < 1/2$. Last time, we learned that $h(t) = \int_0^t f\, dg = \lim_{n\to\infty} \sum_{j=0}^{2^n-1} f(s_j)(g(t_{j+1}) - g(t_j))$ with $s_j \in [t_j, t_{j+1}]$ and $t_j = t \cdot 2^{-n}$, provided that $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$ with $\alpha + \beta > 1$. Alternatively, we can state the following result of Young:

Theorem 2.1. *Given $f \in \mathcal{C}^\alpha, g \in \mathcal{C}^\beta$ with $\alpha + \beta > 1$, there exists a unique $h \in \mathcal{C}^\beta$ such that $h(0) = 0$ and*

$$|h(t) - h(s) - f(s)(g(t) - g(s))| \le [f]_\alpha [g]_\beta |t-s|^{\alpha+\beta}.$$

The idea is that we can approximate g by smooth functions to compute the integral, and if we keep doing this with better approximations, we will get the same answer, regardless of our choice of approximation.

An equivalent way to think about this is if $\mathscr{A} : \mathcal{C}^1 \times \mathcal{C}^1 \to \mathcal{C}^0$ by $\mathscr{A}(f, G) = fG'$, then this $\mathscr{A}$ has a continuous extension to $\widehat{\mathscr{A}} : \mathcal{C}^\alpha \times \mathcal{C}^\gamma \to \mathcal{C}^\gamma$ with $\alpha + \gamma > 0$. Here, $\gamma = \beta - 1$. This gives us a satisfactory candidate for fg', where $f \in \mathcal{C}^\alpha$, $g \in \mathcal{C}^\beta$, $\alpha + \beta > 1$. The Radon-Nikodym theorem says that if a distribution is a measure, then we can multiply it by a function and we get another measure; this, by comparison says we can multiply a distribution (which can be worse than a measure) by a function as long as the function has enough regularity.

As we mentioned last time, Young's integral cannot be used for our equation. Imagine that we have $f \in \mathcal{C}^\alpha, g \in \mathcal{C}^\beta$, and $\alpha + \beta \le 1$ with $\alpha, \beta \in (0, 1)$. What can be said about fg'? We may attempt to make sense of it by replacing g with a smooth approximation g_ε and examine $\lim_{\varepsilon\to 0} fg'_\varepsilon$. It turns out that the limit may not exist or the limit depends on the approximation.

In this context, let us examine the following question: Given Hölder f, g, consider the set $\mathscr{H}$ of h such that $h(0) = 0$ and for some C,

$$|h(t) - h(s) - f(s)(g(t) - g(s))| \le C|t-s|^{\alpha+\beta}.$$

Observe that if $h, \widetilde{h} \in \mathscr{H}$, then $h - \widetilde{h} \in \mathcal{C}^{\alpha+\beta}$. In fact, given any $h^0 \in \mathscr{H}$,

$$\mathscr{H} = \{h^0 + k : k(0) = 0, k \in \mathcal{C}^{\alpha+\beta}\}.$$

Theorem 2.2 (Lyons-Victoire, 1999)**.** *$\mathscr{H} \neq \varnothing$ always.*

The multidimensional version of this theorem was proved by Martin Hairer in 2013 or so. In other words, if $f \in \mathcal{C}^\alpha(\mathbb{R}^d)$, $g \in \mathcal{C}^\beta(\mathbb{R}^d)$, then we have at least one candidate for "$f\nabla g$" (a function multiplied by a distribution). This is basically a distribution that near x, is "close" to $f(x)\nabla g$.

2.2 Integration of functions of Brownian motion

How does stochastic calculus fit into this framework? Let's go back to our original problem

$$\dot{x} = u(x,t) + \sigma(x,t)\xi, \qquad \xi = \dot{B}.$$

Our first attempt is to make sense of $\int_0^t F(B(s))\,dB(s)$.

It is not hard to show (using the strong law of large numbers) that

$$\lim_{n\to\infty} \sum_{j=0}^{2^{n-1}} [B(t_{j+1}) - B(t_j)]^2 = t$$

almost surely. Observe that

$$\int_0^t B\,dB \approx \begin{cases} \sum_i B(t_i)(B(t_{i+1}) - B(t_i)) & \text{Itô } (I) \\ \sum_i B(t_{i+1})(B(t_{i+1}) - B(t_i)) & \text{backward } (II) \\ \sum_i \frac{B(t_{i+1})+B(t_i)}{2}(B(t_{i+1}) - B(t_i)) & \text{Stratonovich } (III). \end{cases}$$

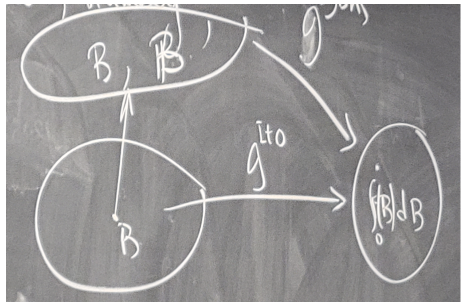
Observe that $II - I \to t$ as $n \to \infty$.

Itô's candidate was to define

$$\int_0^t F(B(s))\,dB(s) = \lim_{n\to\infty} \sum_i F(B(t_i))(B(t_{i+1}) - B(t_i)),$$

where the limit exists in $L^2(\mathbb{P})$. This is a fairly weak type of convergence, as opposed to Young's integral. Indeed, $B \mapsto \mathscr{I}(B) = \int_0^1 F(B(s))\,dB(s)$ is a only a measurable map and is *not* continuous. This is an unsatisfactory feature of Itô's theory.

Lyons made a very important observation, namely if we have a candidate for $\mathbb{B}(s,t) = \int_s^t (B(\theta) - B(s)) \otimes dB(\theta)$ (where the tensor denotes making a matrix out of this), then the map $(B, \mathbb{B}) \mapsto \mathscr{I}(B, \mathbb{B}) = \int_0^t F(B)\, dB$ is now continuous. (Though $\mathbb{B}(s,t)$ must satisfy some algebraic equations known as Chen's relations.)



For this theory, we can replace B with any function (or possibly random rough path) that is in $\mathcal{C}^\alpha$, provided that $\alpha > 1/3$.

2.3 The stochastic heat equation

We are now ready to discuss stochastic partial differential equations.

Example 2.1 (Stochastic heat equation)**.** The stochastic heat equation (SHE) is

$$u_t = \Delta u + \xi,$$

where ξ is white noise in (x,t). By this, we mean that ξ is a Gaussian process, $\mathbb{E}[\xi(x,t)] = 0$, and $\mathbb{E}[\xi(x,t)\xi(y,s)] = \delta_0(x-y, t-s)$ (to be formally defined later). One can show that $\xi \in \mathcal{C}^\alpha$ for any $\alpha < -d/2 - 1$. (Here, we are better off to use a "parabolic" metric, i.e. $|(x,t)-(y,s)|_{\mathrm{par}} = |x-y| + |t-s|^{1/2}$. Then Hölder means $\frac{|f(x,y)-f(y,s)|}{|(x,t)-(y,s)|_{\mathrm{par}}^\alpha}$.)

Because of "parabolic regularity" (which we will discuss later), we expect $u \in \mathcal{C}^{(-d/2+1)-}$. For example, when $d=1$, $u \in \mathcal{C}^{1/2-}$ in the space variable, and it turns out that $u \in \mathcal{C}^{1/4-}$ in the time variable. In higher dimensions, this will not be a function; we have to live with distributions. We can make sense of this PDE by first using Duhamel to write

$$u(x,t) = \int p(x-y,t) u^0(y)\, dy + \int_0^t \int p(x-y, t-s) \underbrace{\xi(y,s)\, dy\, ds}_{W(dy,ds)},$$

where p is a fundamental solution of the heat equation and $W(dy,ds)$ is known as "cylindrical Brownian motion."

3 Important Stochastic PDEs

3.1 The stochastic heat equation

Last time, we considered the stochastic heat equation

$$u_t = \Delta u + \xi, \qquad x \in \mathbb{R}^d, t \in \mathbb{R}$$

where ξ is space time white noise. We stated that we expect $u \in \mathcal{C}^{-d/2+1}$. In particular, $u \in \mathcal{C}^{1/2}$ in x and $\in \mathcal{C}^{1/4}$ in t when $d = 1$, but when $d > 1$ we don't have a function; it will be a distribution.

Later, we will see how a "subcritical" perturbation can be treated after a "renormalization." To explain this, let us first study the scaling properties of the above stochastic heat equation. Recall that ξ is a 0-mean Gaussian with $\mathbb{E}[\xi(x,t)\xi(y,s)] = \delta_0(x-y, t-s)$. So $\lambda \to \infty$, $\lambda^{d+1}\rho(\lambda x, \lambda^2 t) \to \delta_0(x,t)$. Observe that $\lambda^{d+2}\delta_0(\lambda x, \lambda^2 x) = \delta_0(x,t)$. Hence,

$$\widehat{\xi}(x,t) = \lambda^{(d+2)/2}\xi(\lambda x, \lambda^2 t) \overset{d}{=} \xi(x,t).$$

Now we go back to the stochastic heat equation, and if u is a solution, and if $\widehat{u}(x,t) = \lambda^{d/2-1}u(\lambda x, \lambda^2 t)$, then

$$(\widehat{u} - \Delta\widehat{u})(x,t) = \lambda^{d/2+1}(u_t - \Delta u)(\lambda x, \lambda^2 t) = \widehat{\xi} \overset{d}{=} \xi.$$

Thus, $\widehat{u}$ is again a solution of the stochastic heat equation. This is compatible with our guess for the Hölder regularity of the solution, namely $u \in \mathcal{C}^{(1-d/2)-}$ in x and $\in \mathcal{C}^{(1/2-d/4)-}$ in t.

3.2 The SHE with multiplicative noise

This PDE looks like

$$Z_t = \Delta Z + \sigma(Z)\xi$$

for a suitable function $\sigma : \mathbb{R} \to \mathbb{R}$.

Two examples that are particularly important are:

1. $\sigma(Z) = \sqrt{Z}$. This example appears in several models in math biology and population dynamics. Imagine you are modeling fish in a lake. Say each fish has an independent exponentially distributed clock that tells you when it dies. When it dies, you replace the fish with a number of descendants.

   Imagine that each particle travels as a Brownian motion, all independent, and after an exponential random time, a particle is replaced with N many particles with $\mathbb{E}[N] = m$. When $m = 1$, we have a critical regime, and as the initial number of particles goes to infinity, we get a measure-valued process known as **super Brownian motion**. When $d = 1$, this measure has a density Z, and Z solves this SHE with multiplicative noise for $\sigma(Z) = \sqrt{Z}$. This is also associated with **Brownian snake**.

2. $\sigma(Z) = Z$. As we will see shortly, this case is related to stochastic growth models.

It turns out that we can make sense of the SHE with multiplicative noise à la Itô. In other words, we write

$$Z(x,t) = \int p(x-y,t)Z(y,0)\,dt + \int_0^t \int p(x-y,t-s)Z(y,s)\underbrace{\xi(y,s)\,dy\,dx}_{W(dy,ds)}$$

when $d = 1$. Note that we still have the Hölder continuous Z multiplied by the distribution ξ. Hairer treated this PDE in 2013.

3.3 The Kardar-Parisi-Zhang equation

We wish to model stochastic growths. Often we have a random interface separating different phases. If the interface can be represented by a graph of a (height) function $h : \mathbb{R}^d \times [0,T] \to \mathbb{R}$, then the Hamilton-Jacobi PDE of the form

$$h_t = H(h_x) \quad (+\Delta u)$$

would be a good model as a first approximation. To capture the roughness of the interface, we may write

$$h_t = H(h_x) + D\Delta h + \lambda\xi.$$

After some manipulation (expanding h about a linear function), we end up with the KPZ equation:

$$h_t = |h_x|^2 + \Delta h + \xi.$$

This is a far more singular PDE than what we have seen before. Note that when $d = 1$, we expect $h \in \mathcal{C}^{1/2-}$, and $h_x \in \mathcal{C}^{-1/2}$.

The main challenge is to make sense of $|h_x|^2$. Indeed, the KPZ equation is "subcritical" only when $d = 1$. To explain this, let h be a solution to this equation, and set $\widehat{h}(x,t) = \lambda^{d/2-1}h(\lambda x, \lambda^2 t)$. Then

$$\begin{aligned}(\widehat{h}_t - \Delta\widehat{h})(x,t) &= \underbrace{\lambda^{d/2+1}\xi(\lambda x, \lambda^2 t)}_{\widehat{\xi}(x,t)} + \lambda^{d/2+1}|h_x(\lambda x, \lambda^2 t)|^2 \\ &= \widehat{\xi}(x,t) + \lambda^{1-d/2}|\widehat{h}_x(x,t)|^2.\end{aligned}$$

There are a few cases:

1. If $d = 1$, then as $\lambda \to 0$, the nonlinearity disappears. So, locally, the nonlinearity can be ignored!

2. If $d = 2$, this is the *critical regime*. In fact, if we multiply $|h_x|^2$ with a constant of size $\frac{C}{\sqrt{|\log\varepsilon|}}$ (after some smoothing), then we know how to handle the PDE.

3. If $d > 2$, then this is an open problem. We need to replace $\frac{C}{\sqrt{|\log \varepsilon|}}$ with $C\varepsilon^{d/2-1}$.

First observe that if $Z = e^h$ and h solves the KPZ equation, then Z solves the SHE with multiplicative noise. This is called the **Hopf-Cole transform**. This is surprising because the type of singularity we encounter in the KPZ equation is much worse than in the SHE with multiplicative noise. The problem is that the type of solution we had for the SHE with multiplicative noise à la Itô, which means that the usual chain rule must be corrected. Recall that if $\dot{y} = b(x,t) + \sigma dB(t)$, then

$$d\varphi(y) = \varphi'(y)(b\,dt + \sigma\,dB(t)) + \frac{1}{2}\varphi''(y)\sigma^2\,dt,$$

where d means the derivative. Recall that $\sum_j (B(t_{j+1}) - B(t_j))^2 \to t$, so $(dB)^2 = dt$. Thus, we get the Itô correction.

Let's go back to Hopf-Cole and do it carefully. To do this carefully, take a smooth kernel χ with $\int \chi = 1$, set $\xi^\varepsilon(x) = \varepsilon^{-d}\xi(x/\varepsilon)$, and put

$$\xi^\varepsilon(x,t) = \int \chi^\varepsilon(x-y)\xi^\varepsilon(y,t).$$

Then first solve

$$Z_t^\varepsilon = Z_{xx}^\varepsilon + \xi^\varepsilon Z^\varepsilon.$$

Fix x, and treat this equation as a stochastic differential equation in t. Observe that

$$\begin{aligned}
\mathbb{E}[\xi(x,t)^2] &= \mathbb{E}\left[\left(\int \xi(y,t)\chi^\varepsilon(x-y)\,dy\right)^2\right] \\
&= \mathbb{E}\left[\iint \xi(y,t)\xi(y',t)\chi^\varepsilon(x-y)\chi^\varepsilon(x-y')\,dy\,dy'\right] \\
&= \delta_0(t)\int (\chi^\varepsilon(x-y))^2\,dy \\
&= \delta_0(t)\varepsilon^{-1}\underbrace{\left(\int \chi^2\right)}_{\bar{c}}.
\end{aligned}$$

If $h^\varepsilon = \log Z^\varepsilon$, this satisfies

$$h_t^\varepsilon = h_{xx}^\varepsilon + \left[(h_x^\varepsilon)^2 - \frac{1}{2}\bar{c}\varepsilon^{-1}\right] + \xi^\varepsilon.$$

We aimed for the KPZ equation, but letting $h^\varepsilon \to h$, we get that

$$h_t = h_{xx} + (h_x^2 - \infty) + \xi.$$

So we get that this blows up, but we know exactly how.

4 Final Overview of Stochastic PDEs

4.1 The KPZ equation

Last time, we argued that by Itô calculus, we can make sense of the SPDE

$$Z_t = Z_{xx} + Z\xi$$

when $d = 1$. We want to use this solution to come up with a candidate of a solution to the KPZ equation

$$h_t = h_{xx} + |h_x|^2 + \xi.$$

We may use the Hopf-Cole transform to get a solution for this equation utilizing the previous SPDE. To achieve this, we smoothize ξ in the first SPDE by replacing ξ with $\xi^\varepsilon *_x \chi^\varepsilon$, which is white in time and smooth in space. Here, $\chi^\varepsilon(x) = \frac{1}{\varepsilon}\chi(\frac{x}{\varepsilon})$ with χ a smooth function of compact support and total integral 1. Then

$$Z_t^\varepsilon = Z_{xx}^\varepsilon + Z^\varepsilon \xi^\varepsilon.$$

As we saw last time, for fixed x, $\xi(x,t)$ is a multiple of standard white noise with

$$\mathbb{E}[\xi(x,t)\xi(x,s)] = \delta_0(t-s),$$

$$\int (\xi^\varepsilon)^2(y)\,dy = \delta_0(t-s)\varepsilon^{-1} \underbrace{\int \chi^2(y)\,dy}_{\overline{C}} =: \delta_0(t-s)C^\varepsilon.$$

In other words, if B represents a standard Brownian motion, we can represent

$$\xi^\varepsilon(x,t) \overset{d}{=} \sqrt{C^\varepsilon}\dot{B}(t).$$

Writing $z(t) = Z^\varepsilon(x,t)$, we can write the smoothized equation as

$$dz = \underbrace{b(t)}_{Z_{xx}^\varepsilon(x,t)}\,dt + Z^\varepsilon(x,t)\sqrt{C^\varepsilon}\,dB.$$

We now apply Hopf-Cole:

$$d(\underbrace{\log z}_{h^\varepsilon}) = \frac{dz}{z} - \frac{(Z^\varepsilon)^2 C^\varepsilon}{z^2}\,dt$$

(using $(dB)^2 = dt$). Simplifying, we get

$$dh^\varepsilon = \left(\frac{Z_{xx}^\varepsilon}{Z^\varepsilon} - \frac{C^\varepsilon}{2}\right)dt + \sqrt{C^\varepsilon}\,dB.$$

Here,

$$h^\varepsilon = \log Z^\varepsilon, \qquad h^\varepsilon_x = \frac{Z^\varepsilon_x}{Z^\varepsilon}, \qquad h^\varepsilon_{xx} = \frac{Z^\varepsilon_{xx}}{Z^\varepsilon} - (h^\varepsilon_x)^2.$$

Hence,

$$h^\varepsilon_t = h^\varepsilon_{xx} + \left[(h^\varepsilon_x)^2 - \frac{C^\varepsilon}{2}\right] + \xi^\varepsilon.$$

Thus, we can renormalize the KPZ equation by subtracting a constant multiple of $1/\varepsilon$ from the right hand side:

$$h_t = h_{xx} + (h_x^2 - \infty) + \xi$$

4.2 Stochastic quantization

In Euclidean Quantum Field Theory, we need to make sense of probability measures that are formally expressed as

$$\frac{1}{Z} e^{-\mathscr{H}(\phi)}\, D\phi,$$

where ϕ is a field, i.e. $\phi : \mathbb{R}^d \to \mathbb{R}$, and $D\phi$ is a Lebesgue-like measure on the space of ϕs. This may be compared with the following finite dimensional model: $H : \mathbb{R}^N \to \mathbb{R}$ and the minimizer of H correspond to the equilibrium states. If we take into account the thermal fluctuations, we would have equilibrium measures of the form

$$\frac{1}{Z} e^{-H(x)} \underbrace{dx}_{\text{Leb in } \mathbb{R}^N}.$$

Observe that a gradient ODE would allow us to give a dynamical approximation to our equilibrium states. For example, $\dot{x} = -\nabla H(x)$ would allow us to approximate the minimizer of H. As for $\frac{1}{Z} e^{-H(x)}\, dx$, we need to solve

$$\dot{x} = -\nabla H(x) + \dot{B}(t).$$

Then the law of $x(t)$ as $t \to \infty$ is exactly $\frac{1}{Z} e^{-H(x)}\, dx$.

In 1981, Parisi and Wu suggested that a dynamical approximation as in this previous equation would approximate the formal probability measures with a mathematically more tractable model. Indeed, if we have a candidate for an inner product on our function space, then

$$\phi_t = -\partial \mathscr{H}(\phi) + \xi(x,t),$$

which is called the **stochastic quantization**. Hopefully, $\phi(\cdot, t) \approx \frac{1}{Z} e^{-\mathscr{H}(\phi)}\, D\phi$ for large t.

Let's consider some examples:

Example 4.1. Consider

$$\mathscr{H}(\phi) = \int_{\mathbb{R}^d} \left(\frac{1}{2} |\nabla \phi|^2 + V(\phi) \right) dx,$$

where $V : \mathbb{R} \to \mathbb{R}$. We may replace $\mathbb{R}^d$ with a bounded domain with a suitable boundary condition. If we use the L^2 inner product, then

$$(\partial \mathscr{H})_\phi \psi = \int \left(-\Delta \phi + V'(\phi) \right) \psi.$$

Hence, the stochastic quantization equation becomes

$$\phi_t = \Delta_x \phi - V'(\phi) + \xi.$$

This is a perturbation of the SHE. The best we can hope for is a regularity of the form $\phi \in \mathcal{C}^{(1-d/2)-}$, which means that ϕ is a function only when $d = 1$. Hence, $V'(\phi)$ is the main challenge when V' is nonlinear.

4.3 The Gaussian Free Field

Here is a brief history of $\frac{1}{Z} e^{-\mathscr{H}(\phi)} \, D\phi$ and stochastic quantization. First consider the case $V = 0$ (or $V(\phi) = m^2 \phi^2 / 2$). Then what we have for our formal probability measure is a Gaussian measure though in infinite dimension. Using the L^2 inner product and when $V = 0$, what we have is

$$\frac{1}{Z} e^{-\frac{1}{2} \langle (-\Delta) \phi, \phi \rangle}.$$

This is the celebrated **Gaussian Free Field (GFF)**. Its covariance is $(-\Delta)^{-1}$, which has a kernel known as Green's function. In a domain D, we write $G^D(x, y)$ for this kernel: Under GFF,

$$\mathbb{E}[\phi(x) \phi(y)] = G^D(x, y).$$

However, we expect $\phi \in \mathcal{C}^{(1-d/2)-}$, hence not a function when $d > 1$.

For example, when $d = 1$, $D = (0, \infty)$, and we have the boundary condition $\phi(0) = 0$, then

$$G^D(x, y) = \min(x, y).$$

This is the correlation of Brownian motion in $d = 1$. Similarly, for $D = (0, \ell)$ with 0 boundary condition, we get

$$G^D(x, y) = \min(x, y) - \frac{1}{\ell} xy,$$

which corresponds to a Brownian bridge in $(0, \ell)$.

More generally, we have Feynman-Kac

$$\frac{1}{Z}e^{-\int(\frac{1}{2}|\phi'(x)|^2+V(\phi(x)))\,dx}\,D\phi = e^{-\int V(\phi(x))\,dx}\underbrace{\mu_0(d\phi)}_{\text{law of BM}}.$$

Next, consider $d = 2$. In this case, the GFF is "conformally invariant." This has to do with the fact that if $h : D \to D'$ is conformal, then $G^D(z, z') = G^{D'}(h(z), h(z'))$. In fact, ϕ in GFF can be used to study Schramm-Loewner Evolution in critical statistical mechanics ($\dot{z} = e^{\gamma\phi(z)}$). Also, there are models for randomly selected Riemannian metrics that can be expressed as $e^{\gamma\phi(x,y)}(dx^2 + dy^2)$, where ϕ is selected according to the GFF.

Finally, let us go back to the PDE

$$\phi_t = \Delta\phi - V'(\phi) + \xi$$

and examine the existence of a solution when V' is not linear. As a classical example, consider $V(\phi) = \phi^4/4$, so that $V'(\phi) = \phi^3$. Again, it is not clear how to make sense of ϕ^3 when $d \geq 2$, as ϕ is a distribution. To get a feel for this, first let us figure out when this equation is subcritical. Let ϕ solve this equation, and set $\widehat{\phi}(x, t) = \lambda^{d/2-1}\phi$. Then we can readily show

$$\widehat{\phi}_t = \Delta\widehat{\phi} - \lambda^{4-d}\widehat{\phi}^3 + \widehat{\xi}.$$

So the model is subcritical iff $d \leq 3$. The case $d = 2$ was solved back in the late 80s. The case $d = 3$ was solved in 2014 by Hairer. We need to renormalize the equation as

$$\phi_t^\varepsilon = \Delta\phi^\varepsilon - [(\phi^\varepsilon)^3 - c_\varepsilon\phi^\varepsilon] + \xi^\varepsilon$$

with $c_\varepsilon = O(\varepsilon^{-1})$.

5 Integration With Respect to Rough Functions

5.1 Lyons' approach and Chen's relation

Today, we try to solve ODEs of the form $\dot{y} = V(y)\dot{x}$, where V is a $\mathcal{C}^2$ function, and x is a Hölder continuous function, say, of exponent α. If $\alpha > 1/2$, we can study this ODE by first making sense of integrals of the form $\int_0^t V(y(\theta))\,dx(\theta)$. We develop a strategy to deal with such integrals when $1/3 < \alpha \le 1/2$. Let's explain the idea first.

We wish to make sense of

$$y(t) - y(s) = \int_s^t V(y(\theta))\,dx(\theta).$$

To make sense of the right hand side, we may try the following approximation for small $t-s$:

$$\int_s^t V(y(\theta))\,dx(\theta) \approx V(y(s)) \underbrace{\int_s^t dx(\theta)}_{|t-s|^\alpha} + O(|t-s|^{2\alpha}).$$

If we have a very fine mesh for defining our integral, then $\sum_i |t_i - t_{i+1}|^{2\alpha}$ is small only when $2\alpha > 1$. This suggests a finer Taylor expansion of the form

$$\begin{aligned}
\int_s^t V(y(\theta))\,dx(\theta) &= \int_s^t [V(y(s)) + DV(y(s))(y(\theta) - y(s))]\,dx(\theta) + O(|t-s|^{3\alpha}) \\
&= \int_s^t [V(y(s)) + DV(y(s))V(y(s))(x(\theta) - x(s))\,dx(\theta)] + O(|t-s|^{3\alpha}) \\
&= V(y(s)) \int_s^t dx(\theta) + DV(y(s))V(y(s)) \int_s^t (x(\theta) - x(s))\,dx(\theta) \\
&\qquad + O(|t-s|^{3\alpha})
\end{aligned}$$

To make this work, we still need to make sense of

$$\int_s^t (x(\theta) - x(s)) \otimes\, dx(\theta) = \left[\int_s^t (x^i(\theta) - x^i(s))\,dx^j(\theta)\right]_{i,j=1}^{\ell}.$$

Terry Lyons' idea in 1990 was to choose a candidate for $\mathbb{X}(s,t) = \int_s^t (x(\theta) - x(s)) \otimes dx(\theta)$, and given $(x(\cdot), \mathbb{X}(\cdot,\cdot))$, we can make sense of integrals of the form $\int_s^t V(y(\theta))\,dx(\theta)$. For example, given $(x, \mathbb{X})$, we can define

$$\mathscr{I}(\mathbf{x}) = \int_0^T F(\mathbf{x})\,d\mathbf{x}$$

for any C^2 function F, with $\mathscr{I}(\mathbf{x})$ continuous in $\mathbf{x}$.

Theorem 5.1 (Lyons-Victoire)**.** *Given $x \in \mathcal{C}^\alpha$, there exists a function $z \in \mathcal{C}^\alpha$ such that $z(0) = 0$ and*

$$|z(t) - z(s) - x(s) \otimes (x(t) - x(s))| \leq x_0 [x]_\alpha^2 |t - s|^{2\alpha}.$$

Here, $[x]_\alpha = \sup_{s \neq t \in [0,1]} \frac{|x(t)-x(s)|}{|t-s|^\alpha}$.

Here, we want to think of

$$z(t) = \int_0^t x(\theta) \otimes \, dx(\theta),$$

so that

$$z(t) - z(s) = \int_s^t x(\theta) \otimes \, dx(\theta).$$

We also want to think of

$$z(t) - z(s) - x(s) \otimes (x(t) - x(s)) = \mathbb{X}(s,t).$$

Let us write $x(s,t) = x(t) - x(s)$, so that we can write

$$z(s,t) := z(t) - z(s) = \mathbb{X}(s,t) + x(s) \otimes x(s,t).$$

From $s < u < t \implies z(s,u) + z(u,t) = z(s,t)$, we learn that $\mathbb{X}(s,t)$ must satisfy the following formula, known as **Chen's relation**:

$$\mathbb{X}(s,u) + \mathbb{X}(u,t) = \mathbb{X}(s,t) + [x(s) \otimes x(s,t) - x(s) \otimes x(s,u) - x(u) \otimes x(u,t)]$$

Using $x(s,t) = x(s,u) + x(s,t)$, we get

$$= \mathbb{X}(s,t) + x(s,u) \otimes x(u,t).$$

We can now define

$$[(x(\cdot), \mathbb{X}(\cdot,\cdot))]_\alpha := [x]_\alpha + \sup_{s \neq t \in [0,T]} \frac{|\mathbb{X}(s,t)|}{|t-s|^{2\alpha}}.$$

Remark 5.1 (Geometric Rough Path)**.** Roughly, $\dot{z}^{i,j} = x^i \dot{x}^j$. Then

$$\dot{z}^{ij} + \dot{z}^{ji} = x^i \dot{x}^j + x^j \dot{x}^i = \frac{d}{dt}(x^i x^j).$$

If the product rule applies, we expect

$$z^{ij}(s,t) + z^{ji}(s,t) = x^i(t) x^j(t) - x^i(s) x^j(s).$$

In general, this may not be true. For example, Itô calculus is not geometric, while Stratonovich calculus is geometric.

5.2 Convergence of the integral

Theorem 5.2 (Lyons)**.** *Let $(x, \mathbb{X})$ be as above (Chen's relation + $[(x,\mathbb{X})]_\alpha < \infty$), and let $F \in \mathcal{C}^2$. Then we can define*

$$\mathscr{I}(F) = \int_0^t F(\mathbf{x}) \cdot d\mathbf{x} = \lim_{|\pi|\to 0} \underbrace{\sum_i [F(x(t_i)) \cdot x(t_i, t_{i+1}) + DF(x(t_i))^* \mathbb{X}(t_, t_{i+1})]}_{\mathscr{R}(\pi)},$$

where $\pi = \{0 < t_1 < \cdots < t_n < t\}$ and $|\pi| = \max_i |t_{i+1} - t_i|$. Moreover,

$$\left| \int_s^t F(\mathbf{x}) \cdot d\mathbf{x} - (F(x(s)) \cdot x(s,t) + \underbrace{DF(x(s))^*}_{A(s)} \mathbb{X}(s,t)) \right| \leq c(\alpha) \|F\|_{C^2} [(x,\mathbb{X})]_\alpha^3 |t-s|^{3\alpha}.$$

Proof. Take a partition $\pi = \{s < t_0 < \cdots < t_{n-1} < t_n < t = t_{n+1}\}$. Pick some i, and compare $\mathscr{R}(\pi)$ with $\mathscr{R}(\pi - \{t_i\})$:

$$\begin{aligned}
\mathscr{R}(\pi) - \mathscr{R}(\pi - \{t_i\}) &= F(x(t_{i-1}))x(t_{i-1}, t_i) + F(x(t_i))x(t_i, t_{i+1}) \\
&\quad + A(t_{i-1})\mathbb{X}(t_{i-1}, t_i) + A(t_i)\mathbb{X}(t_i, t_{i+1}) \\
&\quad - F(x(t_{i-1}))x(t_{i-1}, t_{i+1}) + A(t_{i-1})\mathbb{X}(t_{i-1}, t_{i+1}) \\
&= y(t_{i-1}, t_i)x(t_i, t_{i+1}) + A(t_{i-1}, t_i)\mathbb{X}(t_i, t_{i+1}) \\
&\quad - A(t_{i-1})x(t_{i-1}, t_i) \otimes x(t_i, t_{i+1}) \\
&= [y(t_{i-1}, t_i)x(t_i, t_{i+1}) - A(t_{i-1})x(t_{i-1}, t_i) \otimes x(t_i, t_{i+1})] \\
&\quad + A(t_{i-1}, t_i)\mathbb{X}(t_i, t_{i+1}).
\end{aligned}$$

So we may estimate the error as

$$\begin{aligned}
|\mathscr{R}(\pi)| &= |[y(t_{i-1}, t_i)x(t_i, t_{i+1}) - A(t_{i-1})x(t_{i-1}, t_i) \otimes x(t_i, t_{i+1})] + A(t_{i-1}, t_i)\mathbb{X}(t_i, t_{i+1})| \\
&\leq \|F\|_{\mathcal{C}^2} |t_{i+1} - t_i|^{3\alpha} [x]_\alpha^3 + \|F\|_{\mathcal{C}^2} |t_{i+1} - t_{i-1}|^{3\alpha} \|F\|_{\mathcal{C}^2} [\mathbb{X}]_{2\alpha}.
\end{aligned}$$

Choose i so that $|t_{i+1} - t_i| \leq 2(t-s)/n$,

$$\mathscr{R}(\pi) - \mathscr{R}(\pi - \{t_i\}) \leq c_0 \frac{|t-s|^{3\alpha}}{n^{3\alpha}} 2^{3\alpha}.$$

Do this inductively to obtain

$$|\mathscr{R}(\pi) - \mathscr{R}(\varnothing)| \leq c_0 |t-s|^{3\alpha}.$$

From our proof, we can also deduce that $\mathscr{R}(\pi)$ converges as $|\pi| \to 0$. $\square$

6 Considerations for Integration Theories

6.1 Definition of Lyons' integral

We are interested in pairs of the form $\mathbf{x} = (x, \mathbb{X})$, where $x \in \mathcal{C}^\alpha$ (i.e. $x : [0,1] \to \mathbb{R}^\ell$ is Hölder of exponent α), and $\mathbb{X} : [0,T]^2 \to \mathbb{R}^{\ell\times\ell}$ (into the $\ell \times \ell$ real matrices) such that

$$x(s,t) = x(t) - x(s), \qquad \mathbb{X}(s,t) = \mathbb{X}(s,u)\mathbb{X}(u,t) + x(s,u) \otimes x(u,t),$$

which is **Chen's relation**. We write

$$\|\mathbf{x}\|_{\alpha,2\alpha} = |x(0)| + \underbrace{\sup_{s\neq t} \frac{|x(t)-x(s)|}{|t-s|}}_{[x]_\alpha} + \sup_{s\neq t} \frac{|\mathbb{X}(s,t)|}{|t-s|^{2\alpha}}.$$

We write $\mathscr{R}^\alpha = \{\mathbf{x} = (x,\mathbb{X}) : \|\mathbf{x}\|_{\alpha,2\alpha} < \infty, \text{Chen's relation holds}\}$. Last time, we proved the following theorem.

Theorem 6.1. *Assume $\alpha \in (1/3, 1/2]$. If $\mathbf{x} \in \mathscr{R}^\alpha$ and $F : \mathbb{R}^\ell \to \mathbb{R}^\ell \in \mathcal{C}^2$, then*

$$\int_0^t F(x) \cdot d\mathbf{x} := \lim_{\substack{|\pi|\to 0 \\ \pi=\{t_0=0<t_1<\cdots<t_{n+1}=t\}}} \sum_{i=0}^n [F(x(t_i)) \cdot x(t_i,t_{i+1}) + DF(x(t_i)) : \mathbb{X}(t_i,t_{i+1})]$$

here if $A = [a_{i,j}]$ and $B = [b_{i,j}]$, then $A : B := \sum_{i,j} a_{i,j}b_{i,j}$, exists, and

$$\left|\int_s^t F(x)\cdot d\mathbf{x} - (F(x(s))\cdot x(s,t) + DF(x(s)) : \mathbb{X}(s,t))\right| \le c_0(\alpha)\|F\|_{\mathcal{C}^2}\|\mathbf{x}\|^2_{\alpha,2\alpha}|t-s|^{3\alpha}.$$

The way to think about $\mathbb{X}(s,t)$ is

$$\mathbb{X}(s,t) = \int_s^t \int_s^\theta dx(\theta') \otimes dx(\theta).$$

6.2 Remarks on integration theories

Remark 6.1. Write $\mathscr{R}^\alpha(x) = \{\mathbb{X} : (x,\mathbb{X}) \in \mathscr{R}^\alpha\}$, with $\alpha > 1/3$. Now if $\mathbb{X}, \mathbb{X}' \in \mathscr{R}^\alpha(x)$, then $W = \mathbb{X}' - \mathbb{X}$, and

$$W(s,u) + W(u,t) = W(s,t).$$

So if $W(t) := W(0,t)$, then we can write $W(s,t) = W(t) - W(s)$. Moreover,

$$\sup_{s\neq t} \frac{|W(t)-W(s)|}{|t-s|^{2\alpha}} < \infty.$$

So $W \in \mathcal{C}^{2\alpha}$. Thus, if $\mathbb{X}^0 \in \mathscr{R}^\alpha(x)$, then

$$\mathscr{R}^\alpha(x) = \{(\mathbb{X}^0(s,t) + W(t) - W(s) : s,t \in [0,T]) : W \in \mathcal{C}^{2\alpha}\}.$$

In particular, if $\alpha > 1/2$, $\mathscr{R}^\alpha(x)$ consists of one element.

Remark 6.2. To generalize our theorem, we define the following function space: Given $x \in \mathcal{C}^\alpha$, let $\mathscr{G}^\alpha(x)$ be the set of pairs $(y, \widehat{y})$ with the following properties:

- $y : [0,T] \to \mathbb{R}^{d\times\ell}$,
- $y \in \mathcal{C}^\alpha$ (could be $\mathcal{C}^\beta$ also),
- $\widehat{y} : [0,T] \to \mathbb{R}^{d\times\ell\times\ell}$,
- $\widehat{y} \in C^\alpha$,
- $$\|(y,\widehat{y})\|_{\alpha,2\alpha} = [y]_\alpha + [\widehat{y}]_\alpha + \sup_{s\neq t} \frac{|y(t) - y(s) - \widehat{y}(s) : (x(t) - x(s))|}{|t-s|^{2\alpha}} < \infty.$$

For example, if $F : \mathbb{R}^\ell \to \mathbb{R}^\ell \in \mathcal{C}^2$ and $x : [0,T] \to \mathbb{R}^\ell$, then $(y, \widehat{y}) = (F(x), DF(x)) \in \mathscr{G}(x)$. We call $\mathscr{G}(x)$ the Gubinelli class of x.

With an identical proof we can show this: If $x = (x, \mathbb{X}) \in \mathscr{R}^\alpha$ and $y = (y, \widehat{y}) \in \mathscr{G}\alpha(x)$, then

$$\int_0^t \mathbf{y} \cdot d\mathbf{x} := \lim_{\substack{|\pi|\to 0 \\ \pi=\{t_0=0<t_1<\cdots<t_{n+1}=t\}}} \left[\sum_i y(t_i)x(t_i,t_{i+1}) + \widehat{y}(t_i) : \mathbb{X}(t_i,t_{i+1})\right].$$

The analogue of the bound in the theorem also holds, provided that $\|F\|_{\mathcal{C}^2}$ is replaced with $\|\mathbf{y}\|_{\alpha,2\alpha}$.

Remark 6.3. If $\alpha > 1/2$, then

$$\int_0^t \mathbf{y} \cdot \, d\mathbf{x} =: \int_0^t y \cdot \, dx$$

because in the above limit definition of the integral, the contribution from $\sum_i \widehat{y}(t_i) : (t_i, t_{i+1})$ is 0, so we can drop it. If this is the case, we refer to it as a **Young integral**.

Remark 6.4. Suppose $\mathbb{X}^0 \in \mathscr{R}^\alpha(x)$, and let $W \in \mathcal{C}^{2\alpha}$ with $\mathbb{X}(s,t) = \mathbb{X}^0(s,t) + W(t) - W(s)$. Now

$$\int_0^t (y,\widehat{y}) \cdot \, d(x,\mathbb{X}) = \int_0^t (y,\widehat{y}) \cdot d(x,\mathbb{X}^0) + \underbrace{\int_0^t \widehat{y} : \, dW}_{\text{Young integral}}.$$

Remark 6.5. We say $\mathbf{x} = (x, \mathbb{X}) \in \mathscr{R}_g^\alpha$, i.e. $\mathbf{x}$ is **(weakly) geometric**, if

$$\mathbb{X}(s,t) + \mathbb{X}^*(s,t) = x(s,t) \otimes x(s,t).$$

Equivalently, we can say

$$\mathbb{X}_{i,j}(s,t) + \mathbb{X}_{j,i}(s,t) = x_i(s,t)x_j(s,t),$$

where $\mathbb{X}_{i,j} = \int_s^t x_i \, dx_j - x_i(s)(x_j(t) - x_j(s))$. Hence, if $\mathscr{R}_g^\alpha(x) = \{\mathbb{X} : (x, \mathbb{X}) \in \mathscr{R}_g^\alpha\}$, then the symmetric part of $\mathbb{X}$ is uniquely determined. Hence if $\mathbb{X}^0 \in \mathscr{R}_g^\alpha$, then

$$\mathscr{R}_g^\alpha(x) = \{\mathbb{X}^0(s,t) + W(t) - W(s) : W \in \mathcal{C}^{2\alpha}, W^* = -W\}.$$

Now consider the corresponding integral:

$$\int_0^t F(x) \cdot \, d(x, \mathbb{X}) = \int_0^t F(x) \cdot \, d(x, \mathbb{X}^0) + \int_0^t DF(\mathbf{x}) : \, dW.$$

Example 6.1. Take any $\mathbf{x} = (x, \mathbb{X}) \in \mathscr{R}^\alpha$, and pick any 1-periodic function $f : [0,1] \to \mathbb{R}^\ell$. If $y_n(t) = n^{-1/2} f(nt)$, then $\mathbf{y} \to 0$. Now consider $x_n = x + y_n \xrightarrow{n\to\infty} x$. Then one can show that the norm is uniformly bounded. Define

$$\mathbb{X}_n = \mathbb{X} + \underbrace{\int_s^t (y_n(\theta) - y_n(s)) \otimes \, dy_n(\theta)}_{\text{classical integral}} \to \mathbb{X} + (t-s)C,$$

where C is an antisymmetric matrix.

Remark 6.6. Start from $(\mathscr{R}^\alpha, \| \cdot \|_{\alpha, 2\alpha})$. Let us define

$$\mathcal{C}_\infty = \left\{ (x, \mathbb{X}) : x \in C^1, \mathbb{X} \text{ is defined by } X(s,t) = \int_s^t (x(\theta) - x(s)) \otimes \, dx(\theta) \right\},$$

where $\int_s^t (x(\theta) - x(s)) \otimes dx(\theta)$ is a classical integral. Write $\mathscr{R}_{\mathrm{sg}}^\alpha$ to be the closure of $\mathcal{C}_\infty$ with respect to $\| \cdot \|_{\alpha, 2\alpha}$. It is not hard to see[3] $\mathscr{R}_{\mathrm{sg}} \subsetneq \mathscr{R}_g^\alpha$. This has to do with the fact that $\mathcal{C}^\alpha$ is not topologically separable. In fact, what is the closure of the set of smooth functions with respect to $\| \cdot \|_\alpha$? The closure is exactly the set of $x : [0,T] \to \mathbb{R}^d$ such that

$$\lim_{\varepsilon \to 0} \underbrace{\sup_{\substack{|s-t|<\varepsilon \\ s \neq t}} \frac{|x(t) - x(s)|}{|t-s|^\alpha}}_{\psi(\varepsilon)} = 0.$$

[3] Not my words.

7 Gubinelli's Sewing Lemma and Tensor Algebra for Increments

7.1 Gubinelli's sewing lemma

The method we have used so far can be used to show that if $x(t) = (f(t), g(t))$ with $f \in \mathcal{C}^\alpha, g \in \mathcal{C}^\beta$, then there exists a unique candidate for $\int_0^t f\,dg$ (Young's theorem), provided that $\alpha + \beta > 1$. (The general case $\alpha + \beta \leq 1$ with $f, g : \mathbb{R}^d \to \mathbb{R}$ will be treated later.) More precisely, we can find a β-Hölder $h : [0, T] \to \mathbb{R}$ such that $h(0) = 0$, and

$$|h(t) - h(s) - \underbrace{f(s)(g(t) - g(s))}_{=:A(s,t)}| \leq c_0|t - s|^{\alpha+\beta}.$$

In fact, Gubinelli's sewing lemma gives the sufficient (even necessary) conditions on A that would guarantee the existence of such an h

Definition 7.1. Given $A : [0, T]^2 \to \mathbb{R}$ and $\gamma > 0$, we say A is **γ-coherent** if

$$|A(s, t) - A(s, u) - A(u, t)| \leq c_0|t - s|^{1+\gamma}$$

for all s, u, t satisfying $0 \leq s \leq u \leq t \leq T$.

Lemma 7.1 (Sewing lemma, Gubinelli). *If A is γ-coherent, then*

$$h(t) = \lim_{|\pi|\to 0} \sum_{i=1}^{n} A(t_i, t_{i+1})$$

exists, where $\pi = \{0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = t\}$ is a partition of the interval $[0, t]$.

Proof. If $\pi = \{s = t_0 < t_1 < \cdots < t_n < t_{n+1} = t\}$ is a partition of $[s, t]$ and if $I(\pi) = \sum_{i=1}^{n} A(t_i, t_{i+1})$, then

$$I(\pi) - I(\pi \setminus \{t_i\}) = |A(t_{i-1}, t_i) + A(t_i, t_{i+1}) - A(t_{i-1}, t_{i-1})| \leq c_0|t_{i+1} - t_{i-1}|^{1+\gamma}.$$

We may choose t_i so that $|t_{i+1} - t_i| \leq \frac{2}{n}|t - s|$. We can repeat our previous argument to show that the limit exists and that

$$|h(t) - h(s) - A(s, t)| \leq c|t - s|^{1+\gamma}, \quad \text{where} \quad c = \sum_{n=1}^{\infty} \left(\frac{2}{n}\right)^{1+\gamma}. \qquad \square$$

Remark 7.1. Observe that if $A(s, t) = f(s)(g(t) - g(s))$, then

$$\begin{aligned}
|A(s, t) - A(s, u) - A(u, t)| &= |f(s)(g(t) - f(s)) - f(s)(g(u) - g(s)) - f(u)(g(t) - g(u))| \\
&= |(f(s) - f(u))(g(t) - g(u))| \\
&\leq [f]_\alpha[g]_\beta|t - s|^{\alpha+\beta}.
\end{aligned}$$

Note that our candidate h' represents fg', and we are comparing fg' with $f(s)g'$:

$$|h(t) - h(s) - f(s)(g(t) - g(s))| = |(h' - f(s)g')(\mathbb{1}_{[s,t]})| \leq c_0|t-s|^{1+\gamma}.$$

Perhaps we set $F_s = f(s)g'$, and the γ-coherence condition requires some kind of regularity of the map $s \mapsto F_s$.

$$\begin{aligned}
A(s,t) - A(s,u) - A(u,t) = &\underbrace{F_s(\mathbb{1}_{[s,t]})}_{F_s(\mathbb{1}_{[s,u]}+\mathbb{1}_{[u,t]})} - F_s(\mathbb{1}_{[s,u]}) - F_u(\mathbb{1}_{[u,t]}) \\
&= F_s(\mathbb{1}_{[u,t]}) - F_u(\mathbb{1}_{[u,t]}) \\
&= (F_s - F_u)(\mathbb{1}_{[u,t]}).
\end{aligned}$$

Perhaps we should write $\varphi = \mathbb{1}_{[0,1]}$ and $\varphi_x^\lambda(\theta) := \lambda^{-1}\varphi(\frac{\theta-x}{\lambda}) = \lambda^{-1}\mathbb{1}_{[x,x+\lambda]}$, which approximates the δ distribution at x. Then $(F_s - F_u)(\mathbb{1}_{[u,t]}) = \lambda(F_s - F_u)(\varphi_u^\lambda)$, where $\lambda = t - u$. Gubinelli's condition means that

$$|(F_s - F_u)(\varphi_u^\lambda)| \leq \lambda^{-1}(|s-u| + \lambda)^{1+\gamma}.$$

This condition is sharp.

7.2 Tensor algebra structure for increments

So far, for a rough path, we need a vector $x(s,t) = x(t) - x(s)$ and a matrix $\mathbb{X}(s,t)$. For $\alpha > 1/k$, we are dealing with a tensor algebra that is truncated at order k. For $k = 3$, we cut it at 3 and only deal with 1 and 2 tensors. Consider the vector space $V = \mathbb{R} \oplus \mathbb{R}^\ell \oplus \mathbb{R}^{\ell\times\ell}$ with elements (λ, v, A) (which we may write as $\lambda + v + A$). We equip V with a multiplication (tensor product)

$$(\lambda + v + A) \otimes (\lambda' + v' + A') = (\lambda a') + (\lambda v' + \lambda' v) + (\lambda A' + \lambda' A + v \otimes v').$$

Note that if $G = \{1 + v + A : v \in \mathbb{R}^\ell, A \in \mathbb{R}^{\ell\times\ell}\}$, then G is closed with respect to $\otimes$. In fact G is a group. Indeed,

$$\begin{aligned}
(1 + v + A)^{-1} &= 1 - (v + A) + (v + A) \otimes (v + A) + \cdots \\
&= 1 - (v + A) + v \otimes v \\
&= 1 - v + (v \otimes v - A).
\end{aligned}$$

Let's take a rough path: $x : [0,T] \to \mathbb{R}^\ell$, $\mathbb{X} : [0,T]^2 \to \mathbb{R}^{\ell\times\ell}$. Given such a path, set

$$\mathbf{x}(s,t) = 1 + \underbrace{x(t) - x(s)}_{x(s,t)} + \mathbb{X}(s,t),$$

so $\mathbf{x} : [0,T]^2 \to G$. Recall Chen's relation,

$$\mathbb{X}(s,t) = \mathbb{X}(s,u) + \mathbb{X}(u,t) + x(s,u) \otimes x(u,t).$$

Observe that

$$\begin{aligned}
\mathbf{x}(s,u) \otimes \mathbf{x}(u,t) &= (1 + x(s,u) + \mathbb{X}(s,u)) \otimes (1 + x(u,t) + \mathbb{X}(u,t)) \\
&= 1 + x(s,t) + (\mathbb{X}(s,t) + \mathbb{X}(s,u)) + (x(s,u) \otimes x(u,t)) \\
&= \mathbf{x}(s,t).
\end{aligned}$$

Thus, Chen's relation is equivalent to saying that $\mathbf{x}(s,t) = \mathbf{x}(s,u) \otimes \mathbf{x}(u,t)$. This says that with respect to $\otimes$, $\mathbf{x}(s,t)$ is an increment. We can also see that $\mathbf{x}(s,t) = \mathbf{x}(0,s) \otimes \mathbf{x}(s,t)$, which gives

$$\mathbf{x}(s,t) = \mathbf{x}(0,s)^{-1} \otimes \underbrace{\mathbf{x}(0,t)}_{\mathbf{x}(t)}.$$

In summary, $\mathscr{R}^\alpha$ is isomorphic to the set of paths $\mathbf{x} : [0,T] \to G$ and, by choosing a suitable metric on G, that are left-invariant with $\mathbf{x}$ being α-Hölder with respect to this metric:

$$[\![x]\!]_\alpha = \sup_{s \neq t} \frac{d_G(x(t), x(s))}{|t-s|^\alpha} < \infty.$$

How about $\mathscr{R}_g^\alpha$? Even in this case, we get the set of α-Hölder paths $\mathbf{x} : [0,T] \to \widehat{G}$, where $\widehat{G}$ is a subgroup of G. Remember that

$$\begin{aligned}
\mathscr{R}_g^\alpha &= \{(x,\mathbb{X}) \in \mathscr{R}^\alpha : \mathbb{X}(s,t) + \mathbb{X}^*(s,t) = x(s,t) \otimes x(s,t)\} \\
&= \left\{(x,\mathbb{X}) \in \mathscr{R}^\alpha : \mathbb{X}(s,t) = \frac{1}{2} x(s,t) \otimes x(s,t) + C(s,d), C^* + C = 0\right\}.
\end{aligned}$$

This suggests that

$$\widehat{G} = \{1 + v + (\tfrac{1}{2} v \otimes v + C) : C^* + C = 0\}.$$

8 Solving ODEs Via Rough Integration

8.1 Solving for the Itô-Lyons map

We now turn to the ODE of the form

$$\dot{y} = \sigma(y)\dot{x}, \qquad y(0) = y^0,$$

where $x \in \mathcal{C}^\alpha$ and σ is sufficiently smooth. Here, $x : [0,T] \to \mathbb{R}^\ell$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times\ell}$, and $y : [0,T] \to \mathbb{R}^d$. We find a unique solution to this ODE, provided that we choose a suitable $\mathbb{X}$ so that $\mathbf{x} = (x, \mathbb{X}) \in \mathscr{R}^\alpha$. The solution we come up with, $y(\cdot) = \mathscr{I}(y^0, \mathbf{x})$ is continuous (even locally Lipschitz) in y^0 and $\mathbf{x}$. $\mathscr{I}$ is known as the **Itô-Lyons map**. Let's make some preparations for this construction. Needless to say that we want to interpret this ODE as

$$y(t) = y^0 + \int_0^t \sigma(y(\theta))\, dx(\theta).$$

Though if $\alpha < 1/2$ (say $\alpha \in (1/3, 1/2]$), we need to lift both $\sigma(y)$ and x to $(\sigma(y), \widehat{\sigma})$, $(x, \mathbb{X})$ with $(\sigma(y), \widehat{\sigma}) \in \mathscr{G}^\alpha(x)$, $(x, \mathbb{X}) \in \mathscr{R}^\alpha$.

Recall that $\mathscr{G}^\alpha(x)$ consists of pairs $(z, \widehat{z})$ (where we intuitively think of $\widehat{z}$ as a "derivative" of z with respect to x) such that $z, \widehat{z} \in \mathcal{C}^\alpha$ and

$$[\![(z,\widehat{z})]\!]_{2\alpha} := \sup_{s\neq t} \frac{|z(t) - z(s) - \widehat{z}(s)(x(t) - x(s))|}{|t-s|^{2\alpha}} < \infty.$$

Indeed, from the integral formulation of this ODE, we expect that if y solves the equation, then $(y, \sigma(y)) \in \mathscr{G}^\alpha(x)$.

Theorem 8.1. *Let $\mathbf{x} = (x, \mathbb{X}) \in \mathscr{R}^\alpha$ for $\alpha \in (1/3, 1/2]$, and assume $\sigma \in \mathcal{C}_b^3$ (bounded derivatives). Then for each y^0, there exists a path $y \in \mathcal{C}^\alpha$ such that $y(0) = y^0$, $(y, \sigma(y)) \in \mathscr{G}^\alpha(x)$, and*

$$y(t) = y^0 + \int_0^t \underbrace{(\sigma(y), \widehat{\sigma}(y))}_{\sigma} \cdot d\underbrace{(x, \mathbb{X})}_{\mathbf{x}}.$$

Here, $\widehat{\sigma}(y) = [\widehat{\sigma}^{ijk}(y)]$ with

$$\widehat{\sigma}^{ijk}(y) = \sum_{r=1}^d \sigma^{i,j}_{y_r}(y)\sigma^{rk}(y).$$

Moreover, $\mathscr{I}(y^0, \mathbf{x})$ is Lipschitz with Lipschitz norm calculated in terms of $\|\sigma\|_{\mathcal{C}^3}$ and $\|\mathbf{x}\|_{\alpha,2\alpha}$.

The idea is to start from $\mathbf{y} = (y, \widehat{y})$ and set

$$\mathcal{F}_{\mathbf{x}}(\widehat{y})(t) = \left(y^0 + \int_0^\cdot (\sigma(y), \widetilde{\sigma}(y, \widehat{y})) \cdot d(x, \mathbb{X}), \sigma(y) \right),$$

where $\widetilde{\sigma}(y, \widehat{y}) = [\widetilde{\sigma}^{ijk}(y, \widehat{y})]$, where

$$\widetilde{\sigma}^{ijk}(y, \widehat{y}) = \sum_{r=1}^{d} \sigma^{ij}_{y_r}(y)\widehat{y}^{rk}.$$

If $\widehat{y}$ is a fixed point of $\mathcal{F}$, then we are done because then the Gubinelli derivative of such $\mathbf{y}$ must be $\sigma(y)$.

8.2 Breakdown of the map $\mathcal{F}$

Let's understand $\mathcal{F}$ better: Throughout, $\mathbf{x} = (x, \mathbb{X}) \in \mathscr{R}^{\alpha}$ is fixed.

Step 1: Recall that for $\mathbf{z} = (x, \widehat{z}) \in \mathscr{G}^{\alpha}(x)$, we can define $w(t) = \int_0^t \mathbf{z}\, dx$, which satisfies

$$|w(t) - w(s) - z(s)(x(t) - x(s)) - \widehat{z}(s)\mathbb{X}(s,t)| \leq c_0([z]_{\alpha}[x]_{\alpha} + [\widehat{z}]_{\alpha}[\mathbb{X}]_{2\alpha})|t-s|^{3\alpha}.$$

This suggests $\mathcal{F}_{\mathbf{x}} : \mathscr{G}^{\alpha}(x) \to \mathscr{G}^{\alpha}(x)$ by $\mathcal{F}^0_{\mathbf{x}}(z, \widehat{z}) = (w, z)$. In fact, $\mathcal{F}^0$ is linear and

$$[\![\mathcal{F}^0_{\mathbf{x}}(\mathbf{y})]\!]_{\alpha,2\alpha} \leq c_0[\mathbf{x}]_{\alpha,2\alpha}[\mathbf{y}]_{\alpha,2\alpha}.$$

Here is the short proof of this:

Proof.

$$\begin{aligned} |w(t) - w(s) - z(s)(x(t) - x(s))| &\leq \|\widehat{z}\|_{L^{\infty}}[\mathbb{X}]_{2\alpha}|t-s|^{2\alpha} \\ &\quad + c_0(\text{what we had before})|t-s|^{3\alpha}. \end{aligned}$$ □

Step 2: Define $\mathcal{F}^1_{\mathbf{x}} : \mathcal{G}^{\alpha}(x) \to \mathcal{G}^{\alpha}(x)$ with $\mathcal{F}^1_{\mathbf{x}}(z, \widehat{z}) = (\sigma(z), D\sigma(z)\widehat{z})$, where

$$(D\sigma(z)\widehat{z})^{ijk} = \sum_{r=1}^{d} \sigma^{ij}_{z_r}\widehat{z}^{rk}$$

and $\mathcal{F}^1$ is bounded if $\sigma \in \mathcal{C}^2$. Here is the proof:

Proof. Using a Taylor expansion for σ,

$$\begin{aligned} &|\sigma(z(t)) - \sigma(z(s)) - D\sigma(z(s))\widehat{z}(s)x(s,t)| \\ &\qquad \leq |D\sigma(z(s))(z(t) - z(s)) - D\sigma(z(s))\widehat{z}(s)x(s,t)| + \|D^2\sigma\|_{L^{\infty}}[z]_{\alpha}|t-s|^{2\alpha} \\ &\qquad \leq \|D\sigma\|_{L^{\infty}}[\mathbf{z}]_{\alpha,2\alpha}|t-s|^{2\alpha} + \|D^2\sigma\|_{L^{\infty}}[z]_{\alpha}|t-s|^{2\alpha} \\ &\qquad \leq \|\sigma\|_{\mathcal{C}^2}[\mathbf{z}]_{\alpha,2\alpha}|t-s|^{2\alpha}. \end{aligned}$$

So we get that

$$[\![\mathcal{F}^1_{\mathbf{x}}(\mathbf{z})]\!]_{\alpha,2\alpha} \leq \|\sigma\|_{\mathcal{C}^2}[\![\mathbf{z}]\!]_{\alpha,2\alpha}.$$ □

Step 3: Next, we define $\mathcal{F} : \mathscr{G}^\alpha(x) \to \mathscr{G}^\alpha(x)$, as $\mathcal{F} = \mathcal{F}^0 \circ \mathcal{F}^1$, so we send

$$(y, \widehat{y}) \mapsto (\sigma(y), D\sigma(y)\widehat{y}) \mapsto \left(\int_0^{\cdot} (\sigma\widehat{\sigma}) \cdot d(x, \mathbb{X}), \sigma(y) \right).$$

Then set

$$\mathcal{F}'(y, \widehat{y}) = \left(y^0 + \int_0^{\cdot} (\sigma, \widehat{\sigma}) \cdot d(x, \mathbb{X}), \sigma(y) \right).$$

We need to turn $\mathcal{F}'$ into a contraction so that it has a fixed point. We achieve this by choosing a sufficiently small interval $[0, t_0)$, and finding a nice invariant subset of $\mathscr{G}^\alpha(x)$. As we will see, t_0 depends on $\|\sigma\|_{C^3}$, so we can repeat the same construction on $[t_0, 2t_0), \ldots$.

How can this be done? First, switch from $\mathscr{G}^\alpha(x)$ to $\widehat{\mathscr{G}^\alpha}(x) = \{(y, \widehat{y}) : y(0) = y^0, \widehat{y}(0) = \sigma(y^0)\}$. This way, we don't need to worry about the difference between a norm and a seminorm; this contraction takes place in a metric space, which is good enough for our purposes. Observe that $(a, \widehat{a}) \in \widehat{\mathscr{G}^\alpha}(x)$, where $a(t) = y^0 + \sigma(y^0)(x(t) - x(0))$ and $\widehat{a}(t) = \sigma(y^0)$. Observe that

$$\underbrace{a(t) - a(s)}_{\sigma(y^0)(x(t)-x(s))} - \underbrace{\widehat{a}(s)}_{\sigma(y^0)}(x(t) - x(s)) = 0.$$

Now set $\mathscr{B} = \{(y, \widehat{y}) \in \widehat{\mathscr{G}^\alpha}(x) : [\![(y - a, \widehat{y} - \widehat{a})]\!]_{\alpha, 2\alpha} \leq 1\}$. The trick is to construct something in a rougher space and then show that it is as regular as you want. We will continue this next time.

9

9.1

10 Kolmogorov's Continuity Theorem for Rough Paths and Candidates for the Lift of Brownian Motion

10.1 Kolmogorov's continuity theorem for rough paths

Recall that if $A(x) = \int_0^T \int_0^T \psi(\frac{|x(t)-x(s)|}{p(|t-s|)})\, dt\, ds$ with $\psi, p : [0,\infty) \to [0,\infty)$ increasing, $\psi(0) = p(0) = 0$ and $\psi(\infty) = \infty$, then

$$|x(t) - x(s)| \leq 8 \int_0^{|t-s|} \psi^{-1}\left(\frac{4A}{\theta^2}\right) p(d\theta).$$

For example, if $\psi(r) = r^q$ and $p(r) = r^{\alpha+1/q}$ with $q > 1$ and $\alpha > 0$, then

$$|x(t) - x(s)| \leq c_0(q,\alpha) A(x)^{1/q} |t-s|^{\alpha-1/q}.$$

In summary, if

$$A(x) = \int_0^T \int_0^T \frac{|x(t)-x(s)|^q}{|t-s|^{\alpha q+1}}\, dt\, ds,$$

then x is Hölder continuous of exponent $\alpha - 1/q$. In particular, if x is randomly selected according to a probability measure $\mathbb{P}$ and $\mathbb{E}[|x(t)-x(s)|^q] \leq c_0|t-s|^{\beta q}$, then

$$\mathbb{E}[A(x)] \leq c_0 \int_0^T \int_0^T |t-s|^{\beta q - \alpha q - 1}\, dt\, ds < \infty$$

if $\beta > \alpha$. In summary, if we have this L^q bound on $x(t) - x(s)$, then x is Hölder of exponent $\gamma \in (0, \beta - 1/q)$. This is also true for $x : [0,T]^d \to \mathbb{R}^\ell$: If $(\mathbb{E}[|x(t)-x(s)|^q])^{1/q} \leq c_0|t-s|^\beta$, then x is Hölder of exponent $\gamma \in (0, \beta - d/q)$.

Here is a version of Kolmogorov's continuity theorem that involves rough paths:

Theorem 10.1. *Let $x : [0,T] \to \mathbb{R}^\ell$ and its lift $\mathbb{X} : [0,T]^2 \to \mathbb{R}^{\ell\times\ell}$ satisfy Chen's relation:*

$$\mathbb{X}(s,t) = \mathbb{X}(s,u) + \mathbb{X}(u,t) + x(s,u) \otimes x(u,t).$$

Let $q \geq 2$, $\beta > 1/q$, and assume that there exists a constant c_0 such that $(\mathbb{E}[|x(s,t)|^q])^{1/q} \leq c_0|t-s|^\beta$ and $(\mathbb{E}[(\sqrt{|\mathbb{X}(s,t)|})^q])^{1/q} \leq c_0|t-s|^\beta$. Then there is a version of $\mathbf{x} = (x, \mathbb{X})$ such that

$$\mathbb{E}\left[\left(\sup_{s\neq t} \frac{x(s,t)}{|t-s|^{\alpha-1/q}}\right)^q + \left(\sup_{s\neq t} \frac{\sqrt{|\mathbb{X}(s,t)|}}{|t-s|^{\alpha-1/q}}\right)^q\right] < \infty,$$

provided that $\alpha < \beta$.

Proof. Without loss of generality, assume $T = 1$. Take a dyadic approximation of $[0,1]$: set $D_n = \{j/2^n : 0, 1, \ldots, 2^n\}$, and let $D = \bigcup_{n=1}^\infty D_n$, which is dense in $[0,1]$. Set

$$A_n = \sup_{t\in D_n} |x(t+2^{-n}) - x(t)| = \sup_{t\in D_n} |x(t, t+2^{-n})|, \qquad B_n = \sup_{t\in D_n} |\mathbb{X}(t, t+2^{-n})|$$

Let $s, t \in D$ with $s < t$, and pick m so that $1/2^{m+1} < |s-t| \leq 1/2^m$. Pick $\theta \in [s,t] \cap D_m$, which exists because $|s-t| \geq 1/2^m$. Then

$$|x(t) - x(s)| \leq |x(t) - x(\theta)| + |x(\theta) - x(s)|.$$

Now write the dyadic expansion $t - \theta = \frac{a_0}{2^m} + \frac{a_1}{2^{m+1}} + \cdots$, so $|x(t) - x(\theta)| \leq \sum_{n \geq m} A_n$. Doing the same with the second term,

$$\leq 2 \sum_{n \geq m} A_n$$

Hence,

$$\begin{aligned} \frac{|x(t) - x(s)|}{|t-s|^\gamma} &\leq |x(t) - x(s)| 2^{(m+1)\gamma} \\ &\leq 2^{\gamma+1} \sum_{n \geq m} A_n 2^{m\gamma} \\ &\leq 2^{\gamma+1} \sum_{n \geq m} A_n 2^{n\gamma}. \end{aligned}$$

So we get the bound

$$\sup \frac{|x(t) - x(s)|}{|t-s|^\gamma} \leq 2^{\gamma+1} \sum_{n=0}^{\infty} A_n 2^{n\gamma}.$$

We want to get a bound on the L^q norm of this:

$$\left(\mathbb{E} \left[\left(\sup \frac{|x(t) - x(s)|}{|t-s|^\gamma} \right)^q \right] \right)^{1/q} \leq 2^{\gamma+1} \sum_n (\mathbb{E}[A_n^q])^{1/q} 2^{n\gamma}.$$

On the other hand,

$$A_n^q = \sup_{t \in D_n} |x(t + 2^{-n}) - x(t)|^q \leq \sum_{t \in D_n} |x(t + 2^{-n}) - x(t)|^q,$$

and taking expectations gives

$$\begin{aligned} \mathbb{E}[A_n^q] &\leq \sum_{t \in D_n} \mathbb{E}[|x(t + 2^{-n}) - x(t)|^q] \\ &\leq c_0^q 2^n 2^{-n\beta q}. \end{aligned}$$

This gives the L^q norm bound

$$(\mathbb{E}[A_n^q])^{1/q} \leq c_0 2^{-n(\beta - 1/q)}.$$

Hence,

$$\left(\mathbb{E}\left[\left(\sup \frac{|x(t)-x(s)|}{|t-s|^\gamma}\right)^q\right]\right)^{1/q} \leq c_0 2^{\gamma+1} \sum_n 2^{-n(\beta-1/q-\gamma)} < \infty$$

if $\gamma < \beta - 1/q$.

As for $\mathbb{X}(s,t)$, we do likewise. Let s, t, θ be as above and use

$$\mathbb{X}(s,t) = \mathbb{X}(s,\theta) + \mathbb{X}(\theta,t) + x(s,\theta) \otimes x(\theta,t).$$

We get

$$|\mathbb{X}(s,t)| \leq 2^{\gamma+1} \sum_n B_n 2^{n\gamma} + \left(\sum_n A_n e^{n\gamma}\right)^2,$$

and we can repeat the above argument.

This would give us the regularity of x (resp. $\mathbb{X}$) restricted to D (resp. D^2). Then set $\widetilde{x}(t) = \lim_{\substack{t_n \to t \\ t_n \in D}} x(t_n)$, and we can show that $x = \widetilde{x}$ almost surely:

$$\begin{aligned}
\mathbb{E}[|x(t) - \widetilde{x}(t)|] &= \mathbb{E}\left[\lim_{n\to\infty} |x(t) - x(t_n)|\right] \\
&\leq \underbrace{\liminf \mathbb{E}[|x(t) - x(t_n)|]}_{\leq c_0 |t-t_n|^\beta} \\
&= 0.
\end{aligned}$$

□

10.2 Candidates for the lift of Brownian motion

We now offer two candidates for the lift of an ℓ-dimensional Brownian motion, namely Itô and Stratanovich. Define

$$\mathbb{B}^{\text{Itô}}(s,t) = A(s,t) - B(s)(B(t) - B(s)),$$

with

$$A(s,t) = \lim_{n\to\infty} \sum_{t_i \text{ dyadic in } [s,t]} B(t_i)(B(t_{i+1}) - B(t)).$$

Define the Stratanovich integral similarly except with

$$A^{\text{Strat}}(s,t) = \lim_{n\to\infty} \sum_{t_i \text{ dyadic in } [s,t]} \frac{B(t_i) + B(t_{i+1})}{2}(B(t_{i+1}) - B(t)).$$

For the sake of definiteness, assume $s = 0$. For diagonal terms, we have

$$A^{\text{Itô}}_{r,r} = \lim_{n\to\infty} \sum_{\{t_i\} = D_n} B_r(t_i)(B_r(t_{i+1}) - B_r(t_i)),$$

$$A_{r,r}^{\text{Strat}} = \lim_{n\to\infty} \sum_{\{t_i\}=D_n} \frac{B_r(t_i)+B_r(t_{i+1})}{2}(B_r(t_{i+1})-B_r(t_i)) = \frac{B(t)^2-B(s)^2}{2}.$$

Observe that

$$(A_{r,r}^{\text{Strat}} - A_{r,r}^{\text{Itô}})(s,t) = \lim \sum_i \frac{1}{2}(B_r(t_{i+1})-B_r(t_i))^2 = \frac{t-s}{2},$$

where the last step is a theorem of Lévy. (The proof is to show that $\mathbb{E}[\sum(B_r(t_{i+1}) - B_r(t_i))^2 - (t_{i+1}-t_i)]^2 \to 0$ as $n\to\infty$.) Hence,

$$A_{r,r}^{\text{Itô}}(s,t) = \frac{B(t)^2-B(s)^2-(t-s)}{2}.$$

It remains to evaluate $A_{r,r'}/A_{r,r'}^{\text{Strat}}$. Basically, we have 2 independent, one dimensional standard Brownian motions, say B and B', and we want to calculate $\lim\sum_i B'(t_i)(B(t_{i+1})-B(t_i))$. Let

$$\mathbb{B}_n(t) = \sum_{i=0}^{\lfloor t2^n\rfloor -1} B'(t_i)B(t_i,t_{i+1}).$$

First assume $t=1$, and let us examine

$$\mathbb{B}_{n+1}-\mathbb{B}_n = \sum_i (B'(t_i)B(t_i s_i) + B'(s_i)B(s_i,t_{i+1}) - B'(t_i)B(t_i,t_{i+1})),$$

where s_i is the midpoint of $[t_i,t_{i+1}]$.

$$= \sum_i B'(t_i,s_i)B(s_i,t_{i+1}).$$

So

$$\begin{aligned}
\mathbb{E}[(\mathbb{B}_{n+1}-\mathbb{B}_n)^2] &= \sum_i \mathbb{E}[B'(t_i,s_i)^2 B(s_i,t_{i+1})^2] \\
&= \sum_i 2^{-2(n+1)} \\
&= 2^{-n-2}.
\end{aligned}$$

Hence, $\mathbb{B}_n$ is Cauchy in L^2.

It turns out that $\mathbb{B}_n$ as a function of time is a martingale, and we can take advantage of this to have a better convergence. First, we set $\mathcal{F}_t$ to be the σ-algebra generated by $(B(s): s\in[0,t])$, and we say $t\mapsto M(t)$ is a martingale if $\mathbb{E}[M(t)\mid \mathcal{F}(s)] = M(s)$ for $s<t$. Using Doob's inequality, we can have convergence that is uniform in t:

$$\left(\mathbb{E}\left[\left|\sup_{t\in[0,T]} M(t)\right|^p\right]\right)^{1/p} \le \frac{p}{p-1}\,\mathbb{E}[|M(T)|^p], \qquad p>1.$$

11 Gaussian Inequalities and Markov Techniques for Lifts of Brownian Motion

11.1 Gaussian-type inequalities

For many stochastic processes of interest, we either can use the Markov property or take advantage of the Gaussian distribution of the realizations. In the former case, many martingales become available, and in the latter case, many Gaussian-type inequalities can be used.

For example, if $x : [0, T] \to \mathbb{R}^d$ is a Gaussian process that is centered (i.e. $\mathbb{E}[x(t)] = 0$ for all t), the process is determined by its correlation, $\mathbb{E}[x(t) \otimes x(s)] = R(s, t)$. For simplicity, let us assume that $x = (x_1, \ldots, x_d)$ with x_i, x_j independent for $i \neq j$. Then $R(s, t)$ is diagonal.

Example 11.1. If X_i have the same law for $i = 1, \ldots, d$, then $R(s, t) = C(s, t)I$, where C is scalar-valued, and I is the identity matrix. Also,

$$\mathbb{E}[|x_i(t) - x_i(s)|^2] = C(t, t) + C(s, s) - 2C(s, t),$$

and if

$$\mathbb{E}[|x_i(t) - x_i(s)|^2] \leq c_0|t - s|^{2\alpha},$$

then we can use Kolmogorov's continuity theorem to assert that $x \in \mathcal{C}^\beta$ for every $\beta < \alpha$. Indeed, this estimate would imply that

$$\begin{aligned}(\mathbb{E}[|x_i(t) - x_i(s)|^{2n}])^{1/2n} &\leq a_n(\mathbb{E}[|x_i(t) - x_i(s)|^2])^{1/2} \\ &\leq \sqrt{c_0}a_n|t - s|^\alpha,\end{aligned}$$

and we can use Kolmogorov's continuity theorem to obtain control on $[x_i]_{\alpha-1/(2n)-\varepsilon}$; this holds for any n. To see this, observe that if X is normal with mean 0 and $\mathbb{E}[X^2] = A$, then $\mathbb{E}[e^{tX}] = e^{\frac{t^2}{2}A}$, so that

$$\mathbb{E}[X^{2n}] = \frac{(2n)!}{n!2^n}(\mathbb{E}[X^2])^n.$$

The moral is that in the Gaussian case, we can bound higher moments in terms of the second moment. The good news is that something similar is also true for martingales.

Rough path theory can be carried out for any Gaussian process, provided that $\mathbb{E}[|x_i(t) - x_i(s)|^2] \leq c_0|t - s|^{2\alpha}$ for some $\alpha > 0$. For example, we can consider a fractional Brownian motion that is specified by the requirement that $\mathbb{E}[|x_i(t) - x_i(s)|^2] = |t - s|^{2H}$, where H is known as the **Hurst index**.

11.2 Brownian motion as a Markov process

How about the Brownian motion as a Markov process? Let $B = (B_1, \ldots, B_d)$, where the B_is are independent standard Brownian motion. As we discussed last time, we can come up with a candidate for

$$\int_s^t B_i\,dB_j = \lim_{n\to\infty} \sum_{k:t_k^n\in[s,t]} B_i(t_k^n) B_j(t_k^n, t_{k+1}^n), \quad \text{where} \quad D_n = \{t_k^n = k/2^n : k \in \mathbb{Z}\}.$$

This is in $L^2(\mathbb{P})$, where $\mathbb{P}$ is **Wiener measure**, a probability measure on $C([0,T];\mathbb{R}^d)$. We had another candidate that we will denote

$$\int_s^t B_i \circ dB_j := \lim_{n\to\infty} \sum_{k:t_k^n\in[s,t]} \frac{B_i(t_k^n) + B_i(t_{k+1}^n)}{2} B_k(t_k^n, t_{k+1}^n).$$

For diagonal terms, we have explicit formulae, namely

$$\int_s^t B_i\,dB_i = \frac{B_i(t)^2 - B_i(s)^2}{2} - \frac{t-s}{2}, \qquad \int_s^t B_i \circ dB_i = \frac{B_i(t)^2 - B_i(s)^2}{2}.$$

Though when $i \neq j$, we have $\int_s^t B_i\,dB_j = \int_s^t B_i \circ dB_j$ because

$$\int_s^t B_i \circ dB_j - \int_s^t B_i\,dB_j = \lim_{n\to\infty} \frac{1}{2} \sum_{t_k^n\in[s,t]} B_i(t_k^n, t_{k+1}^n) B_j(t_k^n, t_{k+1}^n),$$

and

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{t_k^n\in[s,t]} B_i(t_k^n, t_{k+1}^n) B_j(t_k^n, t_{k+1}^n)\right)^2\right] &= \sum_{t_k^n\in[s,t]} \mathbb{E}[B_i(t_k^n, t_{k+1}^n)^2]\,\mathbb{E}[B_j(t_k^n, t_{k+1}^n)^2] \\ &\approx 2^n(t-s)2^{-n}2^{-n} \\ &\to 0. \end{aligned}$$

In summary,

$$\mathbb{B}^{\text{Itô}}(s,t) = \mathbb{B}^{\text{Strat}}(s,t) - \frac{1}{2}(t-s)I,$$

where I is the identity matrix.

However, we need to show that $(B, \mathbb{B}^{\text{Itô}}) \in \mathscr{R}^\alpha$ for any $\alpha < 1/2$. We have done with the B part. We get our estimate for $\mathbb{B}^{\text{Itô}}$ using the fact that $M_{i,j}(t) = \int_0^t B_i\,dB_j$ is a martingale. We write $\mathcal{F}_t$ for the σ-algebra generated by $(B(\theta) : \theta \in [0,t])$. Then $M(t)$ is a **martingale** if $\mathbb{E}[M(t) \mid \mathcal{F}_s] = M(s)$, or $\mathbb{E}[M(t) - M(s) \mid \mathcal{F}_s] = 0$.

For example, $B(t)$ itself is a martingale, and observe that $\mathbb{E}[\int_s^t B_i \, dB_j \mid \mathcal{F}_s] = 0$. Indeed,

$$\begin{aligned}
\mathbb{E}\left[\sum_{k/2^n \in [s,t]} B_i(t_k^n) B_k(t_k^n, t_{k+1}^n) \mid \mathcal{F}_s\right] &= \mathbb{E}\left[\sum_{k/2^n \in [s,t]} B_i(s) B_j(t_k^n, t_{k+1}^n) \mid \mathcal{F}_s\right] \\
&\approx \mathbb{E}[B_i(s) B_j(s,t)] \\
&= 0.
\end{aligned}$$

First, we can show that

$$\mathbb{E}[M_{i,j}(t)^2] = \mathbb{E}\left[\int_0^t B_i(s)^2 \, ds\right],$$

which yields

$$\begin{aligned}
\mathbb{E}[(B_{i,j}^{\text{Itô}})^2] &= \mathbb{E}\left[\left(\int_s^t (B_i(\theta) - B_i(s)) \, dB_j(\theta)\right)^2\right] \\
&= \mathbb{E}\left[\int_s^t (B_i(\theta) - B_i(s))^2 \, d\theta\right] \\
&= \int_s^t (\theta - s) \, d\theta \\
&= \frac{(t-s)^2}{2}.
\end{aligned}$$

Here, if we write $A_{i,j}(t) = \int_0^t B_i(\theta)^2 \, d\theta$, then $M_{i,j}(t)^2 - A_{i,j}(t)$ is again a martingale.[4] We have the following fundamental inequality in this context that is due to Burkholder-Davis-Gundy (Doob's inequality):

Lemma 11.1. *If M and $M^2 - \langle M \rangle = M^2 - [M] = M^2 - A$ are martingales with $M(0) = 0$, define $M^*(t) = \sup_{s \in [0,t]} |M(s)|$. Then*

$$\mathbb{E}[M^*(t)^q] \leq c_q \, \mathbb{E}[A^{q/2}].$$

Now, for our example,

$$\begin{aligned}
\mathbb{E}[|\mathbb{B}^{\text{Itô}}(s,t)|^q] &\leq c \, \mathbb{E}\left[\left|\int_s^t B_i(s,\theta)^2 \, d\theta\right|^{q/2}\right] \\
&\leq c \, \mathbb{E}\left[\left(\sup_{\theta \in [s,t]} |B_i(s,\theta)|\right)^q\right] |t-s|^{q/2}
\end{aligned}$$

[4]This is not a coincidence. For any such martingale, if we square it, there is a monotone function we can subtract to get another martingale.

$$\leq c'|t-s|^{\alpha q}|t-s|^{q/2}.$$

So

$$(\mathbb{E}[|\mathbb{B}^{\text{Itô}}(s,t)|^q])^{1/q} \leq c|t-s|^{\alpha+1/2}.$$

12 Exponential Martingale Bounds and Geometricity of the Stratonovich Integral

12.1 Exponential martingale methods for bounding Brownian motion increments

Our purpose is showing that our candidates

$$\mathbb{B}(s,t) = \lim_{n\to\infty} \underbrace{\sum_{t_i^n \in [s,t]} B(t_i^n) \otimes B(t_i^n, t_{i+1}^n)}_{\mathbb{B}_n(s,t)} - B(s) \otimes B(s,t),$$

$$\widehat{\mathbb{B}}(s,t) = \lim_{n\to\infty} \underbrace{\sum_{t_i^n \in [s,t]} \frac{B(t_i^n) + B(t_{i+1}^n)}{2} \otimes B(t_i^n, t_{i+1}^n)}_{\widehat{\mathbb{B}}_n(s,t)} - B(s) \otimes B(s,t)$$

Last time, we worked out the "quadratic variation" of $\mathbb{B}_n$ and applied the Burkholder-Davis-Gundy inequality to get the desired bound. Alternatively, we can use the so-called exponential martingale to get our bounds. The philosophy is that if we have a martingale $M(t)$ and we want a bound, we need to control a modulus of continuity $\sup_{s\neq t, |s-t|<\delta} |M(t) - M(s)|$. Recall that if X is a centered Gaussian, $\mathbb{E}[e^{\lambda X}] = e^{(\lambda^2/2)\,\mathbb{E}[X^2]}$.

Proposition 12.1. *If we set $X_i = B(t_i^n) - B(s)$, then*

$$\mathbb{E}\left[\exp\left(\lambda \sum_{i=k}^{r-1} X_i B(t_i^n, t_{i+1}^n) - \frac{\lambda^2}{2} X_i^2 (t_{i+1}^n - t_i^n)\right)\right] = 1.$$

Proof.

$$\text{LHS} = \mathbb{E}\left[\exp\left(\lambda \sum_{i=k}^{r-2} X_i B(t_i^n, t_{i+1}^n) - \frac{\lambda^2}{2} X_i^2 (t_{i+1}^n - t_i)\right) e^{\lambda X_{r-1} B(t_{r-1}^n, t_r^n) - \frac{\lambda^2}{2} X_{r-1}^2 (t_r^n - r_{r-1}^n)}\right]$$

Condition on the past up to time t_{r-1}^n. The term on the right just becomes 1 because $B(t_{r-1}^n, t_r^n)$ is the only randomness.

$$= \mathbb{E}\left[\exp\left(\lambda \sum_{i=k}^{r-2} X_i B(t_i^n, t_{i+1}^n) - \frac{\lambda^2}{2} X_i^2 (t_{i+1}^n - t_i)\right)\right]$$

We can do the same thing, picking off one term at a time

$$= \cdots$$

$$= 1.$$

□

Remark 12.1. We may write this as

$$\mathbb{E}[e^{\lambda M_n - \frac{\lambda^2}{2} Z_n}] = 1,$$

where M_n is a martingale, and Z_n is the quadratic variation of M_n.

We wish to expand this expression in λ:

$$1 = \mathbb{E}\left[\sum_{m=0}^{\infty} K_m(M_n, Z_n)\frac{\lambda^m}{m!}\right].$$

From this we want to deduce that $K_0 = 1$ and $\mathbb{E}[K_m(M_n, Z_n)] = 0$ for all $m \geq 1$. This gives nice control on M_n in terms of its quadratic variation Z_n. Indeed, use the expansion:

$$e^{tx - \frac{t^2}{2}} = \sum_{m=0}^{\infty} (\mathrm{He})_m(x)\frac{t^m}{m!},$$

Where $(\mathrm{He})_m(x)$ is the m-th Hermite polynomial. Hermite polynomials satisfy the recursive identity $(\mathrm{He})_{m+1}(x) = x(\mathrm{He})_m(x) - m(\mathrm{He})_{m-1}(x)$. We also have $(\mathrm{He})_m(x) = 1$ and $(\mathrm{He})_1(x) = x$, so it is possible to show that $(\mathrm{He})_m(0) = 0$ if m is odd. We can also show that $(\mathrm{He})_m$ has even powers if m is even and odd powers if m is odd. Moreover, we have the expansion (setting $t = \lambda\sqrt{Z}$ and $x = \frac{M}{\sqrt{Z}}$)

$$e^{\lambda M - \frac{\lambda^2}{2} Z} = \sum_{m=0}^{\infty} K_m(M, Z)\frac{\lambda^m}{m!}, \qquad K_m(M, Z) = (\mathrm{He})_m\left(\frac{M}{\sqrt{Z}}\right)(\sqrt{Z})^m.$$

Observe that

$$K_{2m}(M, Z) = M^{2m} + c_1^m M^{2m-2} Z + \cdots + c_{m-1}^m M^2 Z^{m-1} + c_m^m Z^m.$$

From this an $\mathbb{E}[K_{2m}(M, Z)] = 0$, we learn that

$$\mathbb{E}[M^{2m}] \leq -\sum_{i=1}^{m} c_i^m\, \mathbb{E}[M^{2m-2i} Z^i].$$

Let's Schwarz this![5] Use the weighted Schwarz inequality, $ab \leq \frac{(\varepsilon a)^p}{p} + \frac{(b/\varepsilon)^q}{q}$ to write

$$\begin{aligned}\mathbb{E}[M^{2m-2i} Z^i] &\leq \frac{2m-2i}{2m}(\varepsilon M^{2m-2i})^{2m/(2m-2i)} + (Z^i/\varepsilon)^{m/i}\frac{i}{m} \\ &= \left(1 - \frac{i}{m}\right)\varepsilon^{m/(m-i)} M^{2m} + \left(\frac{1}{\varepsilon}\right)^{m/i}\frac{i}{m} Z^m.\end{aligned}$$

[5] Maybe we shouldn't be using Schwarz as a verb, but this is how verbs are created.

From this, we deduce

$$\mathbb{E}[M^{2m}] \le c_m\, \mathbb{E}[Z^m].$$

In summary, if

$$M = M_n = \sum_{t_i^n \in [s,t]} (B_j(t_i^n) - B_j(s)) B_k(t_i^n, t_{i+1}^n), \qquad B = (B_1, \dots, B_\ell),$$

then

$$\mathbb{E}[M_n] \le c_m\, \mathbb{E}[Z_n^m],$$

where

$$Z_n = \sum_{t_i^n} B_j(s, t_i^n)^2 (t_{i+1}^n - t_i^n).$$

Recall that if $\alpha \in (0, 1/2)$ and if

$$C(B) = \sup_{\substack{s \neq t \\ s,t \in [0,T]}} \frac{|B(s,t)|}{|t-s|^\alpha},$$

then $\mathbb{E}[C(B)^q] < \infty$ for every $q \ge 1$ (and in fact even $\mathbb{E}[e^{c_0 C(B)}] < \infty$). Then

$$\mathbb{E}[Z_n^m] \le \mathbb{E}[C(B)^m |t-s|^{2\alpha m + m}] \le c'_m |t-s|^{2\alpha m + m}.$$

As a result,

$$\left(\mathbb{E}[M_n^{2m}]\right)^{1/(4m)} \le c'_m c_m |t-s|^{(2\alpha+1)/4}.$$

In other words,

$$\|\sqrt{M_n}\|_{L^{4m}(\mathbb{P})} \le c|t-s|^{(2\alpha+1)/4},$$

and by Kolmogorov's theorem,

$$\mathbb{E}\left[\sup_{\substack{s \neq t \\ s,t \in [0,T]}} \frac{|\sqrt{M}_n(s,t)|}{|t-s|^\gamma}\right] < \infty,$$

provided that $\gamma \in (0, \frac{2\alpha+1}{4} - \frac{1}{4m})$. By choosing m large and α close to $1/2$, we can get any $\gamma \in (0, 1/2)$. Thus, we do have a rough path $(B, \mathbb{B})$ in $\mathscr{R}^\gamma$ with $\gamma \in (0, 1/2)$. Since $\widehat{\mathbb{B}}(s,t) = \mathbb{B}(s,t) - \frac{t-s}{2} I$, the same is true for $\widehat{\mathbb{B}}$.

12.2 Geometricity of the Stratonovich lift

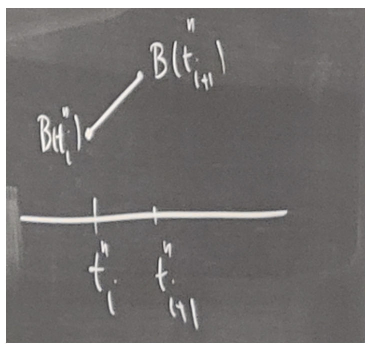
We now claim that $\widehat{\mathbb{B}}$ is geometric and that a smooth approximation of B would lead to the Stratonovich integration. Recall that we want to solve an equation like $\dot{y} = b(y) + \sigma(y)\dot{B}$; we have two candidates for the integrals in the corresponding integral equation, as well. If we replace B by a smooth approximation $B_\varepsilon \xrightarrow{\varepsilon\to 0} B$, then we can solve the equation $\dot{y}_\varepsilon = b_\varepsilon(y) + \sigma_\varepsilon(y)\dot{B}_\varepsilon$ classically. Then $\lim_{\varepsilon\to 0} y_\varepsilon = y$, so

$$\dot{y} = b(y) + \sigma(y)\frac{d}{dt}\widehat{B}.$$

Thus, it will be the Stratonovich integral, not the Itô integral. Note that the regularization should be independent of the path.

To explain this, let us observe that if B is a Brownian motion and $B^{(n)}$ is the linear interpolation

$$B^{(n)}(t) = \sum_{i=0}^{\infty} \mathbb{1}_{[t_i^n, t_{i+1}^n]}(t) \cdot \left[\frac{t - t_i^n}{t_{i+1}^n - t_i^n} B(t_{i+1}^n) + \frac{t_{i+1}^n - t}{t_{i+1}^n - t_i^n} B(t_i^n)\right],$$



then

$$\begin{aligned}
\int_s^t B^{(n)}(\theta) \otimes dB^{(n)}(\theta) &= \sum_{t_i^n \in [s,t]} \int_{t_i^n}^{t_{i+1}^n} B^{(n)}(\theta) \otimes dB^{(n)}(\theta) \\
&= \sum_{t_i^n \in [s,t]} (t_{i+1}^n - t_i^n) \frac{B(t_i^n) + B(t_{i+1}^n)}{2} \otimes \frac{B(t_i^n, t_{i+1}^n)}{t_{i+1}^n - t_i^n} \\
&= \text{Stratonovich approximation.}
\end{aligned}$$

So for $\alpha \in (0, 1/2)$,

$$\left(B^{(n)}, \int B^n \otimes dB^{(n)}\right) \xrightarrow{\mathscr{R}^\alpha} (B, \widehat{\mathbb{B}})$$

because we already know the L^2-convergence, and we have established a uniform bound on $\mathscr{R}^\alpha$ of the approximation. Hence, we have convergence in $\mathscr{R}^\beta$ for $\beta < \alpha$.

Remark 12.2. We can have the following probabilistic interpretation for our approximation that offers another proof of the L^2-convergence. Namely, if $\mathcal{F}_n$ is the σ-algebra generated by $(B(t_i^n) : i = 0, 1, 2, \dots)$, then $B^{(n)} = \mathbb{E}[B \mid \mathcal{F}_n]$. Then $B^{(n)} \to B$ follows from the celebrated Doob's martingale convergence theorem.

13 Making the Jump From Stochastic ODEs to PDEs

13.1 Main results thus far for solving stochastic ODEs

Here are two main results that we have established so far:

1. The "ODE"
$$\begin{cases} \dot{y} = \sigma(y)\dot{x} \\ y(0) = y^0 \end{cases}$$
has a solution that is stable with respect to its input, provided we use the rough-path interpretation for the integrals:
$$y(t) = y^0 + \int_0^t (\sigma(y), \widehat{\sigma})\, d(x, \mathbb{X})$$
with $\widehat{\sigma} = D\sigma(y)\sigma(y)$.

   Moreover, y is a fixed point of the operator
$$\mathscr{I}(\mathbf{y}, \mathbf{x}) = y^0 + \int_0^t (\sigma(y) D\sigma(y)\widehat{y})\, d\mathbf{x}.$$
Using our bounds for the integral, the operator $\mathscr{I}$ is bounded linear in $\mathbf{x}$ and locally Lipschitz in $\mathbf{y}$, and we learn that the solution $X(y^0, \mathbf{x})$ is continuous.

2. If B denotes the standard Brownian motion, then we have two rather natural candidates for its (random) lift, namely $(B, \mathbb{B})$ (Itô) and $(B, \widehat{\mathbb{B}})$ (Stratonovich) in $\mathscr{R}^\alpha$ for any $\alpha \in (0, 1/2)$. Note that our candidate $(B(\cdot), \mathbb{B}(\cdot, \cdot\,; B))$ is in $L^2(\mathbb{P})$ with $\mathbb{P}$ representing the Wiener measure, though $\mathbb{B}$ as a function of B is only measurable.

   In particular, we may approximate B by some nice function, say $B^{(n)}$, and solve
$$\begin{cases} \dot{y} = \sigma(y)\dot{B}^{(n)} \\ y(0) = y^0. \end{cases}$$
Then $\lim_{n\to\infty} y_n = y$, where y solves
$$\dot{y} = \sigma(y)\dot{\widehat{\mathbb{B}}}.$$
Indeed, if for $B^{(n)}$, we choose the linear interpolation of B using dyadic points $D_n = \{i/2^n : i \in \mathbb{Z}\}$ and consider $(B^{(n)}, \mathbb{B}^{(n)})$ by
$$\widehat{\mathbb{B}}^{(n)}(s,t) = \int_s^t B^{(n)} \otimes \dot{B}^{(n)}(\theta)\, d\theta,$$

then as we discussed last time, $\mathbb{B}^{(n)}(s,t)$ is simply the Stratonovich approximation. Hence, in the L^2 sense, $\mathbb{B}^{(n)} \to \widehat{\mathbb{B}}$.

We also know that $\sup_n \|[B^{(n)}, \widehat{\mathbb{B}}^{(n)}]_{\alpha,2\alpha}\|_{L^q(\mathbb{P})} < \infty$. As a result, if we define $\mathbf{B}^{(n)} = (B^{(n)}, \widehat{\mathbb{B}}^{(n)})$ and $\widehat{\mathbf{B}} = (B, \widehat{\mathbb{B}})$ and regard it as a function B, we can show that for $\mathbb{P}$-almost all choices of B,

$$d_\alpha(\mathbf{B}^{(n)}, \widehat{\mathbb{B}}) \to 0,$$

where d_α is the distance with respect to $[\cdot]_{\alpha,2\alpha}$.

In summary, we managed to do Stochastic calculus in two steps:

$$B \xrightarrow{\text{measurable}} (B, \mathbb{C}) \xrightarrow{\text{continuous}} \text{“}\dot{y} = \sigma(y)\tfrac{d}{dt}(B, \mathbb{B})\text{”}$$

Now we want to carry out the program for PDEs.

13.2 Preliminaries for Stochastic PDEs

We start with some notation. We have $\varphi : \mathbb{R}^d \to \mathbb{R}$ or $\varphi : D \to \mathbb{R}$ with some open subset $D \subseteq \mathbb{R}^d$. We will use

$$\|\varphi\|_{L^\infty} = \|\varphi\|_\infty = \sup_x |\varphi(x)|, \qquad \|\varphi\|_{L^\infty(D)} = \sup_{x \in D} |\varphi(x)|$$

to denote the L^∞ norm on $\mathbb{R}^d$ and D, respectively. Given $k = (k_1, \dots, k_d) \in \mathbb{N}_0^d$, we define

$$\partial^k(\varphi) = \partial_{x_d}^{k_d} \cdots \partial_{x_1}^{k_1} \varphi, \qquad |k| = k_1 + \cdots + k_d.$$

We write C^r for the set of functions φ for which ∂^k exists and is continuous for any k with $|k| \le r$. And

$$\|\varphi\|_{C^r} = \sum_{|k| \le r} \|\partial^k \varphi\|_{L^\infty}.$$

We write $\mathcal{D}$ for the set of smooth functions of compact support, and if K is a compact subset of $\mathbb{R}^d$, then $\mathcal{D}(K)$ means the set of $\varphi \in \mathcal{D}$ with $\operatorname{supp}\varphi \subseteq K$. By $\mathcal{D}'$, we mean the set of linear functionals $T : \mathcal{D} \to \mathbb{R}$ which are linear and satisfy

$$|T(\varphi)| \le c_K \|\varphi\|_{C^{r_K}}$$

for some constant c_K and index r_K for every $\varphi \in \mathcal{D}(K)$. Here, r_K is called the **order** of the distribution.

Example 13.1. A 0-th order distribution would be a measure by the Riesz representation theorem.

Next, we wish to discuss $\mathcal{C}^\alpha(\mathbb{R}^d)$ (or $\mathcal{C}^\alpha_{\text{loc}}(\mathbb{R}^d)$) for $\alpha \in \mathbb{R}$. Given a (test) function $\varphi : \mathbb{R}^d \to \mathbb{R}$, we define

$$\varphi_a^\delta(x) = \delta^{-d}\varphi\left(\frac{x-a}{\delta}\right), \qquad (\varphi^\delta := \varphi_0^\delta, \varphi_a := \varphi_a^\delta).$$

Observe that $\int \varphi_a^\delta = \int \varphi$.

Imagine that $u : \mathbb{R}^d \to \mathbb{R}$ is Hölder of exponent α, and take φ from

$$\mathcal{D}_0 = \left\{\varphi \in \mathcal{D} : \operatorname{supp}\varphi \subseteq B(0,1), \int \varphi \neq 0, \|\varphi\|_{L^\infty} \leq 1\right\}.$$

We will use the bracket notation

$$\langle u - u(a), \varphi_a^\delta\rangle = \int (u - u(a))\varphi_a^\delta\, dx.$$

Taking absolute values and making a change of variables, we can write

$$\begin{aligned}
|\langle u - u(a), \varphi_a^\delta\rangle| &= \left|\int (u - u(a))\varphi_a^\delta\, dx\right| \\
&= \left|\int (u(a+\delta z) - u(a))\varphi(z)\, dz\right| \\
&\leq [u]_\alpha \delta^\alpha \int |z| \cdot |\varphi(z)|\, dz.
\end{aligned}$$

Hence, for $u \in \mathcal{C}^\alpha$ with $\alpha \in (0,1]$,

$$[\![u]\!]_{\mathcal{C}^\alpha} := \sup_{\delta \in (0,1]} \sup_{a \in K} \sup_{\varphi} \frac{|\langle u - u(a), \varphi_a^\delta\rangle|}{\delta^\alpha} \leq c[u]_\alpha,$$

so these norms are equivalent by the following proposition:

Proposition 13.1. *If $[\![u]\!]_{\mathcal{C}^\alpha} < \infty$, then $u \in \mathcal{C}^\alpha$.*

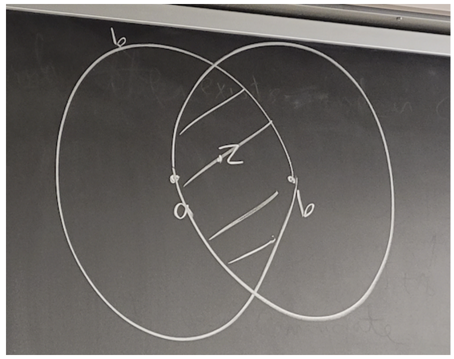
Proof. If $[\![u]\!]_{\mathcal{C}^\alpha} < \infty$,

$$\sup_{a \in K} \delta^{-d} \int_{|z-a|<\delta} |u(z) - u(a)|\, dz \leq c_0 \delta^\alpha.$$

Choose $\delta = |a - b|$ and argue that

$$|u(a) - u(b)| \leq |u(a) - u(z)| + |u(z) - u(b)|$$

for $z \in B(a,\delta) \cap B(b,\delta)$ with $\delta = |a-b|$.



Integrate both sides over $B(a,\delta) \cap B(b,\delta)$ to get

$$
\begin{aligned}
|\underbrace{B(a,\delta) \cap B(b,\delta)}_{B_{a,b}}| \cdot |u(a) - u(b)| &\leq \int_{B_{a,b}} |u(a) - u(z)|\, dz + \int_{B_{a,b}} |u(b) - u(z)|\, dz \\
&\leq \int_{B(a,\delta)} |u(a) - u(z)|\, dz + \int_{B(b,\delta)} |u(b) - u(z)|\, dz \\
&\leq 2c_0 \delta^{\alpha+d}.
\end{aligned}
$$

Hence, $|u(a) - u(b)| \leq c_1 \delta^\alpha$, as desired. □

We want to go beyond $\alpha \in (0,1)$. For example, consider $\alpha \geq 1$. For such α, we first define $n = \max\{m \in \mathbb{N} : m < \alpha\}$. We say $u \in \mathcal{C}^\alpha$ if u has n-many derivatives and if

$$
P_a^u(x) := \sum_{|k| \leq n} (\partial^k u)(a)(x-a)^k, \qquad (x-a)^k := \prod_{i=1}^d (x_i - a_i)^{k_i}, k! := k_1! \cdots k_d!,
$$

then

$$
[\![u]\!]_{\alpha,K} = \sup_{\delta \in (0,1)} \sup_{\varphi \in \mathcal{D}_0} \sup_{a \in K} \frac{\int (u - P_a^u) \varphi_a^\delta \, dx}{\delta^\alpha} < \infty.
$$

One can show that $[\![u]\!]_{\alpha,K} < \infty$ if and only if u possesses n many derivatives and for any k with $|k| = n$, $\partial^k u$ is Hölder of exponent $\alpha - n$.

Basically, we need to choose $\varphi = \partial^k \psi$ for some smooth ψ, and observe that

$$
\|\partial^k \psi\|_{L^\infty} \leq \lambda^{-k-d}.
$$

14 Coherence and Hairer's Reconstruction Theorem

14.1 Examples of coherence

Last time, we discussed $\mathcal{C}^\alpha_{\text{loc}}$ for $\alpha \in (0,1)$ that can be characterized by (if $u \in \mathcal{C}^\alpha_{\text{loc}}$ and K is compact)

$$\sup_{x\in K}\sup_{\delta\in(0,1]}\sup_{\varphi\in\mathcal{D}_0}\frac{|\langle u-u(x),\varphi_x^\delta\rangle|}{\delta^\alpha} =: [u]_{\mathcal{C}^\alpha,K} < \infty.$$

Also, if $\alpha \geq 1$, then we set

$$P_x(y) = P_x^u(y) = \sum_{|k|\leq\alpha}(\partial^k u)(x)\frac{(y-x)^k}{k!},$$

and $u \in \mathcal{C}^\alpha_{\text{loc}}$ means that for K compact,

$$\sup_{x\in K}\sup_{\delta\in(0,1]}\sup_{\varphi\in\mathcal{D}_0}\frac{|\langle u-P_x,\varphi_x^\delta\rangle|}{\delta^\alpha} =: [u]_{\mathcal{C}^\alpha,K} < \infty.$$

For example, if $\alpha \in (1,2)$, then

$$\begin{aligned} u(y) - P_x(y) &= \underbrace{u(y)-u(x)}_{[u(ty+(1-t)x)]_0^1} - Du(x)\cdot(y-x) \\ &= \int_0^1 (Du(ty+(1-t)x) - Du(x))\cdot(y-x)\,dt. \end{aligned}$$

To assert that if $u \in C^1$ and $Du \in \mathcal{C}^{\alpha-1}$, then

$$|u(y)-P_x(y)| \leq c|x-y|^\alpha$$

locally uniformly. Then we can show that the above norm is finite. However, we may use our polynomial approximation expression for our definition of $\mathcal{C}^\alpha_{\text{loc}}$.

Here, we have an example of a function u that is well-approximated by a so-called germ $(P_x : x \in \mathbb{R}^d)$. Indeed, this family enjoys a regularity that we now explore. To find such a regularity, observe

$$P_a(x) = \sum_{|k|<\alpha}\partial^k u(a)\frac{(x-a)^k}{k!},$$

$$\begin{aligned} P_b(x) &= \sum_{|k|<\alpha}\partial^k u(b)\frac{(x-b)^k}{k!} \\ &= \sum_{|k|<\alpha}\left[\sum_{|r|<\alpha-|k|}\partial^{k+r}u(a)\frac{(b-a)^r}{r!} + R_k(a,b)\right]\frac{(x-b)^k}{k!}, \end{aligned}$$

where the error

$$|R_k(a,b)| \lesssim |b-a|^{\alpha-k}.$$

From now on, $f \lesssim g$ mean $f \le cg$ for a constant c. Hence,

$$P_b(x) = \sum_{|m|<\alpha} \frac{\partial^m u(a)}{m!} \underbrace{\left(\sum_{k+r=m} \frac{(b-a)^r}{r!} \frac{(x-b)^k}{k!} m! \right)}_{(x-a)^m} + \sum_{|k|<\alpha} R_k(a,b) \frac{(x-b)^k}{k!}$$

From this, we learn that

$$P_b(x) - P_a(x) = \sum_{|k|<\alpha} R_k(a,b) \frac{(x-b)^k}{k!},$$

and hence

$$\begin{aligned} |\langle P_b - P_a, \varphi_b^\delta \rangle| &\lesssim \sum_{|k|<\alpha} |b-a|^{\alpha-|k|} \delta^{|k|} \\ &\lesssim (\delta + |b-a|)^\alpha. \end{aligned}$$

Here, we have an example of a germ, namely $(P_x : x \in \mathbb{R}^d)$ that is α**-coherent** (which will be defined later).

Let us have another example, namely what we had before in Gubinelli's version (the sewing lemma) of Lyons and Victoire's result: Imagine that we have $A(s,t)$ with

$$|A(s,t) + A(u,t) - A(s,t)| \lesssim |t-s|^{\alpha+\beta}, \qquad s<u<t, \alpha+\beta>1.$$

Then by the sewing lemma, we can find h such that

$$|h(t) - h(s) - A(s,t)| \lesssim |t-s|^{\alpha+\beta}.$$

For example, we may have $A(s,t) = f(s)(g(t)-g(s))$ with $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$. As we stated before, we may consider the germ $(F_s : s \in \mathbb{R})$, where $F_s = f(s)g'$; what the condition A means is this: Observe that $A(s,t) = \langle F_s, \mathbb{1}_{[s,t]} \rangle$. Hence

$$\langle F_u - F_s, \mathbb{1}_{[s,t]} \rangle \lesssim |t-s|^{\alpha+\beta}.$$

If $\varphi = \mathbb{1}_{[0,1]}$,

$$|\langle F_u - F_s, \varphi_s^\delta \rangle| \lesssim \delta^{-1}(|u-s|+\delta)^{\alpha+\beta} = \delta^{-1}(|u-s|+\delta)^{\gamma+1},$$

where $\gamma = \alpha+\beta-1 > 0$. In summary, we have an example of a germ that is γ-coherent, or more specifically $(-1,\gamma)$-coherent.

Motivated by these two examples, we formulate some definitions.

Definition 14.1. By a **germ**, we mean a measurable map $F : \mathbb{R}^d \to \mathcal{D}'$ sending $x \mapsto F_x$.

Definition 14.2. We call a germ $(-\tau, \gamma)$**-coherent** with $\tau = \tau_K$ only depending on a compact set K and with respect to a test function $\phi \in \mathcal{D}$, $\int \varphi \neq 0$, if the following condition is true:

$$\langle F_x - F_y, \varphi_y^\delta \rangle \lesssim \delta^{-\tau_K} (|x-y| + \delta)^{\gamma + \tau_K}$$

uniformly for $x, y \in K$. Here, we assume that $\tau_K \geq 0$ and $\gamma + \tau_K \geq 0$.

We say γ**-coherent** when we mean $(-\tau, \gamma)$-coherent for some τ which does not matter.

14.2 Martin Hairer's reconstruction theorem

Theorem 14.1 (Martin Hairer's reconstruction theorem)**.** *Assume that F is a γ-coherent germ with respect to some $\varphi \in \mathcal{D}$. Then there exists $u \in \mathcal{D}'$ such that*

$$|\langle u - F_x, \psi_x^\delta \rangle| \lesssim \begin{cases} \delta^\gamma & \gamma \neq 0 \\ 1 + |\log \delta| & \gamma = 0. \end{cases}$$

uniformly for $x \in K$ and ψ such that $\operatorname{supp} \psi \subseteq B_1(0)$ *and* $\|\psi\|_{C^r} \leq 1$ *with* $r = r_K$.

Remark 14.1. If $\gamma > 0$ is positive, then the u in the theorem is unique.

Proof. If u and u' satisfy the same inequality, and $T = u - u'$, then $|T(\psi_x^\delta)| \lesssim \delta^\gamma$. Let us take $f \in L^1_{\mathrm{loc}}$ and consider $\zeta \in \mathcal{D}$ and consider $f * \zeta$. Here,

$$(f * \zeta)(x) = \int \zeta(x-y) f(y)\, dy = \int (\tau_y \zeta)(x)\, f(y)\, dy.$$

We claim that

$$T(f * \zeta) = T\left(\int T_y \zeta f(y)\, dy \right) = \int T(\tau_y \zeta) f(y)\, dy.$$

This can be done by Riemann approximation of the integral. Now

$$T(\zeta) = \lim_{\delta \to 0} T(\zeta * \psi^\delta) = \lim_{\delta \to 0} \int T(\tau_y \psi^\delta) \zeta(y)\, dy = 0,$$

as $|T(\psi^\delta)| \lesssim \delta^\gamma$. □

15 Bounds for Germs

15.1 Condition for coherence of germs

Ultimately, we wish to find a distribution $u \in \mathcal{D}'$ that is well-approximated by a germ. Recall that a **germ** is $F : \mathbb{R}^d \to \mathcal{D}'$ that is measurable.

Proposition 15.1. *Let F be a germ, and assume that there exists a constant c, a compact set K, exponents γ and r, and a distribution u such that*

$$|(u - F_x)(\phi_x^\delta)| \le c\delta^\gamma$$

for all $x \in K$, $\delta \in (0, 1]$, and $\phi \in \mathcal{D}$ such that $\operatorname{supp}\phi \subseteq B_1(0) = \{x : |x| \le 1\}$ and $\|\phi\|_{C^r} \le 1$. (Here, $F_x := F(x)$.) Then

$$|(F_x - F_y)(\phi_y^\delta)| \le 2c\delta^{-\tau}(|x - y| + \delta)^{\gamma+\tau},$$

provided that $\delta \in (0, 1/2]$, ϕ is as before, and $|x - y| \le 1/2$. Here, we may choose $\tau = d + r$.

Remark 15.1. This basically says that F is (τ, γ)-coherent. Observe that

$$\delta^{-\tau}(|x - y| + \delta)^{\gamma+\tau} \lesssim \begin{cases} \delta^{-\tau}|x - y|^{\gamma+\tau} = \delta^\gamma(\frac{|x-y|}{\delta})^{\gamma+\tau} & |x - y| > \delta \\ \delta^\gamma & |x - y| < \delta, \end{cases}$$

so this second case is an improvement. Also observe that if F is (τ, γ)-coherent and $\tau \le \tau'$, then it is also (τ', γ)-coherent.

Proof. Observe that we can write

$$\begin{aligned} |(F_x - F_y)(\phi_y^\delta)| &\le |(F_x - u)(\phi_y^\delta)| + |(u - F_y)(\phi_y^\delta)| \\ &\le |(F_x - u)(\phi_y^\delta)| + c \underbrace{\delta^\gamma}_{\le \delta^\tau(|x-y|+\delta)^{\gamma+\tau}}. \end{aligned}$$

It remains to bound $|(u - F_x)(\phi_y^\delta)|$. Note that $x \ne y$ in general. Observe that

$$\begin{aligned} \phi_y^\delta(z) &= \frac{1}{\delta^d}\phi\left(\frac{z - y}{\delta}\right) \\ &= \frac{1}{\delta^d}\phi\left(\frac{(z - x) - (y - x)}{\delta}\right) \\ &= \frac{1}{\delta^d}\phi\left(\frac{z - x - \frac{y-x}{|y-x|+\delta}(|y - x| + \delta)}{(|y - x| + \delta)\frac{\delta}{|y-x|+\delta}}\right) \end{aligned}$$

Denote $\varepsilon := |y-x| + \delta$, $\varepsilon' := \frac{\delta}{|y-x|+\delta}$, and $a := \frac{y-x}{|y-x|+\delta}$.

$$
\begin{aligned}
&= \frac{1}{\varepsilon^d}\frac{1}{(\varepsilon')^d}\phi\left(\frac{\frac{z-x}{\varepsilon} - a}{\varepsilon'}\right) \\
&= \frac{1}{\varepsilon^d}\phi_a^{\varepsilon'}\left(\frac{z-x}{\varepsilon}\right)
\end{aligned}
$$

Denote $\psi := \phi_a^{\varepsilon'}$.

$$
= \psi_\varepsilon^x(z).
$$

Now observe that by definition, our new test function

$$
\psi(z) = (\varepsilon')^{-d}\phi\left(\frac{z-a}{\varepsilon'}\right), \qquad a = \frac{y-x}{|y-x|+\delta},
$$

so let's examine the support of ψ: $\operatorname{supp}\psi \subseteq B_{\varepsilon'+a}(0)$, if $\operatorname{supp}\phi \subseteq B_1(0)$. Note that $\varepsilon' + |a| = \frac{|y-x|}{|y-x|+\delta} + \frac{\delta}{|y-x|+\delta} = 1$.

We can also rephrase the condition of $\|\phi\|_{C^r} \leq 1$ as just giving a factor of $\|\phi\|_{C^r}$ in the inequality in the hypothesis of the theorem.

We can now argue that

$$
\begin{aligned}
|(u - F_x)(\phi_y^\delta)| &= |(u - F_x)(\psi_x^\varepsilon)| \\
&\leq c\varepsilon^\gamma \|\psi\|_{C^r}.
\end{aligned}
$$

On the other hand, $\|\psi\|_{C^r} \leq (\varepsilon')^{-(d+r)}$. Hence,

$$
\begin{aligned}
|(u - F_x)(\phi_y^\delta)| &\leq c(|y-x|+\delta)^\gamma \left(\frac{\delta}{|y-x|+\delta}\right)^{-d-r} \\
&= c\delta^{-\tau}(|y-x|+\delta)^{\gamma+\tau},
\end{aligned}
$$

where $\tau = d + r$, as desired. □

15.2 Uniform bounds on germs

We now address the following question: Assume that

$$
\delta^{-\gamma} \sup_{x\in K} \sup_{\|\phi\|_{C^r}\leq 1} (u - F_x)(\phi_x^\delta) \leq c
$$

for $\delta \in (0, 1]$.

Proposition 15.2. *Suppose $F = (F_x : x \in \mathbb{R}^d)$ is a $(-\tau, \gamma)$-coherent germ with respect to ϕ.*[6] *Then there exists $\eta = \eta_K$ such that $|F_x(\phi_x^\delta)| \lesssim \delta^{-\eta}$ uniformly in a compact set K and uniformly for $\delta \in [0, 1]$.*

[6] Later, we will see that coherence with respect to 1 ϕ implies coherence with respect to other functions.

The important part is that we can choose η independent of x.

Proof. Fix $a \in K$, and observe that

$$|F_a(\phi_x^\delta)| \le c_0 \|\phi_x^\delta\|_{C^r} \le c_1 \delta^{-d-r} \|\phi\|_{C^r}$$

for ϕ such that $\operatorname{supp} \phi_x^\delta$ is in some compact set. This is the case if $\delta \in (0,1]$ and $x \in K$. We now use the coherence to assert that for $x \in K$,

$$\begin{aligned} |F_x(\phi_x^\delta)| &\le |(F_a - F_x)(\phi_x^\delta)| + |F_a(\phi_x^\delta)| \\ &\le c\delta^{-\tau}(\underbrace{|a-x|}_{\operatorname{diam} K} + \delta)^{\gamma+\tau} + c_2 \delta^{-d-r}. \end{aligned}$$

We are done if we choose $\eta = \max\{\tau, d+r\}$. $\square$

15.3 Preparation for proving the reconstruction theorem

We now focus on the proof of the reconstruction theorem of Hairer.[7] As a preparation, we start with a test function ϕ with $\int \phi \neq 0$ and switch to a new test function $\widehat{\phi}$ so that

$$\int \widehat{\phi} = \int \phi, \quad \text{but} \int \widehat{\phi}(x) x^r \, dx = 0 \quad \text{for} \quad 0 < |r| \le \ell - 1.$$

In fact, what we have in mind is

$$\widehat{\phi} = \sum_{i=0}^{\ell-1} c_i \phi^{\lambda_i}, \qquad \phi^\lambda(x) := \lambda^{-d} \phi\left(\frac{x}{\lambda}\right).$$

In other words, given distinct positive $\lambda_0, \ldots, \lambda_{\ell-1}$, we can find $c_0, \ldots_{\ell+1}$ such that for $\widehat{\phi}$ defined this way, the integral conditions hold. Indeed,

$$\int \widehat{\phi} = \sum_{i=0}^{\ell-1} c_i \int \phi$$

$$\begin{aligned} \int \widehat{\phi} x^r \, dx &= \sum_{i=1} c_i \int x^r \phi^{\lambda_i}(x) \, dx \\ &= \sum_{i=0}^{\ell-1} c_i \lambda_i^{|r|} \int x^r \phi(x) \, dx. \end{aligned}$$

[7] Hairer's original proof used wavelets, which we will not use.

So we need

$$\begin{cases} \sum_{i=0}^{\ell-1} c_i = 1, \\ \sum_{i=0}^{\ell-1} c_i \lambda_i^r \text{ for } r = 1, \ldots, \ell - 1, \end{cases} \qquad \text{or} \quad \underbrace{\begin{bmatrix} 1 & \cdots & 1 \\ \lambda_0 & \cdots & \lambda_{\ell-1} \\ \vdots & & \vdots \\ \lambda_0^{\ell-1} & \cdots & \lambda_{\ell-1}^{\ell-1} \end{bmatrix}}_{A} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{\ell-1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

In fact, there is an explicit formula for A^{-1}, and the answer is

$$c_i = \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i}.$$

Note that we may choose the λ_is small enough so that $\operatorname{supp} \widehat{\phi} \subseteq B(0, 1/2)$.

Out of this $\widehat{\phi}$, we now build another test function of the form

$$\widetilde{\phi} = \widehat{\phi}^2 - \widehat{\phi}^{1/2}.$$

We will use this to prove the reconstruction theorem next time.

16 Proof of Hairer's Reconstruction Theorem

16.1 Motivation for multiresolution analysis

We wish to show the following reconstruction theorem of Martin Hairer:

Theorem 16.1. *If F is a γ-coherent germ, then there exists a distribution T such that*

$$|\langle T - F_x, \varphi_x^\delta \rangle| \lesssim \delta^\gamma, \qquad \gamma \neq 0.$$

This is uniform for $\delta \in (0,1]$, $x \in K$, $\operatorname{supp} \varphi \subseteq B_1(0)$, $\|\varphi\|_{C^r} \leq 1$.

Last time, we showed that T is unique if $\gamma > 0$. However, if $\gamma < 0$, then we can add a distribution S to T, provided that

$$|\langle S, \varphi_x^\delta \rangle| \lesssim \delta^\gamma,$$

which means $S \in \mathcal{C}^\gamma$.

To give an idea about the strategy of the proof, we first discuss Hairer's original proof that uses wavelet expansion. In fact, the proof we presented for $d = 1$ uses the wavelet $\mathbb{1}_{[0,1]}$, i.e. the Haar basis. Recall that if $f \in \mathcal{C}^\alpha, g \in \mathcal{C}^\beta$, then

$$\begin{aligned}
\int_s^t f g' \, d\theta &= \underbrace{(fg')}_{T}(\mathbb{1}_{[s,t]}) \\
&\approx \sum_{t_i^n \in [s,t]} f(t_i^n)(g(t_{i+1}^n) - g(t_i^n)) \\
&= \sum_{t_i^n \in [s,t]} f(t_i^n) g'(\mathbb{1}_{[t_i^n, t_{i+1}^n]}) \\
&= \sum_{t_i^n \in [s,t]} F_{t_i^n}(\mathbb{1}_{[t_i^n, t_{i+1}^n]}).
\end{aligned}$$

And we have shown that this converges if $\alpha + \beta > 1$. For our extension, we replace $\mathbb{1}_{[s,t]}$ with $\varphi \in \mathcal{D}$, and the Haar basis may be replaced with a basis using a multiresolution analysis (MRA) of Mallat.

16.2 Multiresolution analysis

Here is a quick review of MRA:

Definition 16.1. We say $\phi \in L^2(\mathbb{R})$ is a **scaling function** or a **(father) wavelet**[8] if the following conditions are true: First, let $\phi_a^n(x) = 2^{n/2} \phi(2^n(x - a))$, where $n \in \mathbb{Z}$, $a \in 2^{-n}\mathbb{Z}$ so that $\|\phi_a^n\|_{L^2} = \|\phi\|_{L^2}$. Also set $V_n = \operatorname{span}\{\phi_a^n : a \in \Lambda_n = 2^{-n}\mathbb{Z}\}$.

[8] There are also mother wavelets.

(i) $V_n \subseteq V_{n+1}$ (it suffices to have $V_0 \subseteq V_1$)

(ii) $\{\phi(\cdot - k) : k \in \mathbb{Z}\}$ is an orthonormal basis for V_0 (hence $\{\phi_a^n : a \in \Lambda_n\}$ is an orthonormal basis for V_n)

(iii) $L^2(\mathbb{R}) = \overline{\bigcup_n V_n}$.

Example 16.1. We can take, for example, $\phi = \mathbb{1}_{[0,1]}$ to get functions of the form $\phi_a^n = \mathbb{1}_{[t_i^n, t_{i+1}^n]}$. Also, $V_0 = \{\phi(\cdot - k) : k \in \mathbb{Z}\}$.

Remark 16.1. It can be proved that there is no such ϕ which is smooth and has compact support. However, if we only require that ϕ has a certain number of derivatives, it is possible to construct one.

Remark 16.2. We may find W_n such that $V_{n+1} = V_n \oplus W_n$ (W_n is the orthogonal complement of V_n inside V_{n+1}).

Proposition 16.1. *There exists ψ such that if $\psi_a^n(x) = 2^{n/2}\psi(2^n(x-a))$, then*

$$W_n = \operatorname{span}\{\psi_a^n : a \in \Lambda_n\}.$$

This ψ is called the ***(mother) wavelet****.*

Remark 16.3. In fact, it suffices to find $\psi \in V_1$ so that ψ is orthogonal to the integer translates of ϕ, and $W_0 = \operatorname{span}\{\psi(\cdot - k) : k \in \mathbb{Z}\}$. Indeed,

$$V_0 \subseteq V_1 \iff \phi(x) = \sqrt{2}\sum_{r\in\mathbb{Z}} a_r \phi(2x - r) \text{ for coefficients } a_r,$$

And ψ is simply given by

$$\psi(x) = \sqrt{2}\sum_{r\in\mathbb{Z}} b_r \phi(2x - r), \qquad b_r = (-1)^r a_{1-r}.$$

Example 16.2. When $\phi = \mathbb{1}_{[0,1]}$, we may take to be 1 on $[0, 1/2]$ and -1 on $[-1/2, 0)$.

Here is the proof of $\psi \perp V_0$:

Proof. Observe that

$$\begin{aligned}
\phi_\ell(x) &= \phi(x - \ell) \\
&= \sum_r a_r(\sqrt{2}\phi(2x - 2\ell - r)) \\
&= \sqrt{2}\sum_r a_{r-2\ell}\phi(2x - r).
\end{aligned}$$

Hence,

$$\begin{aligned}\langle \psi, \phi_\ell \rangle &= \sum_r a_{r-2\ell} b_r \\ &= \sum_r a_{r-2\ell} (-1)^r a_{1-r}\end{aligned}$$

Denote $1 - s = r - 2\ell$

$$= -\sum_s a_{1-s} (-1)^s a_{s-2\ell},$$

which implies that $\langle \psi, \phi_\ell \rangle = 0$. $\square$

Theorem 16.2 (Ingrid Daubechies)**.** *For every k, there exists a scaling function $\phi \in C^k$ of compact support. Moreover, any polynomial of degree k is in V_0.*

16.3 Strategy of Hairer's proof of the reconstruction theorem

Assuming this theorem of Daubechies, we are now ready to describe Hairer's strategy for the proof. Again, we wish to find a distribution T such that $\langle T - F_x, \varphi_x^\delta \rangle \lesssim \delta^\gamma$. Here, is the recipe for constructing T: When $\gamma > 0$, $T = \lim_{n\to\infty} T_n$ (this means for every $\psi \in \mathcal{D}$, $T(\psi) = \lim_{n\to\infty} T_n(\psi) = \lim_{n\to\infty} \int T_n(x)\psi(x)\,dx$), where

$$T_n(x) = \sum_{a\in\Lambda_n} \langle F_a, \phi_a^n \rangle \phi_a^n(x).$$

How about $\gamma < 0$? In this case, the convergence fails. Recall that if $n > 0$,

$$V_n = V_{n-1} \oplus W_{n-1} = V_{n-2} \oplus W_{n-1} \oplus W_{n-2} = \cdots = V_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_{n-1}.$$

Hence, $L^2 = \overline{V_0 \oplus \bigoplus_{n=0}^\infty W_n}$, or more generally,

$$L^2 = \overline{V_m \oplus \bigoplus_{n=m}^{\infty} W_m}.$$

So for any u,

$$u = \sum_{a\in\Lambda_m} \langle u, \phi_a^m \rangle \phi_a^m + \sum_{n=m}^\infty \sum_{a\in\Lambda_m} \langle u, \psi_a^n \rangle \psi_a^n.$$

Our candidate for T is

$$T = \sum_{a\in\Lambda_m} \langle F_a, \phi_a^m \rangle \phi_a^m + \sum_{n=m}^\infty \sum_{a\in\Lambda_m} \langle F_a, \psi_a^n \rangle \psi_a^n.$$

16.4 Proof of the reconstruction theorem without wavelet expansions

We now present a proof that does not use wavelet expansions. We achieve this by using a suitable $\rho \in \mathcal{D}$. If we choose ρ correctly, then

$$T_n = F_x(\widehat{\rho}_x^n), \qquad \text{where } \widehat{\rho}_x^n(y) = 2^{dn}\rho(2^n(x-y)) = \rho_x^{2^{-n}}(y).$$

For $\gamma > 0$, the limit $\lim_n T_n$ will exist, but for $\gamma < 0$, we will throw away a "bad term" which will not matter. We will finish the explanation next time.

17 Proof of Hairer's Reconstruction Theorem Without Using Wavelets

17.1 Scaling and translation of convolutions

Given a γ-coherent germ $(F_x : x \in \mathbb{R}^d)$, we wish to find a distribution T such that

$$\langle T - F_x, \varphi_x^\delta \rangle \lesssim \delta^\gamma,$$

locally uniformly in x. Recall that coherence means

$$\langle F_x - F_y, \varphi_y^\delta \rangle \lesssim \delta^{-\tau}(|x-y| + \delta)^{\gamma+\tau},$$

locally uniformly in x, y. If it is also uniform in φ with $\|\varphi\|_{C^r} \le 1$ and $\operatorname{supp}\varphi \subseteq B_1(0)$, then we can find $\beta \ge 0$ such that

$$\langle F_x, \varphi_x^\delta \rangle \lesssim \delta^{-\beta},$$

locally uniformly.

We now give a proof of the existence of T using a single test function φ with $\int \varphi \neq 0$. Here is the strategy for constructing our T. We choose a suitable $\rho \in \mathcal{D}$ with $\int \rho = 1$ and define $\widehat{\rho}_x^n = \rho_x^{2^{-n}} = 2^{dn}\rho(2^n(x-y))$ (recall that $\psi_x^\delta(y) := \delta^{-d}\psi(\frac{y-x}{\delta})$). We will construct ρ that can be represented as $\rho = \psi * \varphi$ for suitable test function ψ and φ that will be determined later. But for now, let us make some observations.

Proposition 17.1.

$$(\psi * \varphi)^\delta = \psi^\delta * \varphi^\delta.$$

Proof.

$$\begin{aligned}
-\delta^{-d}(\psi * \varphi)(\tfrac{x}{\delta}) &= \delta^{-d}\int \psi(\tfrac{x}{\delta} - z)\varphi(z)\,dz \\
&= \delta^{-2d}\int \psi(\tfrac{x}{\delta} - \tfrac{z}{\delta})\varphi(\tfrac{z}{\delta})\,dz \\
&= \int \psi^\delta(x-z)\varphi^d(z)\,dz.
\end{aligned}$$

□

Proposition 17.2.

$$(\psi * \varphi)_x(\cdot) = \int \psi_z(\cdot)\varphi_x(z)\,dz$$

Proof.

$$\begin{aligned}
(\psi * \varphi)_x(y) &= \psi * \varphi_x(y) \\
&= \int \psi(y-z)\varphi_x(z)\,dz \\
&= \int \psi_z(y)\varphi_x(z)\,dz.
\end{aligned}$$

□

17.2 Construction of T as a limit

In fact, for a carefully selected ρ, we set

$$T_n(x) = F_x(\widehat{\rho}_x^n), \qquad T = \lim_{n\to\infty} T_n,$$

where this limit means $T(\zeta) = \lim_{n\to\infty} \int T_n(x)\zeta(x)\,dx$. For $\gamma > 0$, we show that the limit does exist and satisfies our requirement. For $\gamma < 0$, we need to first get rid of some diverging terms. Again, our ρ takes the form $\rho = \psi * \varphi$ (with ψ and φ to be picked later). To prove our convergence, write

$$T = T_\infty = T_1 + \sum_{n=1}^{\infty}(T_{n+1} - T_n)$$

and show that $|\langle T_{n+1} - T_n, \zeta\rangle| \lesssim 2^{-n\alpha}$ for some $\alpha > 0$. Indeed,

$$\begin{aligned} T_{n+1}(x) - T_n(x) &= F_x(\widehat{\rho}_x^{n+1} - \widehat{\rho}_x^n) \\ &= F_x(\widehat{m}_x^n), \end{aligned}$$

where $m = \rho^{1/2}(y) - \rho(y) = 2^d\rho(2y) - \rho(y)$. Observe that since $\rho = \psi * \varphi$, then

$$m = \rho^{1/2} - \rho = \psi^{1/2} * \varphi^{1/2} - \psi * \varphi.$$

If we chose $\psi = \varphi^2$, then

$$m = \varphi * \varphi^{1/2} - \varphi^2 * \varphi = \varphi * (\varphi^{1/2} - \varphi^2) =: \varphi * \xi.$$

Our goal is bounding $F_x(\widehat{m}_x^n)$. By our propositions,

$$\begin{aligned} F_x(\widehat{m}_x^n) &= F_x\left(\int \widehat{\varphi}_z^n \widehat{\xi}_x^n\,dz\right) \\ &= \int F_x(\widehat{\varphi}_z^n)\widehat{\xi}_z^n(z)\,dz \\ &= A_n + B_n, \end{aligned}$$

where

$$A_n = \int F_z(\widehat{\psi}_z^n)\widehat{\xi}_x^n(z)\,dz, \qquad B_n = \int (F_x - F_z)(\widehat{\psi}_z^n)\widehat{\xi}_x^n(z)\,dz.$$

Given $\xi \in \mathcal{D}$,

$$\begin{aligned} \langle A_n, \zeta\rangle &= \iint F_z(\widehat{\psi}_z^n)\widehat{\xi}_x^n(z)\zeta(x)\,dz\,dx \\ &= \int F_z(\widehat{\varphi}_z^n)(\widehat{\xi}^n + \xi)(z)\,dz. \end{aligned}$$

Recall that $\xi = \varphi^{1/2} - \varphi^2$. Imagine that φ satisfies $\int \varphi = 1$, $\int \varphi x^r \, dx = 0$ for $0 < |r| \leq \ell$. Hence, $\int \xi x^r \, dx = 0$ for $0 \leq |r| \leq \ell$. For such φ, we can assert

$$\widehat{\xi}_x^n + \zeta(z) = \int \widehat{\xi}^n(x-z)\zeta(z)\, dz = \int \widehat{\xi}^n(x-z)\underbrace{(\zeta(z) - P_x^\ell(z))}_{=O(|x-z|^{\ell+1})}\, dz = O(2^{-n(\ell+1)}),$$

where $P_x^\ell(z)$ is the Taylor expansion up to degree ℓ at x. Thus,

$$|\langle A_n, \zeta\rangle| \lesssim 2^{n(\beta-\ell-1)},$$

which is exponentially small if ℓ is sufficiently large. In summary, we have $A = A_\infty = A_1 + \sum_{n=1}^\infty (A_{n+1} - A_n)$ converges as a distribution.

We now turn to the B_ns; this is the one that only converges is $\gamma > 0$. Observe that

$$|\langle B_n, \zeta\rangle| = \left| \iint (F_x - F_z)(\widehat{\varphi}_z^n)\widehat{\xi}_x^n(z)\, dz \zeta(x)\, dz \right|$$

Since $|x-z|$ and δ are both of order 2^{-n},

$$\lesssim 2^{-n\gamma},$$

which is exponentially small if $\gamma > 0$.

In summary, the limit exists, and we have our candidate for T. It remains to verify that

$$|\langle T - F_x, \zeta_x^\delta\rangle| \lesssim \delta^\gamma,$$

locally uniformly. To prove this, observe that since $\rho = \varphi * \psi$, we can write

$$\begin{aligned} T(x) &= \lim_{n\to\infty} T_n(x) \\ &= \lim_{n\to\infty} F_x(\widehat{\varphi * \psi}^n)_x \\ &= \lim_{n\to\infty} F_x \widehat{\varphi}_y^n \widehat{\psi}_x^n(y)\, dy \\ &= \lim_{n\to\infty} F_x(\widehat{\varphi}_y^n)\psi_x^n(y)\, dy. \end{aligned}$$

Also,

$$\begin{aligned} F_x(\xi_x^\delta) &= \lim_{n\to\infty} F_x((\xi^\delta * \widehat{\varphi}^n)(x)) \\ &= \lim_{n\to\infty} F_x\left(\int \widehat{\varphi}_y^n \xi_x^\delta(y)\, dy\right) \\ &= \lim_{n\to\infty} \int F_x(\widehat{\varphi}_y^n)\xi_x^\delta(y)\, dy. \end{aligned}$$

We will complete the proof next time.

18 Proving the Bounds in Hairer's Reconstruction Theorem

18.1 Recap: Constructing a candidate in Hairer's reconstruction theorem

Theorem 18.1 (Hairer's reconstruction theorem). *Let F be γ-coherent. Then there exists a distribution $T = \mathcal{R}(F)$ such that*

$$|(T - F_x)(\psi_x^\delta)| \lesssim \begin{cases} \delta^\gamma & \gamma \neq 0, \\ \log \frac{1}{\delta} & \gamma = 0. \end{cases}$$

Here, the bound is uniform over $x \in K$, $\delta \in (0,1]$, $\psi \in \mathcal{D}$ with $\|\psi\|_{C^r} \leq 1$.

We want to think of T as the value of some continuous operator $\mathcal{R}$. So far, we have a candidate for T when $\gamma > 0$. Here is an overview of what we have seen so far. We start from $\varphi \in \mathcal{D}$ such that $\int \varphi = 1$ and $\int \varphi x^k \, dx = 0$ for $0 < |k| < r$. From this φ, we constructed a suitable test function ρ of the form $\rho = \eta * \varphi$ with $\eta = \varphi^2$ and $\rho^{1/2} - \rho = \zeta * \varphi$, where $\zeta = \varphi^{1/2} - \varphi^2$. Recall that

$$\varphi^\delta(x) = \delta^{-d}\varphi(x/\delta), \qquad \varphi_a^\delta(x) = \delta^{-d}\varphi((x-a)/\delta), \qquad \varphi_a(x) = \varphi(x-a).$$

Observe that since $\int \zeta = 0$, we have $\int \zeta P \, dx = 0$ for any polynomial P of degree at most $r - 1$.

Here is the idea behind the construction of T: Indeed if we define convolution by

$$(T * \phi)(X) = T(\widetilde{\phi}_x), \qquad \widetilde{\phi}(z) = \phi(-z),$$

and if $\int \phi = 1$, it can be shown that $\lim_{\delta \to 0} T * \phi^\delta = T$. Recall that $\widehat{\rho}_x^n(y) = 2^{dn}\rho(2^n(y - x)) = \rho_x^{2^{-n}}(y)$, and since

$$\lim_{n\to\infty} T(\widehat{\rho}_x^n) = T,$$

from this we guess tat a good approximation for T satisfying the theorem is simply

$$T_n(x) = F_x(\widehat{\rho}_x^n).$$

Last time, we showed that indeed $T_n(x)$ converges when $\gamma > 0$, where convergence means that for any $\psi \in \mathcal{D}$, $\lim_n \langle T_n(x), \psi \rangle$ exists. The very form of ρ allows us to have the following representation:

$$T_n = T_1 + \sum_{k=1}^{n-1} (T_{k+1} - T_k),$$

where

$$T_{k+1}(x) - T_k(x) = F_x(\widehat{m}_x^k) = \int F_x(\widehat{\varphi}_y^k)\widehat{\zeta}_x^k(y) \, dy, \qquad m = \rho^{1/2} - \rho.$$

We can write this as

$$T_{k+1}(x) - T_k(x) = \underbrace{\int (F_x - F_y)(\widehat{\varphi}_y^k)\widehat{\xi}_x^k(y)\,dy}_{B_k} + \underbrace{\int F_y(\widehat{\varphi}_y^k)\widehat{\zeta}_x^k(y)\,dy}_{A_k}.$$

Last time, we showed that $\sum_k A_k$ converges no matter γ is. However, the bound for the second term is $|B_k| \lesssim 2^{-\gamma k}$, so for B_k, we get a pointwise bound that would imply the pointwise convergence only when $\gamma > 0$. In fact, when $\gamma \le 0$, our candidate for T is $\lim_{n\to\infty} T_1 + \sum_{k=1}^{n-1} A_k = T_1 + \sum_{k=1}^{\infty} A_k$.

18.2 Proof of the bounds in the reconstruction theorem

We now try to prove that $(T - F_a)(\psi_a^\delta) \lesssim \delta^\gamma$. Here, a is fixed. We first focus on the case of $\gamma \le 0$. Again, our T is the limit of $S_n = T_1 + \sum_{k=1}^{n-1} A_k$. To compare this with F_a, observe that

$$F_a = \lim_{n\to\infty} F_a(\widehat{\rho}_\cdot^n).$$

That is,

$$F_a(\psi) = \lim_{n\to\infty} \int F_a(\widehat{\rho}_x^n)\psi(x)\,dx.$$

In the same manner, we may write

$$F_a = G_1 + \sum_{k=1}^{\infty}(G_{k+1} - G_k), \qquad G_k(x) := F_a(\widehat{\rho}_x^k).$$

Also, we may find

$$\Gamma_n(x) = G_1(x) + \sum_{k=1}^{n-1}(G_{k+1}(x) - G_k(x)),$$

so that $\lim_{n\to\infty} \Gamma_n = F_a$. We wish to compare Γ_n to $S_n = T_1 + \sum_{k=1}^{n-1} A_k$. Observe that

$$\begin{aligned} C_k(x) &= G_{k+1}(x) - G_k(x) \\ &= F_a(\widehat{\rho}_x^{k+1} - \widehat{\rho}_x^k) \\ &= F_a(\widehat{m}_x^k) \\ &= \int F_a(\widehat{\varphi}_y^k)\widehat{\zeta}_x^k(y)\,dy. \end{aligned}$$

We wish to estimate

$$|\langle A_k(x) - C_k(x), \psi_a^\delta\rangle| = \left| \iint (F_y - F_a)(\widehat{\varphi}_y^k)\zeta_x^k(y)\,dy\,\psi_a^\delta(x)\,dx \right|$$

$$
\begin{aligned}
&= \left| \int (F_y - F_a)(\widehat{\varphi}_y^k)(\widehat{\zeta}^k * \psi_a^\delta)(y)\, dy \right| \\
&\lesssim \int 2^{k\tau}(|y-a| + 2^{-k})^{\gamma+\tau} |(\zeta^k * \psi_a^\delta)(y)|\, dy
\end{aligned}
$$

As a warmup, observe that we have the bound

$$
\leq 2^{k\tau}(\delta + 2^{-k+1})^{\gamma+\tau} \|\psi^k * \psi_a^\delta\|_{L^1} \leq 2^{k\tau}(\delta + 2^{-k+1})^{\gamma+\tau} \|\zeta\|_{L^1} \|\psi\|_{L^1}.
$$

Hence, if $\gamma < 0$,

$$
\begin{aligned}
\sum_{k: 2^{-k} \geq \delta} |\langle A_k - C_k, \psi_a^\delta \rangle| &\lesssim \sum_{k: 2^{-k} \geq \delta} 2^{-k\gamma} \\
&= \sum_{k \leq |\log \delta|} 2^{-k\gamma} \\
&= \begin{cases} |\log \delta| & \gamma = 0 \\ (2^{-\gamma})^{|\log \delta|} = \delta^\gamma & \gamma < 0. \end{cases}
\end{aligned}
$$

Next, we concentrate on $\sum_{k: 2^{-k} < \delta} |\langle A_k - C_k, \psi_a^\delta \rangle|$. To control this, we need a better estimate on $(\widehat{\zeta}^k * \psi_a^\delta)(y)$, which has a support contained in $B_a(2^{-k+\delta})$. Recall that $\int \zeta P = 0$ for any polynomial P of degree $< r$. Now

$$
\int \widehat{\zeta}^k(y) \psi_a^\delta(x)\, dx = \int \widehat{\zeta}_x^k(y)(\psi_a^\delta(x) - P_y(x))\, dx,
$$

where $P_y(x)$ is the Taylor polynomial of ψ_a^δ at y of degree $r-1$. Hence,

$$
\begin{aligned}
\left| \int \widehat{\zeta}^k(y) \psi_a^\delta(x)\, dx \right| &\lesssim \int |y-x|^r \|\psi_a^\delta\|_{C^r} \|\widehat{\xi}_x^k(y)|\, dx \\
&\lesssim 2^{-kr} \|\psi_a^\delta\|_{C^r} \\
&\lesssim 2^{-kr} \delta^{-d-r} \|\psi\|_{C^r}.
\end{aligned}
$$

Hence,

$$
|\langle A_k - C_k, \psi_a^\delta \rangle| \lesssim 2^{k\tau}(\delta + 2^{-k+1})^{\gamma+\tau} 2^{-kr} \delta^{-d-r} \qquad \underbrace{(2^{-k} + \delta)^d}_{\text{volume of } B_a(2^{-k}+\delta)} .
$$

Thus,

$$
\sum_{k: 2^{-k} \leq \delta} |\langle A_k - C_k, \psi_a^\delta \rangle| \lesssim \sum_{k: 2^{-k} \leq \delta} 2^{-k(r-\tau)} \delta^{\gamma+\tau-r}
$$

Provided $r > \tau$, we get

$$\begin{aligned} &\lesssim \delta^{r-\tau}\delta^{\gamma+\tau-r} \\ &= \delta^{\gamma}. \end{aligned}$$

This completes the proof when $\gamma \leq 0$.

How about when $\gamma > 0$? We already know that the tail

$$\sum_{k:2^{-k}\leq\delta} |\langle A_k - C_k, \psi_a^\delta\rangle| \lesssim \delta^\gamma.$$

We now argue that

$$\sum_{k:2^{-k}\leq\delta} |\langle B_k, \psi_a^\delta\rangle| \lesssim \delta^\gamma$$

when $\gamma > 0$. Observe that

$$\begin{aligned} |\langle B_k, \psi_a^\delta\rangle| &= \left|\iint (F_x - F_y)(\widehat{\varphi}_y^k)\widehat{\zeta}_x^k(y)\,dy\,\psi_a^\delta(x)\,dx\right| \\ &\lesssim \left|\int 2^{k\tau}(|x-y| + 2^{-k})^{\gamma+\tau}\widehat{\zeta}_x^k(y)\,dy\,\psi_a^\delta(x)\,dx\right| \\ &\lesssim 2^{-k\gamma}\|\zeta\|_{L^1}\|\psi\|_{L^1}. \end{aligned}$$

Hence,

$$\sum_{k:2^{-k}\leq\delta} |\langle B_k, \psi_a^\delta\rangle| \lesssim \sum_{k:2^{-k}\leq\delta} 2^{-k\gamma} \lesssim \delta^\gamma.$$

19 Finishing Hairer's Reconstruction Theorem and Introduction to Regularity Structures

19.1 Finishing the proof of Hairer's reconstruction theorem

We have been proving the following theorem.

Theorem 19.1 (Hairer's reconstruction theorem)**.** *If G is γ-coherent, then there is a distribution T such that*

$$|(T - F_x)(\psi_x^\delta)| \lesssim \begin{cases} \delta^\gamma & \gamma \neq 0 \\ |\log \delta| & \gamma = 0. \end{cases}$$

Proof. Last time, we proved this when $\gamma \leq 0$. The proof we offered last time would yield the following estimate for $\gamma > 0$: Recall that

$$T = \lim_{n\to\infty} T_n, \qquad T_n(y) = F_y(\widehat{\rho}_y^n), \qquad T = T_1 + \sum_{k=1}^{\infty}(T_{k+1} - T_k),$$

$$F_x = \lim_{n\to\infty} G_n, \qquad G_n(y) = F_x(\widehat{\rho}_y^n), \qquad F_x = G_1 + \sum_{k=1}^{\infty}(G_{k+1} - G_k).$$

Last time, we proved that

$$\sum_{k:2^{-k}\leq\delta} |\langle (T_{k+1} - T_k) - (G_{k+1} - G_k), \psi_x\delta\rangle| \lesssim \delta^{-\gamma}, \qquad \gamma > 0.$$

It remains to show

$$|\langle T_n - G_n, \psi_x^\delta\rangle| \lesssim \delta^\gamma,$$

provided that $\delta \approx 2^{-n}$, or more specifically, $2^{-n} \leq \delta < 2^{-n+1}$. Observe that

$$T_n(y) - G_n(y) = (F_y - F_x)(\widehat{\rho}_y^n), \qquad \rho = \varphi * \eta, \quad \eta = \varphi^2.$$

Hence,

$$(T_n - G_n)(y) = \int (F_y - F_x)(\widehat{\varphi}_z^n)\widehat{\eta}_y^n(z)\, dz$$

Now

$$|\langle (T_n - G_n), \psi_x^\delta\rangle| = \iint \underbrace{(F_y - F_x)}_{F_y - F_z + F_z - F_x} (\widehat{\varphi}_z^n)\widehat{\eta}_y^n(z)\psi_x^\delta(y)\, dy\, dz$$

Using the coherence,

$$\lesssim \iint [\underbrace{2^{n\tau}(|y-z| + 2^{-n})^{\gamma+\tau}}_{2^{-\gamma n}} + \underbrace{2^{n\tau}(|x-z| + 2^{n})^{\gamma+\tau}}_{2^{-\gamma}}]|\widehat{\eta}_y^n(z)\psi_x^\delta(y)|\, dy\, dz$$

$$\lesssim 2^{-\gamma n}\|\eta\|_{L^1}\|\psi\|_{L^1}.$$

We are done. $\square$

19.2 Remarks about the reconstruction theorem

Remark 19.1. The way we constructed $T = \lim_{n\to\infty} F_x(\widehat{\rho}_n^n) = \mathcal{R}(F)$ is linear in F. Moreover, if we define

$$|||F|||_{K,\varphi} = \sup_{x,y\in K} \sup_{\delta\in(0,1]} \frac{(F_x - F_y)(\varphi_y^\delta)}{\delta^{-\tau}(|x-y|+\delta)^{\gamma+\tau}},$$

where τ and γ depend on the compact set K, then

$$|(T - F_x)(\psi_x^\delta)| \lesssim |||F|||\delta^\gamma$$

uniformly over $\psi \in \mathcal{D}_r$, $x \in K$, $\delta \in (0,1]$, where

$$\mathcal{D}_r = \{\psi \in \mathcal{D} : \|\psi\|_{C^r} \le 1, \operatorname{supp}\psi \subseteq B_1(0)\}.$$

Remark 19.2. As an example, take $f \in \mathcal{C}^\alpha(\mathbb{R}^d)$, $g \in \mathcal{C}^\beta(\mathbb{R}^d)$, $\alpha, \beta \in (0,1)$, and set $F_x = f(x)\nabla g$. Observe that if $g \in \mathcal{C}^\beta$, then $\nabla g \in \mathcal{C}^{\beta-1}(\mathbb{R}^d)$. By $\mathcal{C}^\tau(\mathbb{R}^d)$ with $\tau < 0$, we mean this: First, pick $r = r(\tau)$ to be the smallest positive integer r such that $-\tau < r$ (or $\tau > -r$). Define

$$[T]_{K,\tau} := \sup_{\delta\in(0,1)} \sup_{\varphi\in\mathcal{D}_r} \frac{|T(\varphi_x^\delta)|}{\delta^T},$$

$$\mathcal{C}^\tau_{\text{loc}} := \{T : [T]_{K,\tau} < \infty \text{ for every } K\}.$$

Then $g \in \mathcal{C}^\beta \implies \nabla g \in \mathcal{C}^{\beta-1}$.

Now

$$\begin{aligned}(F_x - F_y)(\varphi_x^\delta) &= (f(x) - f(y))\nabla g(\varphi_x^\delta) \\ &= -(f(x) - f(y))g(\nabla\varphi_x^\delta) \\ &= \delta^{-1}(f(x) - f(y))g((\nabla\varphi)_x^\delta).\end{aligned}$$

Since we are dealing with ∇g, we can replace g by $g - g(x)$ (subtracting a constant). Hence, if $|x-y| \le 1$, then

$$\begin{aligned}|(F_x - F_y)(\varphi_x^\delta)| &\lesssim \delta^{-1}[f]|x-y|^\alpha[g]_\beta\delta^\beta \\ &\le [f]_\alpha[g]_\beta\delta^{-1}(|x-y|+\delta)^{\alpha+\beta} \\ &= [f]_\alpha[g]_\beta\delta^{-1}(|x-y|+\delta)^{\gamma+1},\end{aligned}$$

where $\gamma = \alpha + \beta - 1$. Thus, F is $(-1, \gamma)$-coherent.

Use the theorem to assert that there exists some operator $\Gamma(f,g) = \mathcal{R}(F)$ such that

$$|(\Gamma(f,g) - f(x)\nabla g)(\psi_x^\delta)| \lesssim [f]_\alpha[g]_\beta\delta^\gamma.$$

Note that since $\mathcal{R}$ is linear in F, Γ is bilinear and continuous in (f,g). In fact, $\Gamma(f,g)$ is unique if $\gamma > 0$. On the other hand, if $f,g \in C^1$, then $T(y) = f(y)\nabla g(y)$ also satisfies the above inequality. By uniqueness, $\Gamma(f,g) = f\nabla g$ for f,g smooth. The same comment does *not* apply to the case of $\gamma \leq 0$.

Remark 19.3. Our result can be extended to Besov spaces $\mathscr{B}^{\gamma}_{p,q}$. Roughly, in $\mathscr{B}^{\gamma}_{p,q}$ we replace the uniform norm in x with L^p norm and uniform in $\delta \in (0,1)$ with $L^q(\frac{1}{\delta}\, d\delta)$.

19.3 Introduction to regularity structures

For our purposes, we often have various terms in our PDE that involve a local description of various different exponents. To do this in a systematic way, we introduce the theory of **regularity structures**. Here is the set-up.

(i) There is a discrete set $A \subseteq \mathbb{R}$ that is bounded below. Roughly, each α in A represents terms that are in $\mathcal{C}^\alpha$ in our PDE. We always assume $0 \in A$.

(ii) For each α, we have a Banach space T_α with norm $\|\cdot\|_\alpha$. For $T_0 = \mathbb{R} = \mathrm{span}(\mathbb{1})$. Fpr $T_0 = \mathbb{R} = \mathrm{span}(\mathbb{2}) = \langle \mathbb{1} \rangle$,.

(iii) We also consider a group G of linear, continuous transformations $\Gamma : T \to T$ where $T = \bigoplus_{\alpha \in A} T_\alpha$. Moreover, we assume $\tau \in T_\alpha$, $\Gamma\tau - \tau \in \bigoplus_{\beta < \alpha} T_\beta$.

(i)-(iii) yields a structure (A, T, G).

We need a model to turn this abstract stuff into real stuff: $(\pi_x, \Gamma_{x,y} : x, y \in \mathbb{R}^d)$. Here, π_x is a bounded, linear map from $T \to \mathcal{D}'$ with each $\Gamma_{x,y} \in G$ satisfying

$$\pi_x \Gamma_{x,y} \tau = \pi_y \tau.$$

In short,

$$\pi_x \Gamma_{x,y} = \pi_y.$$

Example 19.1 ($e^{2+\gamma}, \gamma \in (0,1)$)**.** Let $d = 1$, $A = \{0,1,2\}$, $T_0 = \langle \mathbb{1} \rangle$, $T_1 = \langle X \rangle$, $T_2 = \langle X^2 \rangle$. Then $T = \{\tau = c_0\mathbb{1} + c_1 X + c_2 X^2 : c_0, c_1, c_2 \in \mathbb{R}\}$, and

$$G = \{\Gamma_h : h \in \mathbb{R}\}, \quad \text{where}\Gamma_h \tau = (c_0\mathbb{1} + (X + h\mathbb{1})^2 + c_2(X + h\mathbb{1})^2).$$

Then

$$\begin{aligned}\Gamma_h \tau - \tau &= c_1 h \mathbb{1} + 2c_2 h X + c_2 h^2 \mathbb{1} \\ &= (c_1 h + c_2 h^2)\mathbb{1} + 2c_2 h X.\end{aligned}$$

20 Regularity Structures

20.1 Regularity structures and their relation to coherence

We have proved the reconstruction theorem, which says that if you have some local regularity, i.e. coherence, then you can construct a distribution which serves as a local approximation. Last time, we discussed the regularity structure, a bookkeeping device for discussing obstructions we deal with in a PDE: Namely, we have a discrete $A \subseteq \mathbb{R}$ with $0 \in A$ and $\min A > -\infty$ (the members of this set represent the homogeneity and hence regularity of various terms you have to deal with). We have a Banach space $T = \bigoplus_{r \in A} T_r$ with Banach spaces T_r with $\|\cdot\|_r$; we are mostly interested in when A is a finite set and T_r are Euclidean spaces. We always assume $T_0 = \mathbb{R}$; the dimension of these spaces will be the number of beasts of that type you have to deal with.

If we have a Taylor expansion at a point x we can re-expand to turn it into a Taylor expansion at a point y. We express this idea in this setting by a group G of linear, continuous transformations $\Gamma : T \to T$. Moreover, for $\tau \in T_r$, $\Gamma\tau - \tau \in \bigoplus_{s<r} T_s$. We also write

$$T_{<r} = \bigoplus_{s<r} T_s, \qquad T_{\leq r} = \bigoplus_{s \leq r} T_s.$$

Definition 20.1. Let $\mathcal{L}(T) = \mathcal{L}(T; \mathcal{D}')$ be the set of linear, continuous maps $L : T \to \mathcal{D}'$. We say $M = (\Pi, \Gamma)$ is a **model** for (A, T, G) if $\Pi : \mathbb{R}^d \to \mathcal{L}(T)$ and $\Gamma : \mathbb{R}^d \times \mathbb{R}^d \to G$ with the following properties (denoting $\Pi_x = \Pi(x), \Gamma_{x,y} = \Gamma(x, y)$):

1. $\Pi_x \tau = \pi_y \Gamma_{x,y} \tau$

2. $\Gamma_{x,y}\Gamma_{y,z} = \Gamma_{x,z}$.

3. If $\tau \in T_\alpha$, then

$$\sup_{\delta \in (0,1]} \sup_{\varphi \in \mathcal{D}_r} \sup_{x \in K} \frac{|(\Pi_x \tau)(\varphi_x^\delta)|}{\|\tau\|_\alpha \delta^\alpha} < \infty,$$

   where K is a compact set, and $\mathcal{D}_r = \{\varphi \in \mathcal{D} : \operatorname{supp} \varphi \subseteq B_1(0), \|\varphi\|_{C^r} \leq 1\}$, where r is the smallest integer that is more than $-\min A$.

4. $\|\Gamma_{x,y}\tau\|_\beta \lesssim |x - y|^{\alpha - \beta} \|\tau\|_\alpha$.

Π turns an abstract symbol into a distribution in a way that respects all this linear structure.

Definition 20.2. Next, we define $\mathcal{C}_M^\gamma$ to be the set of functions $f : \mathbb{R}^d \to T_{<\gamma}$ such that

$$\|f(x) - \Gamma_{x,y} f(y)\|_\alpha \lesssim |x - y|^{\gamma - \alpha}$$

for $\alpha < \gamma$.

Hence, we may define the norm

$$[f]_\alpha = \sup_{\alpha<\gamma} \sup_{x\neq y\in K} \frac{\|f(x) - \Gamma_{x,y} f(y)\|_\alpha}{|x-y|^{\gamma-\alpha}}.$$

Theorem 20.1. *For every γ, there exists an operator $\mathcal{R}_\gamma : \mathcal{C}_M^\gamma \to \mathcal{D}'$ which is linear and continuous and which satisfies*

$$|(\mathcal{R}_\gamma f - \Pi_x f(x))(\varphi_x^\delta)| \lesssim \begin{cases} \delta^\gamma & \gamma \neq 0 \\ |\log \delta| & \gamma = 0, \end{cases}$$

uniformly for $\psi \in \mathcal{D}_r, \delta \in (0,1], x \in K$.

Proof. To simplify the notation, we write $f_x = f(x)$ and $\Pi_x = \Pi(x)$. Now define $F_x = \Pi_x(f_x)$. We can achieve the desired result if we can show that the germ $(F_x : x \in \mathbb{R}^d)$ is γ-coherent. Indeed,

$$\begin{aligned} |(F_x - F_y)(\varphi_y^\delta)| &= |(\Pi_x f_x - \Pi_y f_y)(\varphi_y^\delta)| \\ &= |(\Pi_x f_x - \Pi_x \Gamma_{x,y} f_y)(\varphi_y^\delta)| \\ &= |\Pi_x (f_x - \Gamma_{x,y} f_y)(\varphi_y^\delta)| \end{aligned}$$

Recall that $f : \mathbb{R}^d \to \bigoplus_{\alpha<\gamma} T_\alpha$, so $\|f_x - \Gamma_{x,y} f_y\| \lesssim |x-y|^{\gamma-\alpha}$. Here, $P_\alpha \tau$ means the α-component of τ, i.e. $\tau = \sum_{\alpha\in A}(P_\alpha \tau)$.

$$\begin{aligned} &= \left|\prod_x P_\alpha (f_x - \Gamma_{x,y} f_y)(\varphi_y^\delta)\right| \\ &\leq \sum_{\alpha<\gamma} \left|\prod_x \sum_{\alpha<\gamma} P_\alpha (f_x - \Gamma_{x,y} f_y)(\varphi_y^\delta)\right| \\ &\lesssim \sum_{\alpha<\gamma} \delta^\alpha \|f_x - \Gamma_{x,y} f_y\|_\alpha \\ &\lesssim \sum_{\alpha<\gamma} \delta^\alpha |x-y|^{\gamma-\alpha} \\ &= \delta^{-r} \sum_{\alpha<\gamma} \delta^{\alpha+r} |x-y|^{\gamma-\alpha} \\ &= \delta^{-r} \sum_{\alpha<\gamma} (\delta + |x-y|)^{\gamma+r} \\ &\lesssim \delta^{-r} (\delta + |x-y|)^{\gamma+r}, \end{aligned}$$

which is exactly the definition of coherence.[9] □

[9] Professor Rezakhanlou described this as a "one-line proof."

20.2 An example: Taylor series

Example 20.1. Assume $A = \{0, 1, 2, \dots\}$, and let $T = \mathbb{R}[X_1, \dots, X_d]$ be the space of polynomials of variables $X_1, \dots, X_d$ with real coefficients. Again, we use $X^k = X_1^{k_1} \cdots X_d^{k_d}$, where $k = (k_1, \dots, k_d) \in \mathbb{N}^d$ and $|k| = k_1 + \cdots + k_d$. Now T_r is the set of homogeneous polynomials of degree r, $\operatorname{span}(\{X^k : |k| = r\}) = \langle X^k : |k| = r\rangle$, and $T_{\le r}$ is the set of polynomials of degree $\le r$. (By $S = \langle \tau^1, \dots, \tau^e ll\rangle$, we mean $S = \operatorname{span}(\tau^1, \dots \tau^\ell)$ and $\tau^1, \dots, \tau^\ell$ are linearly independent.)

Next, $G = \{\Gamma_h : h \in \mathbb{R}^d\}$. Formally,

$$\Gamma_h(P(X)) = P(X + h\mathbf{1}).$$

For example, $(X + h\mathbf{1})^k = \prod_{i=1}^d (X_i + h\mathbf{1})^{r_i}$, using the convention that $X_i\mathbf{1} = \mathbf{1}X_i = X_i$. If $\deg P = r$, then $\deg(\Gamma_h P - P) < r$.

We now define a model for this:

$$\Pi_a(P(X))(x) = P(x - a).$$

Observe that

$$\begin{aligned}\langle \Pi_a(X^k), \varphi_a^\delta\rangle &= \int \Pi_a(X^k)(x)\varphi_a^\delta(x)\\ &= \int (x-a)^k \varphi_a^\delta \, dx\end{aligned}$$

Use a change of variables.

$$= \left(\int P(x)\varphi(x)\, dx\right)\delta^{|k|}$$

Next, we discuss $\mathcal{C}_M^\gamma$ with $\gamma = n + \gamma_0$, where $n \in \mathbb{N}$ and $\gamma_0 \in (0, 1)$. Let $f \in \mathcal{C}_M^\gamma$. Then $f(x) = \sum_k c_k(x)X^k \in \bigoplus_{r<\gamma} T_r$ is a polynomial of degree n. We claim that we must have that $c_k = \frac{1}{k!}\partial^k c_0$ with $c_0 \in \mathcal{C}^\gamma$ and c_k must be Hölder of exponent γ_0 when $|k| = n$.

This is the same flavor as in the Whitney extension theorem. First, we have the Tietze extension theorem. If we have a closed subset of a decent topological space, we can extend a continuous function on the closed subset to the whole set without increasing the norm of it. The Whitney extension theorem achieves this with derivatives by assigning polynomials to each point and showing that they must be related via Taylor expansion.

21 Two Examples of Regularity Structures

We discuss two models for our theory before treating our ill-posed PDEs.

21.1 Finite Taylor Polynomials

Example 21.1. Let $A = \mathbb{N}$, $T = \mathbb{R}[X_1, \ldots, X_d]$, with $T_r = \langle X^k : |k| = r \rangle$, and $\| \cdot \|_r$ the standard Euclidean norm. Recall that $\Gamma_{x,y} = \Gamma_{x-y}$, with $\Gamma_h(P(X)) = P(X + h\mathbf{1})$. For our model, $(\Pi_a(P(X)))(x) = P(x - a)$. This gives the model $M = (\Pi, \Gamma)$. We now specify $\mathcal{C}_M^\gamma = \{f : \mathbb{R}^d \to \bigoplus_{r<\gamma} T_r \mid \|f(x) - \Gamma_{x,y} f(y)\|_r \lesssim |x - y|^{\gamma - r}\}$.

We claim that for any $\gamma > 0$, $\mathcal{C}_M^\gamma$ is isomorphic to $\mathcal{C}^\gamma(\mathbb{R}^d)$. Let us assume that $\gamma = n + \gamma_0$ with $n \in \mathbb{N}$ and $\gamma_0 \in (0, 1)$. Then $f \in \mathcal{C}_M^\gamma$ means that $f(x)$ is a polynomial of degree at most n i.e. $f(x) = \sum_{k:|k|\le n} a_k(x) X^k$, with (setting $h = x - y$ so that $x = y + h$)

$$\left\| \sum_{k:|k|\le n} a_k(y+h) X^k - \sum_{k:|k|\le n} a_k(y)(X + h\mathbf{1})^k \right\|_r \lesssim |h|^{\gamma - r}.$$

For example, if $r = n$,

$$\sum_{k:|k|\le n} |a_k(y+h) - a_k(y)| \lesssim |h|^{\gamma_0},$$

which means that when $|k| = n$, $a_k(y)$ is γ_0-Hölder. More generally,

$$\sum_{|\ell| = r} \left| a_\ell(y+h) - \sum_{\substack{k:k\ge \ell \\ |k|\le n}} \binom{k}{\ell} a_k(y) h^{k-\ell} \right| \lesssim |h|^{\gamma - r}.$$

To ease the notation, assume $d = 1$ and $r = n - 1$. Then we get

$$|a_{n-1}(y+h) - a_{n-1}(y) - n a_n(y) h| \lesssim |h|^{\gamma_0 + 1}$$

Divide by h and send $h \to 0$ to arrive at: a_{n-1} is differentiable, and $\frac{d}{dy} a_{n-1} = n a_n$. Inductively, we can show that

$$a_k(y) = \frac{1}{k!} \partial^k a_0(y).$$

In summary,

$$f(x) = \sum_{|k|\le n} a_k(x) X^k \in \mathcal{C}_M^{\gamma_0 + h} \iff a_0 \in \mathcal{C}^\gamma(\mathbb{R}^d), f(x) = \text{Taylor expansion of deg } n \text{ of } a_0.$$

Remark 21.1 (Whitney expansion)**.** Imagine that a closed set $K \subseteq \mathbb{R}^d$ is given and we assign a polynomial $f(x)$ as above to each $x \in K$. If for $x \in K$ the bound

$$\left\| \sum_{k:|k|\leq n} a_k(y+h)X^k - \sum_{k:|k|\leq n} a_k(y)(X+h\mathbf{1})^k \right\|_r \lesssim |h|^{\gamma-r}.$$

holds, then $f(x)$ can serve as a candidate for the Taylor expansion of a suitable function $a_0 : \mathbb{R}^d \to \mathbb{R}$ such that for $x \in K$, $f(x)$ is indeed its Taylor expansion.

21.2 The Gubinelli derivative

Example 21.2. Pick $\alpha \in (1/3, 1/2)$, and choose $A = \{\alpha-1, 2\alpha-1, 0, \alpha\}$. Note that $r = 1$, i.e. the integer -1 is the best lower bound for A. We define

$$T_0 = \langle \mathbf{1} \rangle, \qquad T_\alpha = \langle X_1, X_2, \ldots, X_\ell \rangle,$$

$$T_{\alpha-1} = \langle \dot{X}_1, \dot{X}_2, \ldots, \dot{X}_\ell \rangle, \qquad T_{2\alpha-1} = \langle \dot{\mathbb{X}}^{i,j} : 1 \leq i,j \leq \ell \rangle,$$

so $T = \bigoplus_{\beta \in A} T_\beta$ has $\dim T = (\ell+1)^2$. Here, these are all just formal symbols, but we have written the notation to be suggestive. Next, $G = \{\Gamma_h : h \in \mathbb{R}^\ell\}$ with

$$\Gamma_h \mathbf{1} = \mathbf{1}, \qquad \Gamma_h X = X + h\mathbf{1},$$

$$\Gamma_h \dot{X} = \dot{X}, \qquad \Gamma_h \dot{\mathbb{X}} = \dot{\mathbb{X}} + h \otimes \dot{X}.$$

Next, we define a model. Given a rough path $\mathbf{x} = (x, \mathbb{X}) \in \mathscr{R}_{\alpha,2\alpha}$, i.e. $x : [0,T] \to \mathbb{R}^\ell$, $\mathbb{X}(s,t) \in \mathbb{R}^{\ell\times\ell}$, and Chen's relation. We build a model as follows:

$$(\Pi_s \mathbf{1})(t) = 1, \qquad (\Pi_s X_i)(t) = x_i(s,t) = x_i(t) - x_i(s),$$

$$(\Pi_s(\dot{X}_i))(\underbrace{\psi}_{\in\mathcal{D}}) = (\dot{x}_i)(\psi) = -\int \dot{\psi}(t) x_i(t)\, dt,$$

$$(\Pi_s \dot{\mathbb{X}}^{i,j})(\psi) = (\mathbb{X}_t^{i,j}(s,\cdot))(\psi) = -\int \dot{\psi}(t)\mathbb{X}^{i,j}(s,t)\, dt.$$

Next, we have

$$\Gamma_{s,s'} = \Gamma_{x(s',s)}.$$

We need to verify a number of things:

- $\langle \Pi_s, \varphi_s^\delta \rangle = \int ((x(t)-x(s))\varphi(\frac{t-s}{s})\frac{1}{\delta} \lesssim \delta^\alpha$, which follows from $x \in \mathcal{C}^\alpha$.

- Similarly,

$$(\Pi_s(\dot{\mathbb{X}})(\varphi_s^\delta) = -\int \frac{d}{dt}\varphi_s(t)\mathbb{X}(s,t)\,dt = -\frac{1}{\delta}\int \dot{\varphi}(\theta)\mathbb{X}(s,s+\delta\theta)\,d\theta.$$

  Hence,

$$|(\Pi_s\dot{\mathbb{X}})(\varphi_s^\delta)| \lesssim [\mathbb{X}]_{2\alpha}\delta^{2\alpha-1}\|\varphi\|_{C^1}.$$

- Next, we need to check $\Pi_{s'} = \Pi_s\Gamma_{s,s'}$. Indeed,

$$(\Pi_{s'}\dot{\mathbb{X}})(\psi) = -\int \dot{\psi}(t)\mathbb{X}(s',t)\,dt,$$

$$(\Pi_s\Gamma_{s,s'}\dot{\mathbb{X}})(\psi) = -\int \dot{\psi}(t)(\mathbb{X}(s,t) + x(s',s)\otimes x(t))\,dt$$

  Since ψ is of 0 average,

$$= -\int \dot{\psi}(t)(\mathbb{X}(s,t) + x(s',s)\otimes x(s,t))\,dt$$

$$= -\int \dot{\psi}(t)(\mathbb{X}(s,t) + \mathbb{X}(s',s) + x(s',s)\otimes x(s,t))\,dt$$

  By Chen's relation,

$$= \int \dot{\psi}(t)\mathbb{X}(s',t)\,dt,$$

  as desired.

Next, we examine $\mathcal{C}_M^{2\alpha}$. Assume $Y \in \mathcal{C}_M^{2\alpha}$ is of the form

$$Y(t) = y(t)\mathbf{1} + \widehat{y}(t)\cdot X.$$

We claim that $T \in \mathcal{C}_M^{2\alpha}$ if and only if $\mathbf{y} = (y,\widehat{y}) \in \mathscr{G}_{\alpha,2\alpha}(\mathbf{x})$ (i.e. $\widehat{y}$ is a Gubinelli derivative of y). Indeed,

$$\|Y(t) - \Gamma_{t,t'}Y(t')\|_r \lesssim |t-t'|^{2\alpha-r}.$$

This is

$$\|y(t)\mathbf{1} + \widehat{y}(t)\cdot X - (y(t')\mathbf{1} + \widehat{y}(t')\cdot(X + x(t',t)\nVdash))\| \lesssim |t-t'|^{2\alpha-1}.$$

Choose $r = \alpha$. Then

$$|\widehat{y}(t) - \widehat{y}(t')| \lesssim |t-t'|^{\alpha},$$

i.e. $\widehat{y} \in \mathcal{C}^\alpha$. Next, choose $r = 0$. We get

$$|y(t) - y(t') - \widehat{y}(t')x(t',t)| \lesssim |t-t'|^{2\alpha}.$$

Imagine that we want to make sense of $\mathbf{y}\cdot d\mathbf{x}$. This should really be the realization of $Y\cdot\dot{X}$. We will give a candidate for the abstract multiplication and recover our previous results using the reconstruction theorem.

22 Applying Regularity Structures to Rough Path Theory and Singular PDEs

22.1 Recovering a previous theorem as an application of the reconstruction theorem

For our rough path theory, we choose $A = \{\alpha - 1, 2\alpha - 1, 0, \alpha\}$ with $\alpha \in (1/3, 1/2)$. Here, $T_\alpha = \langle X_1, \dots, X_\ell \rangle$, where we think of $X = (X_1, \dots, X_\ell)$ as an abstract candidate for the path $x(\cdot) \in \mathcal{C}^\alpha$, $T_{\alpha-1} = \langle \dot{X}_1, \dots, \dot{X}_\ell \rangle$, and $T_{2\alpha-1} = \langle \dot{\mathbb{X}}^{i,j} : 1 \le i, j \le \ell$. We think of $\dot{\mathbb{X}} = [\mathbb{X}^{i,j}] = X \otimes \dot{X}$. From this, we have

$$\Gamma_h X = X + h\mathbf{1}, \qquad \Gamma_h \dot{X} = \dot{X}, \qquad \Gamma_h (X \otimes \dot{X}) = X \otimes \dot{X} + h \otimes \dot{X}.$$

Recall that $f : \mathbb{R}^d \to \bigoplus_{\alpha < \gamma} T_\alpha \in \mathcal{C}^\gamma_M$ means $\|f(s) - \Gamma_{st} f(t)\|_\alpha \lesssim |s-t|^{\gamma-\alpha}$. So if we decrease the index, the regularity required would be rougher. Last time, we argued that if $Y(t) = y(t)\mathbf{1} + \widehat{y}(t) \cdot X \in \mathcal{C}^{2\alpha}_M$, then the pair $\mathbf{y}(t) = (y(t), \widehat{y}(t)) \in \mathscr{G}^\alpha(x)$, i.e.

$$|\widehat{y}(t) - \widehat{y}(s)| \lesssim |t-s|^\alpha, \qquad |y(t) - y(s) - \widehat{y}(s) x(s,t)| \lesssim |t-s|^{2\alpha}.$$

Now we want to examine another algebraic manipulation in our abstract setting, namely we wish to make sense of $Y \cdot \dot{X}$, which we want to think of as $(y\mathbf{1} + \widehat{y}X) \cdot \dot{X} = y\dot{X} + \widehat{X} \otimes \dot{X}$. Because of this, consider

$$(Y \cdot \dot{X})(t) = y\dot{X} + \widehat{y}\dot{\mathbb{X}}.$$

Proposition 22.1. *$(y, \widehat{y}) \in \mathscr{G}^\alpha(x)$ if and only if $Y \cdot \dot{X} \in \mathcal{C}^{3\alpha-1}_M$.*

Proof.

$$\begin{aligned}(Y \cdot \dot{X})(s) - \Gamma_{s,t}(Y \cdot \dot{X})(t) &= (y(s)\dot{X} + \widehat{y}(s)\dot{\mathbb{X}}) - (y(t)\dot{X} + \widehat{y}(t)\dot{\mathbb{X}} + \widehat{(}s)x(t,s)\dot{X}) \\ &= (y(s) - y(t) - \widehat{y}(s)x(t,s))\dot{X} + (\widehat{y}(s) - \widehat{y}(t))\dot{\mathbb{X}}\end{aligned}$$

For the first coefficient, we want the estimate

$$|y(s) - y(t) - \widehat{y}(s)x(t-s)| \lesssim |t-s|^{\gamma-(\alpha-1)} = |t-s|^{2\alpha}.$$

This is exactly the estimate for the Gubinelli derivative. Similarly, we want

$$|\widehat{y}(s) - \widehat{y}(t)| \lesssim |t-s|^{\gamma-(2\alpha-1)} = |t-s|^\alpha.$$

This gives the equivalence. $\square$

Now we wish to apply our reconstruction theorem to $Y \cdot \dot{X}$. More precisely, there exists some operator $J^{3\alpha-1}_M$ on $\mathcal{C}^{3\alpha-1}_M$ such that $W := J^{3\alpha-1}_M(Y \cdot \dot{X})$ satisfies

$$|(W - \Pi_t(y(t)\dot{X} + +\widehat{y}(t)\dot{\mathbb{X}}))(\psi^\delta_t)| \lesssim \delta^{3\alpha-1}.$$

Equivalently,

$$|W(\psi_t^\delta) - (y(t)\dot{x} + \widehat{y}(t)\mathbb{X}_t(t,\cdot))(\psi_t^\delta)| \lesssim \delta^{3\alpha-1}.$$

This is indeed the first theorem we proved in this class, namely given $\mathbf{y} = (y, \widehat{y}) \in \mathscr{G}^\alpha(x)$ and $\mathbf{x} = (x, \mathbb{X}) \in \mathscr{R}_{\alpha,2\alpha}$, there exists $z \in \mathcal{C}^\alpha$ such that

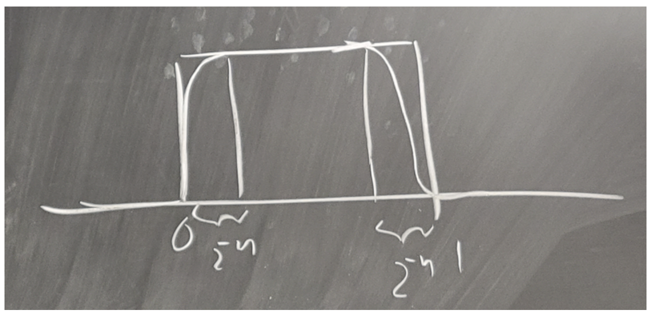
$$|z(s) - z(t) - y(t)(x(s) - x(t)) - \widehat{y}(t)\mathbb{X}(t,s)| \lesssim |t-s|^{3\alpha-1}$$

with $\dot{z} = W$.

To derive this theorem from estimate above it, we need to allow a ψ that is of the form $\psi(t) = \mathbb{1}_{[0,1]}(t)$ sot that $\psi_t^\delta(s) = \frac{1}{\delta}\mathbb{1}_{[t,t+\delta]}(s)$. This can be achieved by writing

$$\mathbb{1}_{[0,1]} = \sum_{n=0}^{\infty} \varphi_n(t) + \psi_n(t),$$

where ψ_n, φ_n are smooth with compact support, $\operatorname{supp}\varphi_n \subseteq [0, 2^{-n}]$, and $\operatorname{supp}\psi_n \subseteq [1 - 2^{-n}, 1]$.



22.2 Applying regularity structure theory to understand a singular PDE

We now turn our attention to one of our singular PDE, say the KPZ equation

$$\begin{cases} h_t = h_{xx} + h_x^2 + \xi - C \\ h(x,0) = h^0(x), \end{cases}$$

where ξ is white noise. As we argued before, if $\xi^\varepsilon = \xi *_x \chi^\varepsilon$ with $\chi^\varepsilon(x) = \frac{1}{\varepsilon}\chi(\frac{x}{\varepsilon})$, then the corresponding PDE

$$h_t^\varepsilon = h_{xx}^\varepsilon + (h_x^\varepsilon)^2 + \xi^\varepsilon - C_\varepsilon$$

is well-posed, and $\lim_{\varepsilon\to 0} h^\varepsilon$ exists only if $C_\varepsilon \approx C/\varepsilon$, where $C = \frac{1}{2}\int \chi^2$ (a theorem due to Martin Hairer).

To achieve this, we first build an abstract version of our PDE and usr it to have an abstract solution that is continuous with respect to its input (which in cludes a well-selected version of ξ). Indeed, if we write $\mathcal{P}$ for the operator/kernel $(\partial_t - \partial_x^2)^{-1}$, then

$$h = \mathcal{P} * (h_x^2 + \xi - C) + \mathcal{P} * h^0 = \mathcal{F}(h).$$

Then we would make sense of $\mathcal{F}$ in a suitable way, show that $\mathcal{F}$ has a fixed point, and this would be our candidate for the solution. For this, we need some preparations.

Definition 22.1. Given a regularity structure (A, T, G), we say $V \subseteq T$ is a **sector** if $V = \bigoplus_{\alpha \in A} V_\alpha$ with subspaces $V_\alpha \subseteq T_\alpha$ and $G(V) \subseteq V$.

Definition 22.2. If $\mathcal{L} = \sum_{|k|=r} a_k \partial^k$ is a differential operator, we say $\widehat{\mathcal{L}} : V \to T$ **represents** $\mathcal{L}$ if the following conditions hold:

- If $\tau \in V_\alpha$, then $\widehat{\mathcal{L}}\tau \in T_{\alpha - r}$.
- $\widehat{\mathcal{L}}\Gamma_h = \Gamma_h \widehat{\mathcal{L}}$.
- $\Pi_a \widehat{\mathcal{L}}\tau = \mathcal{L}(\Pi_a \tau)$.

We can also talk about products. In other words, we want to be able to multiply $f \in \mathcal{C}^\alpha_M$ and $g \in \mathcal{C}^\beta_M$ to get $f \odot g \in \mathcal{C}^{\alpha \wedge \beta}_M$.

Recall that i we have a distribution F, then we can talk about

$$F * K \text{ ``=''} \int F(y) K(x-y)\, dy = F(y)\widetilde{K}(y-x)\, dy = \int F(y) \widetilde{K}_x(y)\, dy,$$

where $\widetilde{K}(y) = K(-y)$, which suggests that we should define

$$(F * K)(\varphi) := F(\widetilde{K} * \varphi).$$

For our purposes, we need to examine the regularity of $F * K$. A Schauder-type estimate allows us to show that if K is singular at 0 with singularity of the form $|x|^{\alpha - d}$, then

$$F \in \mathcal{C}^\gamma \implies F * K \in \mathcal{C}^{\gamma + \alpha}.$$

Here is the precise statement:

Theorem 22.1. *Assume that $K : \mathbb{R}^d \to \mathbb{R}$ with the following conditions:*

1. $\operatorname{supp} K \subseteq B_1(0)$
2. *$K \in \mathcal{C}^\infty$ (this can be relaxed), and $|\partial^\ell K(x)| \leq c_\ell |x|^{\alpha - d - |\ell|}$ for all x.*

Then

$$F \in \mathcal{C}^\gamma \implies F * K \in \mathcal{C}^{\gamma + \alpha}$$

for all $\gamma \in \mathbb{R}$, though for $\gamma \in \mathbb{Z}$, we need to replace the Hölder spaces with Hölder-Zygmund spaces.

For the proof, we need a suitable candidate for function spaces that are equivalent to Hölder spaces (and its variant would yield Besov spaces), except when $\gamma \in \mathbb{Z}$. For $\gamma < 0$, we have already discussed this; if r is the smallest integer such that $r + \gamma \geq 0$, then define

$$[u]_{\gamma,K} = \sup_{x \in K} \sup_{\alpha \in (0,1]} \sup_{\varphi \in \mathcal{D}_r} \frac{|u(\varphi_x^\delta)|}{\delta^\gamma}$$

where $\mathcal{D}_r = \{\varphi : \|\varphi\|_{C^r} \leq 1, \operatorname{supp} \varphi \subseteq B_1(0)\}$, and let

$$\mathcal{C}^\gamma_{\text{loc}} = \{u : [u]_{\gamma,K} < \infty \text{ for all compact } K\}.$$

As for $\gamma > 0$ with $\gamma = n + \gamma_0$ and $n \in \mathbb{N}$, define

$$[u]_{\gamma,K} = \sup_{x \in K} \sup_{\alpha \in (0,1]} \sup_{\varphi \in \mathcal{D}^n} \frac{|\langle u, \varphi_x^\delta \rangle|}{\delta^\gamma},$$

where $\mathcal{D}^n$ is the set of $\varphi \in \mathcal{D}$ such that $\int \varphi P(x)\, dx = 0$ for all polynomials P with $\deg P \leq n$. It requires proof to show that when $\gamma \neq \mathbb{Z}$, then these equivalent to the Hölder norms.

23 Norms and Schauder Estimates for Hölder-Zygmund Spaces

23.1 Equivalence of definitions of Hölder-Zygmund spaces

Recall that for our function spaces, we are using Hölder (or rather Hölder-Zygmund) spaces, and one of the simplest regularity estimates that are available for elliptic/parabolic PDE are the Schauder-type estimattes. For example, if K is smooth off of 0, with

1. $\operatorname{supp} K \subseteq B_1(0)$
2. $|\partial^k K(x)| \lesssim |x|^{\beta-d-|k|}$,

our Schauder estimates assert that

$$u \in \mathcal{C}^\alpha \implies K * u \in \mathcal{C}^{\alpha+\beta}.$$

Recall that for $\alpha < 0$, $\mathcal{C}^\alpha$ consists of $u \in \mathcal{D}'$ such that

$$[u]_{\alpha,K} = \sup_{x\in K} \sup_{\delta\in(0,1]} \sup_{\varphi\in\mathcal{D}_r} \frac{|u(\varphi_x^\delta)|}{\delta^\alpha} < \infty \qquad \text{for every compact } K,$$

where $\mathcal{D}_r$ is the set $\varphi \in \mathcal{D}$ with $\operatorname{supp}\varphi \subseteq B_1(0)$ and $\|\varphi\|_{C^r} \leq 1$. Here, r is the smallest integer such that $r + \alpha > 0$. As for $\alpha = n + \alpha_0$ with $n \in \mathbb{N}$ and $\alpha \in (0,1)$, $\mathcal{C}^\alpha$ consists of functions u such that $\partial^k u$ exists for $|k| \leq n$, and $\partial^k u \in \mathcal{C}^{\alpha_0}$, the Hölder continuous functions of exponent α_0.

We now define

$$[u]_{\alpha,K} = \sup_{x\in K} \sup_{\delta\in(0,1]} \sup_{\varphi\in\mathcal{D}^{(n)}} \frac{|u(\varphi_x^\delta)|}{\delta^\alpha} < \infty \qquad \text{for every compact } K,$$

where now $\mathcal{D}^{(n)}$ is the set of $\varphi \in \mathcal{D}$ such that $\operatorname{supp}\varphi \subseteq B_1(0)$, $\|\varphi\|_{L^\infty} \leq 1$, and $\int \varphi P = 0$ for every polynomial P of degree at most n. Let us write $\widehat{\mathcal{C}}^\alpha$ for the set of distributions u for which $[u]_{\alpha,K} < \infty$ for every compact K.

Proposition 23.1. *$\widehat{\mathcal{C}}^\alpha = \mathcal{C}^\alpha$, and they are isomorphic.*

Proof. $\mathcal{C}^\alpha \subseteq \widehat{\mathcal{C}}^\alpha$ by Taylor expansion. For the converse, assume that $u \in \widehat{\mathcal{C}}^\alpha$, and pick $\rho \in \mathcal{D}$ with $\int \rho = 1$. Recall that $u = \lim_{\varepsilon\to 0} u * \rho^\varepsilon$ and that $u * \rho^\varepsilon$ is a smooth function for each ε. Then using a telescoping sum, we can write

$$\begin{aligned} u - u(\rho_x^\delta) &= \sum_{n=0}^{\infty} (u(\rho_x^{2^{-n-1}\delta}) - u(\rho_x^{2^{-n}\delta})) \\ &= \sum_{n=0}^{\infty} (u(\rho_x^{2^{-n-1}\delta}) - u(\rho_x^{2^{-n}\delta})) \end{aligned}$$

$$= \sum_{n=0}^{\infty} u((\rho^{1/2} - \rho)_x^{2^{-n}\delta}).$$

Let us first assume that $n = 0$, i.e. $\alpha \in (0, 10$, so that $\rho^{1/2} - \rho \in \mathcal{C}^{(0)}$. This would allow us to assert that

$$|u((\rho^{1/2} - \rho)_x^{2^{-n}\delta})| \lesssim (2^{-n}\delta)^{\alpha}.$$

This, in particular, implies that the above sum is uniformly convergent, so u must be a function. On the other hand, this estimate implies that

$$|u(x) - u(\rho_x^{\delta})| \lesssim \delta^{\alpha}.$$

Finally,

$$\begin{aligned} |u(x) - u(y) &\lesssim \delta^{\alpha} + |u(\rho_x^{\delta}) - u(\rho_y^{\delta})| \\ &= \delta^{\alpha} + |u(\rho_x^{\delta} - \rho_y^{\delta})|. \end{aligned}$$

Observe that

$$(\rho - \rho_a)_x^{\delta} = \rho_x^{\delta} - \rho_{x+\delta a}^{\delta}.$$

Hence, if we choose $a = \frac{y-x}{\delta}$, then we get $\rho_x^{\delta} - \rho_y^{\delta}$. So we may select $\delta = |x - y|$ to assert

$$|u(\rho_x^{\delta} - \rho_y^{\delta})| \lesssim \delta^{\alpha} = |x - y|^{\alpha}.$$

Thus,

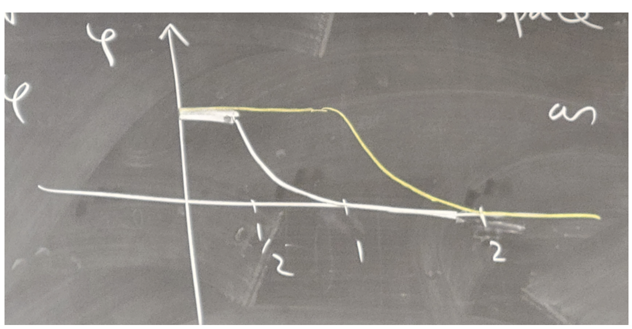
$$|u(x) - u(y)| \lesssim |x - y|^{\alpha},$$

as desired.

This completes the proof when $n = 0$. For higher n, we integrate by parts. For example, if $n = 1$, then we can show that $\partial_{x_i} u \in \mathcal{C}^{\alpha}$ for each i by the previous case. $\square$

23.2 Schauder estimates for Hölder-Zygmund spaces

Now we focus on our Schauder estimate. We use a Paley-Littlewood type expansion of the kernel K but in the space variable. To prepare for this, start with a smooth function φ with $\varphi = 1$on $(0, 1/2]$ and $\operatorname{supp} \varphi \subseteq [0, 1]$. Then set $\psi(x) - \varphi(x/2) - \varphi(x)$, so that $\operatorname{supp} \psi \subseteq [1/2, 2]$.



Further define $\psi_n(x) = \psi(2^n x)$. Observe that

$$\begin{aligned}\sum_{n=0}^{\infty} \psi_n(x) &= \sum_{n=0}^{\infty} (\varphi(2^{n-1}x) - \varphi(2^n x)) \\ &= \varphi(2^{-1}x),\end{aligned}$$

which equals 1 in $(0,1]$. Observe that $\operatorname{supp} \psi_n \subseteq [2^{-n-1}, 2^{-n+1}]$, and if we write

$$K = \sum_{n=0}^{\infty} (\underbrace{K(x)\psi_n(|x|)}_{K_{n-1}}) = \sum_{n=-1}^{\infty} K_n(x)$$

with $K_n \subseteq \{x : |x| \in [2^{-n-2}, 2^{-n}\} \subseteq B_{2^{-n}}(0)$, then by our assumption on K,

$$|K_n(x)| \lesssim 2^{-n(\beta-d)}, \qquad |\partial^k K_n(x)| \lesssim 2^{-n(\beta-|k|-d)}.$$

More conveniently, we can think of this as

$$K = \sum_{n=-1}^{\infty} 2^{-n\beta}(2^{n\beta} K_n),$$

where the part in the parentheses looks like a Dirac delta. Here, the kernel K is a function with $K \in C^\infty(\mathbb{R}^d \setminus \{0\})$, $\operatorname{supp} \subseteq B_1(0)$, $|\partial_k K(x)| \lesssim |x|^{\beta-d-|k|}$.

We can now study $u \mapsto u * K$, where $u \in \mathcal{C}^\alpha$. We wish to show that $u * K \in \mathcal{C}^{\alpha+\beta}$.

$$\begin{aligned}(u * K)(\xi) &= \text{“}\iint u(x-y)K(y)\xi(x)\,dx\,dy\text{”} \\ &= \text{“}\iint u(y)\widetilde{K}(y-x)\xi(x)\,dx\,dy\text{”} \\ &= u(\widetilde{K} * \xi) \\ &= \sum_{n=-1}^{\infty} 2^{-n\beta} u(2^{n\beta}\widetilde{K}_n * \xi).\end{aligned}$$

We wish to use $u \in \mathcal{C}^\alpha$ to get an estimate for $u(2^{n\beta}\widetilde{K}_n * \xi)$, where $\xi = \psi_a^\delta$ with ψ that satisfies certain conditions. For example, if $\alpha + \beta < 0$, then $\psi \in \mathcal{D}_r$; otherwise, we need some polynomial condition. We wish to use two pieces of information, $u \in \mathcal{C}^\alpha$ and the bounds on K_n. $\operatorname{supp} \xi \subseteq B_\delta(a)$, and $\operatorname{supp}(2^{n\beta} K_n) \subseteq B_{2^{-n}}(0)$, and we have the bounds $|\xi| \lesssim \delta^{-d}$ and $|2^{n\beta} K_n| \lesssim (2^{-n})^d$, which can be used to assert that $\operatorname{supp} \xi * K_n \subseteq B_{\delta+2^{-n}}(a)$ and

$$|u(2^{-\beta} K_n * \xi)| \lesssim (\delta + 2^{-n})^\alpha.$$

We are assuming $\alpha + \beta < 0$, so

$$
\begin{aligned}
|(u * K)(\xi)| &\lesssim \sum_n (\delta + 2^{-n})^\alpha 2^{-n\beta} \\
&= \sum_{n:2^{-n}<\delta} + \sum_{n:2^{-n}>\delta} \\
&\leq \delta^\alpha \sum_{2^{-n}<\delta} 2^{-n\beta} + \sum_{2^{-n}>\delta} 2^{-n(\alpha+\beta)} \\
&\lesssim \delta^\alpha \delta^\beta + \sum_{n \leq \log_2 \frac{1}{\delta}} (2^{-(\alpha+\beta)})^n \\
&\lesssim \delta^{\alpha+\beta} + (2^{-(\alpha+\beta)})^{\log_2 \frac{1}{\delta}} \\
&= 2\delta^{\alpha+\beta},
\end{aligned}
$$

as desired.

Next, let us examine the case of $\alpha + \beta > 0$. In this case, we assume that $\int \psi P \, dx = 0$, P is a polynomial, and $\deg P \leq \alpha + \beta$. We need to use the latter condition to improve our second bound (even when $\alpha < 0$). This can be achieved by subtracting a suitable polynomial P from $2^{n\beta} K_n$ because ξ is orthogonal to such polynomials. We omit the rest of the details.

24 Setup for Solving the KPZ Equation

24.1 Kernel of the KPZ equation

Last time, we showed that if $u \in \mathcal{C}^\alpha$, then $u * K \in \mathcal{C}^{\alpha+\beta}$, where K is a function that has the following properties:

(i) $\operatorname{supp} K \subseteq B_1(0)$, and K is smooth off of 0.

(ii) $|\partial^k K(x)| \lesssim |x|^{\beta-d-|k|}$.

For example, when K is the kernel of $(-\Delta)^{-1}$ and $d \geq 3$, then we have our estimate for $\beta = 2$, except that its kernel $c_0|x|^{2-d}$ is not of compact support. However, we can express our kernel as $K + \widehat{K}$, with K as above and $\widehat{K}$ a smooth function so that $u * \widehat{K}$ is smooth. Moreover, instead of convolution, we can also integrate againsta kernel $K(x, y)$, and for our Schauder estimate, we need K to behave smoothly away from the diagonal, and near the diagonal as above.

For our KPZ equation, we need a Schauder estimate for the operator $(\partial_t - \Delta)^{-1}$. Its kernel, $K(x,t) := (4\pi t)^{-d/2} e^{-|x|^2/(4t)} \mathbb{1}_{\{t>0\}}$ does not look like what we have had so far. Though we can achieve a similar claim with identical proof, provided that we follow the parabolic scaling, treating time as 2.

For one thing, we may use the metric

$$d((x,t),(y,s)) = |(x-y, t-s)|_{\text{par}} = |x-y| + \sqrt{|t-s|}$$

and denote

$$\widetilde{\varphi}^{\delta}_{(y,s)}(x,t) = \frac{1}{\delta^{d+2}} \varphi\left(\tfrac{x-y}{\delta}, \tfrac{t-s}{\delta^2}\right),$$

where the $\sim$ means that we are using parabolic scaling. We can also discuss the size of a multiindex by

$$k = (k_1, \ldots, k_d, \underbrace{k_{d+1}}_{\text{time variable}}), \qquad |k|_{\text{par}} = k_1 + \cdots + k_d + 2k_{d+1}.$$

With these conventions, we may take a kernel $K(x,t)$ and assume

$$|\partial^k K(x,t)| \lesssim |(x,t)|_{\text{par}}^{\beta-(d+2)-|k|_{\text{par}}}.$$

Moreover, if $\alpha < 0$, then $\widetilde{\mathcal{C}}^\alpha(\mathbb{R}^{d+1})$ would consist of distributions F such that

$$[F]_{\alpha,K} = \sup_{(x,t)\in K} \sup_{\delta\in(0,1]} \sup_{\varphi\in D_r} \frac{|F(\widetilde{\varphi}^{\delta}_{(x,t)})|}{\delta^\alpha} < \infty.$$

In particular, we will have our Schauder estimate for such a kernel K, in the sense that if $u \in \widetilde{\mathcal{C}}^\alpha$, then $K * u \in \widetilde{\mathcal{C}}^{\alpha+\beta}$.

For example, the bound above holds for the heat kernel $K(x,t) = (4\pi t)^{-d/2} e^{-|x|^2/(4t)} \mathbb{1}_{\{t>0\}}$ for $\beta = 2$. Here are some details:

$$t^{-d/2} e^{-\frac{1}{4}(\frac{|x|}{\sqrt{t}})^2} \lesssim (|x| + \sqrt{t})^{2-(d+2)} = (|x| + \sqrt{t})^{-d} = \left(\frac{|x|}{\sqrt{t}} + 1\right)^{-d} t^{-d/2}.$$

This is equivalent to

$$e^{-z^2/4} \lesssim (z+1)^{-d}, \quad \text{or} \quad (z+1)^d \lesssim e^{z^2/4}.$$

Taking $\frac{d}{dt}$ gives

$$t^{-d/2-1} e^{-\frac{1}{4}(\frac{|x|}{\sqrt{t}})^2} \lesssim (|x| + \sqrt{t})^{-d-2} = (\sqrt{t})^{-d-2} \left(\frac{|x|}{\sqrt{t}} + 1\right)^{-d-2}.$$

Then we can expand the left hand side to get that

$$\frac{t^{-d/2}}{t} \frac{|x|^2}{t} e^{-\frac{1}{4}(\frac{|x|}{\sqrt{t}})^2} \lesssim (\sqrt{t})^{-d-2} \left(\frac{|x|}{\sqrt{t}} + 1\right)^{-d-2}.$$

Then perform induction.

24.2 Regularity considerations for white noise

Return to the KPZ equation

$$\begin{cases} h_t = \Delta h + |h_x|^2 + \xi \\ h(x,0) = h^0(x), \end{cases}$$

which can be written as

$$h = K * h^0 + K * (|h_x|^2 + \xi),$$

where ξ is the white noise. Let us examine the regularity of ξ. Recall that $\xi(x,t)$ is Gaussian with

$$\mathbb{E}[\xi(\varphi)] = 0, \qquad \mathbb{E}[(\xi(\varphi))^2] = \int \varphi^2 \, dx \, dt.$$

Hence,

$$\begin{aligned} \mathbb{E}[(\xi(\widetilde{\varphi}^\delta_{(x,s)}))^2] &= \int (\widetilde{\varphi}^\delta_{(x,s)})^2 \, dx \, dt \\ &= \int \left(\frac{1}{\delta^{d+2}}\right)^2 \varphi(\tfrac{y-x}{\delta}, \tfrac{t-s}{\delta}) \, dt \, dy \\ &= \delta^{-(d+2)} \int \varphi^2. \end{aligned}$$

We learn that

$$(\mathbb{E}[|\xi(\widetilde{\varphi}^{\delta}_{(x,s)})|^2]])^{1/2} = \delta^{-(d+2)/2}\|\varphi\|_{L^2},$$

hence

$$(\mathbb{E}[|\xi(\widetilde{\varphi}^{\delta}_{(x,s)})|^{2q}])^{1/(2q)} = c_q\delta^{-(d+2)/2}\|\varphi\|_{L^2}.$$

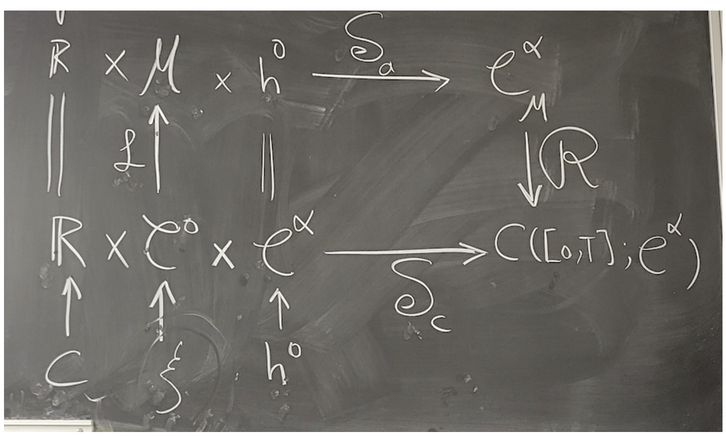
One can show that if ξ is any random Schwartz distribution with $(\mathbb{E}[(\xi(\widetilde{\varphi}^{\delta}_{(x,s)}))^{2q}])^{1/(2q)} \lesssim \delta$, then $\xi \in \widehat{\mathcal{C}}^{-\alpha-1/(2q)}$ as in Kolmogorov's theorem. Accepting this for now, we learn that $|xi \in \widetilde{\mathcal{C}}^{-(d+2)/2-\varepsilon}(\mathbb{R}^{d+1})$ for any $\varepsilon > 0$. Here, we are using parabolic scaling. As a result, we can use our Schauder estimate to assert that if K is the heat kernel, then $K * \xi \in \widetilde{\mathcal{C}}^{-d/2+1-\varepsilon} =: \widetilde{\mathcal{C}}^{-d/2+1}(\mathbb{R}^{d+1})$. For example, when $d = 1$, then $K * \xi \in \widetilde{\mathcal{C}}^{1/2-}$, which really means $\mathcal{C}^{1/2-}$ in space and $\mathcal{C}^{1/4-}$ in time.

24.3 Strategy for solving the KPZ equation

We wish to solve the KPZ equation

$$\begin{cases} h_t = \Delta h + |h_x|^2 + \xi + C \\ h(x,0) = h^0(x), \end{cases}$$

where we should really solve this as we vary the constant C. If we choose a smooth function for ξ, then we can solve this equation classically. Let us write $\mathcal{S}_c(C, \xi, h^0)$ for the classical solution. Here is the picture of what this will look like when with lift it:



Here is our strategy: We build a regularity structure that would allow us to solve the KPZ equation in abstract space, once we have a recipe for the meaning of h_x^2 so that this abstract solution is indeed a continuous operator. However, we still need to build our regularity structure. For this, let us now focus on our operator $F \mapsto F * K$, where K is the heat kernel. We claim that if our regularity structure (A, G, T) is "rich enough," then we can build an operator $\mathcal{K} : \mathcal{C}_M^\gamma \to \widetilde{\mathcal{C}}_M^{\gamma+2}$ such that

$$\mathcal{R}(\mathcal{K}f) = K * \mathcal{R}f.$$

Here, $\widetilde{\mathcal{C}}_M^\gamma = \{f : \mathbb{R}^{d+1} \to \bigoplus_{\alpha<\gamma} T_\alpha : |\Gamma_{yx}f(x) - f(y)| \lesssim |x-y|_{\text{par}}^{\gamma-\alpha}\}$, and we have the reconstruction theorem:

Theorem 24.1 (Reconstruction theorem)**.**

$$|(\mathcal{R}f - \Pi_x f(x))(\widetilde{\varphi}_x^\delta)| \lesssim \delta^\gamma.$$

As a warm-up, first let us assume that the kernel K is smooth (no singularity at 0), and assume that our regularity structure has a sector consisting of polynomials: a subspace $\overline{T}$ of T such that $T_n = \langle X^k : |k| = n\rangle$. Then, since $K * F$ is smooth for any distribution F,

$$(\mathcal{K}f)(a) = \sum_k \frac{1}{k!}(\partial^k K * \mathcal{R}f)X^k.$$

Next time, we will cover the general case.

25 Multiplication of Abstract Candidates

25.1 Motivation: Necessity of multiplication in the solution for KPZ

We have formulated a general strategy for treating (subcritical) ill-posed PDEs. Our strategy is to isolate the bad parts, interpret them in an abstract setting, come up with an abstract solution, and use reconstruction theory to give an actual solution. We now would like to describe this strategy in detail for the KPZ equation

$$\begin{cases} h_t = h_{xx} + h_x^2 + \xi - C \\ h(x,0) = h^0(x), \end{cases}$$

where $h : \mathbb{R} \times [0,T] \to \mathbb{R}$ and ξ is white noise. We would like to construct a solution as a fixed point of a suitable operator

$$h = P * (h_x^2 + \xi - C) + P * h,$$

where P is the heat kernel, and by $P * h$, we mean h integrated against P. Last time, we discussed Schauder-type estimates that give regularity for the expression $f \mapsto P * f$ in the sense that if $f \in \mathcal{C}^\alpha_{\text{par}}$, then $P * f \in \mathcal{C}^{\alpha+2}_{\text{par}}$. We argued last time that there would be a multi-layer type Schauder estimate that is applicable for general regularity structure under some natural conditions. We would be able to come up with an operator P such that if $f \in \mathcal{C}^\gamma_M$ ($f : \mathbb{R}^d \to \bigoplus_{\alpha<\gamma} T_\alpha$), then our reconstruction theorem would turn f into $\mathcal{R}f$, which is a distribution that is well approximated by $\Pi_x f(x)$ near x. Moreover,

$$\mathcal{R}(\mathcal{P}f) = P * \mathcal{R}f,$$

and, as we will see later, we can rewrite $\mathcal{P} = \mathscr{I} + \widehat{\mathscr{I}}$, where $\widehat{\mathscr{I}}$ would be a polynomial like dealing with the Taylor approximation of the smooth part of $\mathcal{P} * \mathcal{R}f$. So, in some sense, only the $\mathcal{I}$ part of $\mathcal{P}$ would capture the true nature of the singularity of the kernel $\mathcal{P}$.

To solve this heat kernel equation, we first formulate an abstract variant that can be solved as a fixed point of some nice continuous operator. In other words, the solution we are looking for can be expressed as $h = \mathcal{R}H$, where H solves an equation of the form

$$\begin{aligned} H &= \mathcal{P}((\partial H)^2 + \Xi) + (\mathcal{P} * h^0)\mathbf{1} \\ &= \mathscr{I}((\partial H)^2 + \Xi) + \widehat{\mathscr{I}}((\partial H)^2 + \Xi) + (P * h^0)\mathbf{1}. \end{aligned}$$

Here, Ξ represents ξ in the abstract setting,[10] ∂ represents the spatial derivative (should satisfy $\Pi_a(\partial\tau) = \frac{\partial}{\partial x}(\Pi_a\tau)$), and $(\partial H)^2$ is a candidate for $(\partial H)(\partial H)$.

What do we mean by multiplying two members of our Banach space T? Basically, our regularity structure must be rich enough so that such multiplication can be carried out. Here is our general definition for any multiplication type operation.

[10] Professor Rezakhanlou is using Θ in the lectures instead of Ξ because he doesn't like letters with 3 connected components. On a keyboard, I have no such objection.

Definition 25.1. Given a regularity structure (A, T, G) and two sectors V and $\overline{V}$ (i.e. $V = \bigoplus_{\alpha \in A} V_\alpha$, with V_α a subspace of T_α and with each V_α invariant under G), we say $\star : V \times \overline{V} \to T$ sending $(\tau, \overline{\tau}) \mapsto \tau \star \overline{\tau}$ is a **multiplication** if the following conditions are true:

1. $\star$ is bilinear.

2. If $\tau \in V_\alpha$ and $\overline{\tau} \in \overline{V}_{\alpha'}$, then $\tau \star \overline{\tau} \in T_{\alpha + \alpha'}$.

3. If $\Gamma \in G$, then $\Gamma(\tau \star \overline{\tau}) = (\Gamma \tau) \star (\Gamma \overline{\tau})$.

Example 25.1. Take $\tau = X^k$ and $\overline{\tau} = X^{\overline{k}}$. Then $\tau \star \overline{\tau} = X^{k + \overline{k}}$.

Recall that $f \in \mathcal{C}^\gamma_M$ means that $\|f(x) - \Gamma_{x,y} f(y)\|_\beta \lesssim |x-y|^{\gamma - \beta}$.

Proposition 25.1. *Let $f_1 \in \mathcal{C}^\gamma_M$ with $f_1 : \mathbb{R}^d \to R$ and $f_1(x) \in \bigoplus_{\alpha_1 \le \alpha < \gamma_1} T_\alpha$, and Let $f_2 \in \mathcal{C}^\gamma_M$ with $f_2 : \mathbb{R}^d \to R$ and $f_2(x) \in \bigoplus_{\alpha_2 \le \alpha < \gamma_2} T_\alpha$. Define $(f_1 \star f_2)(x) = f_1(x) \star f_2(x)$. Then $f_1 \star f_2 \in \mathcal{C}^\gamma_M$ with $(\gamma_1 + \alpha_2) \min(\gamma_2 + \alpha_1)$.*

Proof.

$$\begin{aligned}
\|\Gamma_{x,y}(f_1 \star f_2)(y) - (f_1 \star f_2)(x)\|_\beta &= \|(\Gamma_{x,y} f_1(y)) \star (\Gamma_{x,y} f_2(y)) - f_1(x) \star f_2(x)\|_\beta \\
&= \|(\Gamma_{x,y} f_1(y) - f_1(x)) * (\Gamma_{x,y} f_2(y) - f_2(x)) \\
&\quad + (\Gamma_{x,y} f_1(y) - f_1(x)) \star f_2(x) \\
&\quad + f_1(x) \star (\Gamma_{x,y} f_2(y) - f_2(x))\| \\
&\lesssim \sum_{\beta_1 + \beta_2 = \beta} (|x-y|^{\gamma_1 - \beta_1} |x-y|^{\gamma_2 - \beta_2} \\
&\qquad + |x-y|^{\gamma_1 - \beta_1} + |x-y|^{\gamma_2 - \beta_2}) \\
&= \sum_{\beta_1 + \beta_2 = \beta} |x-y|^{\gamma_1 + \gamma_2 - \beta} + |x-y|^{\gamma_1 + \beta_2 - \beta} \\
&\qquad + |x-y|^{\gamma_2 + \beta_1 - \beta} \\
&\lesssim |x-y|^{[(\gamma_1 + \alpha_2) \wedge (\gamma_2 + \alpha_1)] - \beta}.
\end{aligned}$$

□

25.2 Basis for a multiplicatively closed regularity structure

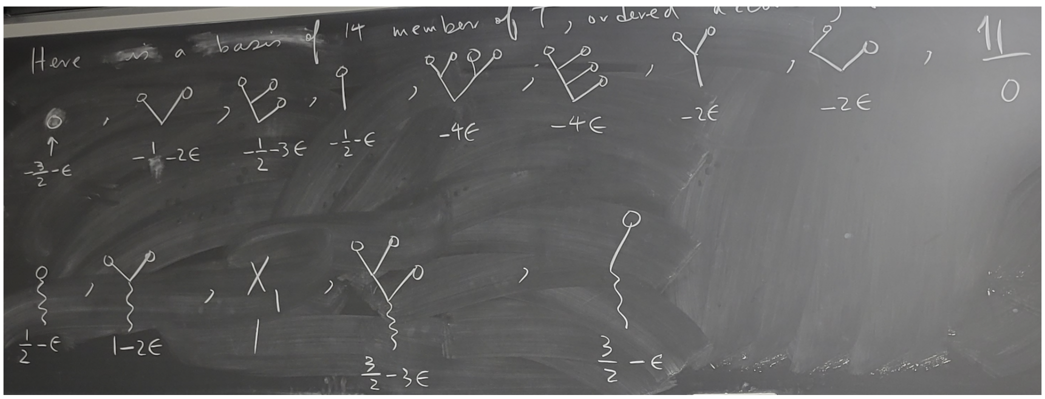
We now use our fixed point equation to guess what regularity structure we need. As is done in mathematical physics, we will use graphical notation.[11] We use $\circ$ for Ξ and $\wr$ for the operator $\mathscr{I}$, so that $\mathscr{I}(\Xi)$ is $\wr^\circ$. Because of this, we would also have a component which is like $\mathscr{I}((\partial \wr^\circ)^2)$. This involves $\partial \mathscr{I} =: \mathscr{I}'$. Graphically, we use $|$ for $\mathscr{I}'$ so that $\partial \wr^\circ = |^\circ$.

[11] I hate this.

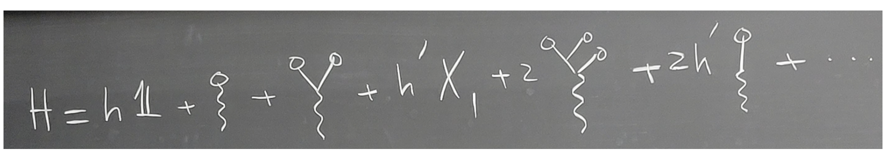
Here is a table of some of the terms:

degree	expression	model
0	$\mathbf{1}$	constants
$-\frac{3}{2}-\varepsilon$	$\Xi, \circ$	ξ
$\frac{1}{2}-\varepsilon$	$\mathscr{I}(\Xi), \wr^\circ$	$P * \xi$
$-\frac{1}{2}-\varepsilon$	$\mathscr{I}'(\Xi), \vert^\circ$	$(P * \xi)_x$
$-1-2\varepsilon$	$(\mathscr{I}'(\Xi))^2,\ {}^o\vee^o$	$(P * \xi)_x^2$
$1-2\varepsilon$	$\mathscr{I}((\mathscr{I}'(\Xi))^2),$	$P * (P * \xi)_x^2$
1	X_1	$x - a$

Here is a basis of 14 members of T (ignoring polynomials of higher order), ordered according to their degrees:[12]



This will give us



[12] There's no way I'm trying to recreate all of these in LaTeX.

26 Fixed Point Operators for Solving Abstract Regularity Structure PDEs

26.1 Fixed point operators for solving our ill-posed PDEs

We are interested in ill-posed problems like:

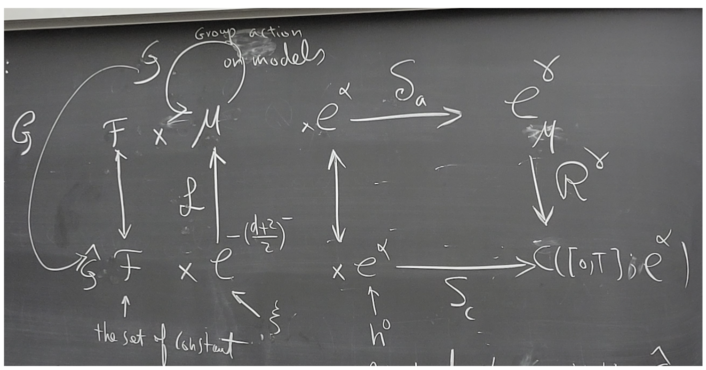
$$h_t = \Delta h + |h_x|^2 + \xi - C,$$

where ξ is white noise. This is subcritical iff $d \leq 1$. If we have

$$u_t = \Delta u - u^3 + \xi + C_1 + C_2 u,$$

this model is subcritical if $d \leq 3$. Here, we can vary the constants, so we are dealing with a family $\mathcal{F}$ of differential equations.

The general strategy for subcritical models is summarized in the following diagram:



We need to find a group action $\mathcal{G}$ on our model so that if $\xi^\varepsilon = \xi * \rho^\varepsilon$, then

$$\lim_{\varepsilon \to 0} \mathcal{S}_a M_\varepsilon(\mathcal{L}(\xi^\varepsilon))$$

exists, where M_ε is a suitable family of members of $\mathcal{G}$. This $\mathcal{G}$ would lead to a suitable $\widehat{\mathcal{G}}$ on $\mathcal{F}$. In our stochastic setting, since our distributions are all Gaussian, Wick's trick would allow us to discover what $\mathcal{G}$ is. Let us now focus on constructing $\mathcal{S}_a$ as we did last time.

As we discussed before, we consider the weak formulation

$$h_t = p * (|h_x|^2 + \xi) + \overline{h},$$

where p is the heat kernel and $\overline{h}$ solves the heat equation:

$$\begin{cases} \overline{h}_t = \overline{h}_{xx} \\ \overline{h}(x,0) = h^0(x). \end{cases}$$

For the other problem, we have

$$u = p * (-u^3 + \xi) + \overline{u}$$

with

$$\begin{cases} \overline{u}_t = \Delta \overline{u} \\ \overline{u}(x,0) = u^0(x). \end{cases}$$

Last time, we argued that $f \mapsto p * f$ can be lifted to a suitable operator $\mathcal{K}$ that can be decomposed as $\mathcal{K} = \mathscr{I} + \widehat{\mathcal{K}}$, where $\widehat{\mathcal{K}}$ is polynomial like annd $\mathscr{I}$ is somewhat local. Ideally, we could like to have this: An operator $\mathcal{K} : T \to T$ or $\mathcal{K} : \mathcal{C}^\alpha \to \mathcal{C}^{\alpha+2}$ so that

$$\begin{cases} \Pi_x(\mathcal{K}\tau) = p * \Pi_x \tau \quad (f \in \mathcal{C}^\alpha \ \Pi_x(\mathcal{K}f)(x) = p * \Pi_x f(x)) \\ \Gamma \mathcal{K} \tau = \mathcal{K} \Gamma \tau. \end{cases}$$

Such $\mathcal{K}$ would not exist. Here is the problem: if $\tau \in T_\alpha$, then $|(\Pi_x \tau)(\varphi_x^\delta)| \lesssim \delta^\alpha$. So $\mathcal{K}\tau \in T_{\alpha+2}$, and we must have an estimate of the form $|(\Pi_x(\mathcal{K}\tau))(\varphi_x^\delta)| \lesssim \delta^{\alpha+2}$. The problem is that in general, there is no reason for $p * \Pi_x \tau$ to vanish like $\delta^{\alpha+2}$ near the point x. This can be resolved if we subtract a suitable Taylor expansion. Motivated by this, we may define $\mathscr{I}$ by the following recipe. If $\tau \in T_\alpha$,

$$\Pi_x(\mathscr{I}\tau)(y) = p * \Pi_x \tau(y) - \sum_{k:|k|<\alpha+2} \frac{\partial^k (p * \Pi_x \tau)}{k!}(y-x)^k.$$

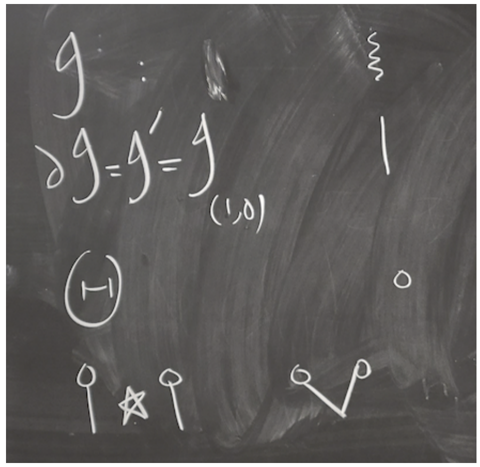
Because of this, we do not expect to have $\Gamma \mathscr{I} \tau = \mathscr{I} \Gamma \tau$, but we do have that $(\Gamma \mathscr{I} - \mathscr{I} \Gamma)(\tau)$ is in a sector of polynomials. (This should be compared with the differentiation operator: If "∂" is the lift of $\frac{\partial}{\partial x_1}$, then we do expect $\Pi_x(\partial \tau) = \frac{\partial}{\partial x_1}(\Pi_x \tau)$ and $\partial \Gamma = \Gamma \partial$.)

26.2 Using graphical notation with regularity structures to solve abstract PDEs

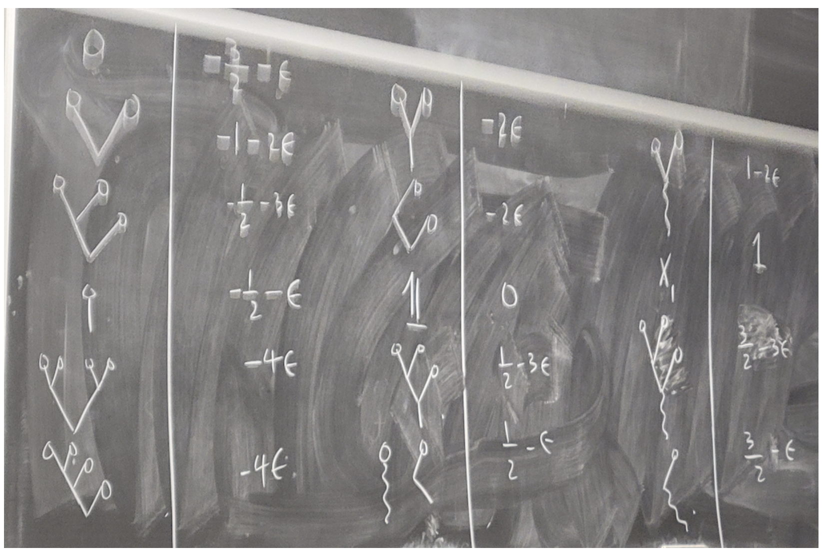
In the abstract version,

$$\begin{cases} H = \mathscr{I}((\partial H)^2 + \Xi) + \text{Polynomial part from } \widehat{K} + \overline{h}\mathbf{1} \\ \mathcal{U} = \mathscr{I}(\Xi - U^3) + \text{Polynomial part} + \overline{u}\mathbf{1}, \end{cases}$$

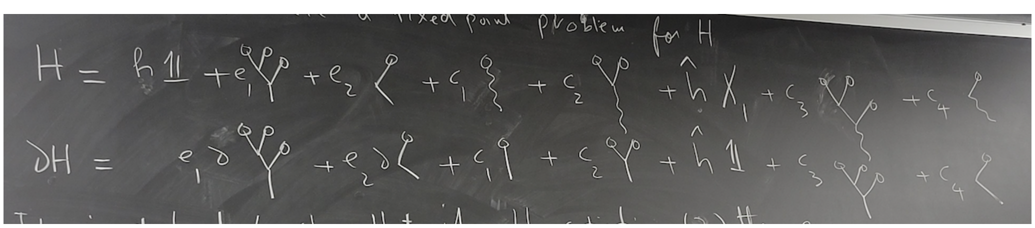
where Ξ represents white noise, we use the following graphical notation.[13]



Here are all the terms we need to discuss H:

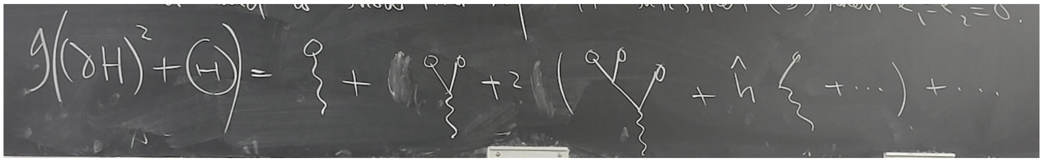


We want to formulate a fixed point problem for H.

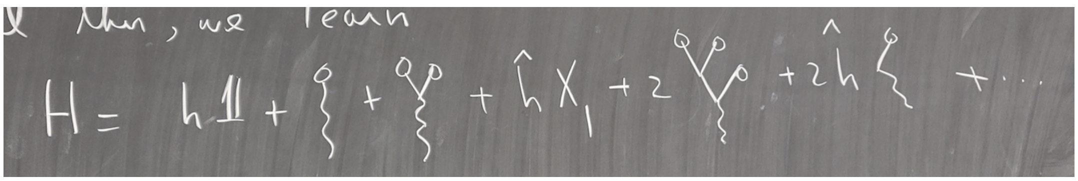


[13]I've decided to stop trying to type out any graphical notation. From here on out, it will all be pictures.

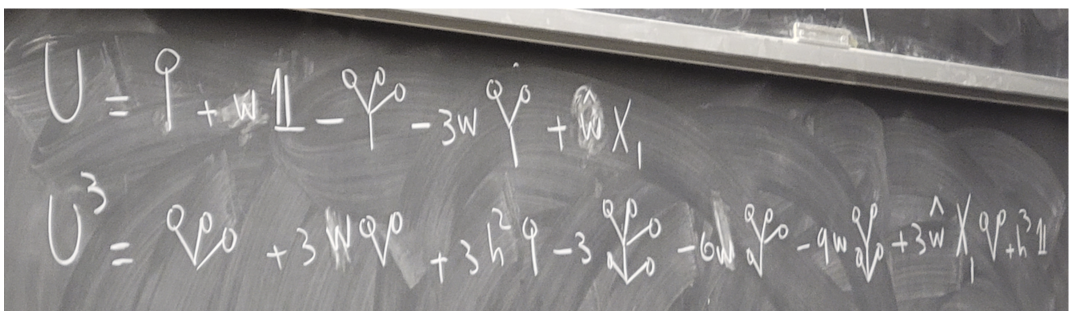
It can be shown that if H satisfies the abstract equation, then $e_1 = e_2 = 0$.



Here, we have noted that by comparing coefficients, we can see that $c_1 = c_2 = 1$, $c_3 = 2$, and $c_4 = \widehat{h}$. From all this, we learn that

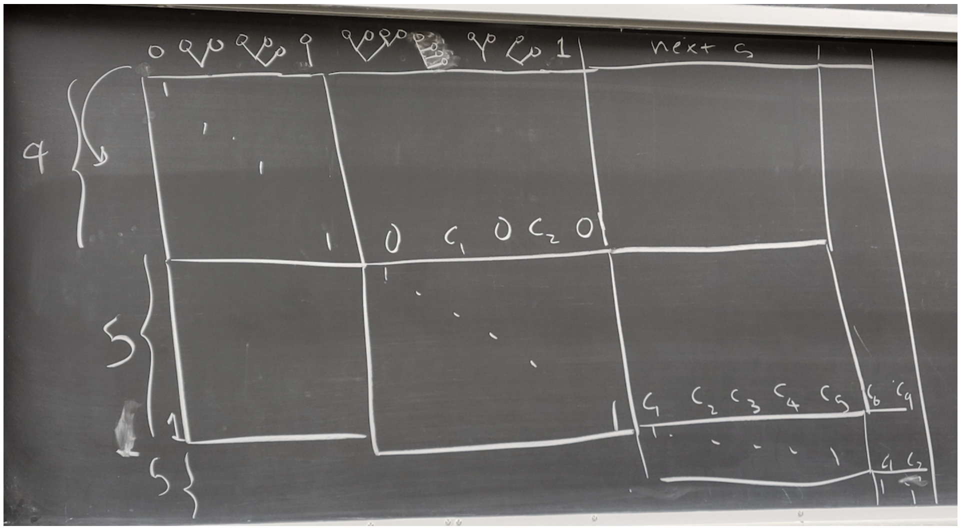


We can play a similar game with the abstract equation for $\mathcal{U}$. To have simpler notation, we write $|$ for $\mathscr{I}$ (instead of $\wr$). We get



We still need to find the group G. This is a suitable set of transformations $\Gamma : T \to T$.

This group of 16×16 matrices is 7-dimensional.



27 Algebraic Structure in Our Regularity Structure

27.1 Products structures in rough path theory

What kind of algebra leads to our group G and the form $\Gamma_{x,y}$? We first discuss Hopf algebras.[14] In fact, Connes and Kreimer have observed that a suitable Hopf algebra on the polynomials of "decorated trees" can be used to explain renormalization phenomena in quantum field theory and Feynman diagrams.

To motivate the role of such algebras, let us go back to our rough path theory first. Indeed, if we have a path $x = (x^1, \dots, x^\ell) : [0,T] \to \mathbb{R}^\ell$, then we need a candidate for

$$\langle \mathbf{x}(s,t), e_{i_1 \cdots i_r} \rangle = \int_s^t \int_s^{s_1} \int_s^{s_r} dx^{i_1}(x_1) \cdots dx^{i_r}(s_r) =: a_{i_1,\dots,i_r}.$$

What we have in mind is that we choose a lift for the path x that is tensor-valued, and this condition yields the $e_{i_1 \cdots i_r}$ component of such a tensor. More precisely, if the space of tensors $T(\mathbb{R}^d) = \mathbb{R} \oplus \ell^\ell \oplus \cdots \oplus (\mathbb{R}^\ell)^{\otimes n} \oplus \cdots$, then

$$\overline{x}(s,t) = \sum_{i_1,\dots,i_r} a_{i_1,\dots,i_r} e_{i_1,\dots,i_k},$$

where $e_{i_1,\dots,i_r} = e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_r}$ and the $e_1 \cdot e_\ell$s form a basis for $\mathbb{R}^\ell$. As we have seen before, if $x \in \mathcal{C}^\alpha$ with $\frac{1}{n} < \alpha \leq \frac{1}{n-1}$, then we can truncate our tensor at level $n-1$. Note that if $x \in \mathcal{C}^\alpha$, then the type of regularity we have is

$$|\langle \mathbf{x}(s,t), e_{i_1 \cdots i_r} \rangle| \lesssim |s-t|^{r\alpha}.$$

Recall that Chen's relation becomes $\mathbf{x}(s,u) \otimes \mathbf{x}(u,t) = \mathbf{x}(s,t)$, which allows us to only consider $\mathbf{x}(0,t)$ because $\mathbf{x}(s,t) = \mathbf{x}(0,s)^{-1} \otimes \mathbf{x}(0,t)$.

Recall that when $\frac{1}{3} \leq \alpha < \frac{1}{2}$, for a metric path, we have

$$\mathbb{X}(s,t) + \mathbb{X}^*(s,t) = \mathbf{x}(s,t) \otimes \mathbf{x}(s,t),$$

or

$$\langle \mathbf{x}(s,t), e_{i,j} + e_{j,i} \rangle = \langle \mathbf{x}(s,t), e_i \rangle \langle \mathbf{x}(s,t), e_j \rangle.$$

However, for low α, the geometric condition becomes

$$\langle \mathbf{x}(s,t), e_{i_1,\dots,i_r} \rangle \langle \mathbf{x}(s,t), e_{j_1 \cdots j_\ell} \rangle = \langle \mathbf{x}(s,t), e_{i_1,\dots,i_r} \sqcup\!\sqcup e_{j_1,\cdots,j_\ell} \rangle,$$

where $a \sqcup\!\sqcup b$ means the **shuffle product** of a and b:

$$e_{i_1 \cdots i_r} \sqcup\!\sqcup e_{j_1 \cdots j_\ell} = \sum e_{k_1 \cdots k_{r+\ell}},$$

where $k_1, \cdots, k_{r+\ell}$ is obtained from $i_1, \dots, i_r, j_1, \dots, j_\ell$ by interleaving them without changing the original order. So there are exactly $\frac{(k+\ell)!}{k!\ell!}$ many terms.

[14] Historically, Hopf was studying homology and cohomology on Lie groups, where the additional multiplication of the Lie group gave extra algebraic structure to the homology.

Example 27.1.

$$e_i \sqcup\!\sqcup e_j = e_{i,j} + e_{j,i}$$

Example 27.2.

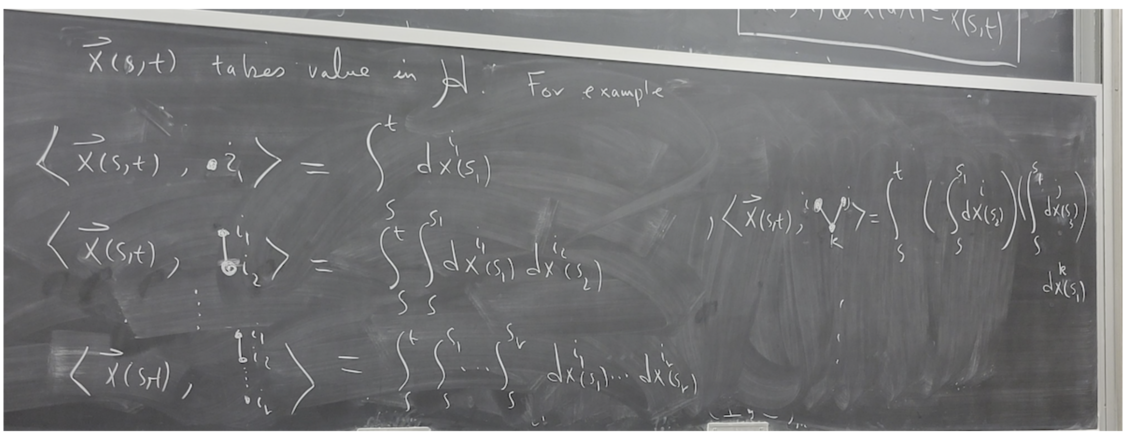
$$e_i \sqcup\!\sqcup e_{j,k} = e_{i,j,k} + e_{j,i,k} + e_{j,k,i}$$

In summary, for a geometric path, we have two products on $T(\mathbb{R}^\ell)$:

$$\begin{cases} \mathbf{x}(s,u) \otimes \mathbf{x}(u,t) = \mathbf{x}(s,t) \\ \langle \mathbf{x}(s,t), a \rangle \langle \mathbf{x}(s,t), b \rangle = \langle \mathbf{x}(s,t), a \sqcup\!\sqcup b \rangle. \end{cases}$$

How about for a nongeometric path? This has been worked out by Gubinelli and involves Connes-Kreimer's Hopf algebra (the theory of branched paths). In this case, the right space is not the tensor algebra, rather the algebra of polynomials of decorated binary trees with decorations/labels selected from $\{1, \ldots, \ell\}$. Write $\mathcal{H}$ for this space. Now $\underline{\mathrm{x}}(s,t)$ takes values in $\mathcal{H}$.

Example 27.3.



What happens to Chen's relation? There is an other product, the convolution product $\star$ that would allow us to represent Chen's relation as

$$\mathbf{x}(s,u) \star \mathbf{x}(u,t) = \mathbf{x}(s,t).$$

Before we define this, let us discuss the notion of Hopf algebras first.

27.2 Hopf algebras

Let $\mathcal{H}$ be an algebra with unit $\mathbf{1}$ and product $\cdot$. Also suppose we have an algebra on $\mathcal{H}^*$. Let us write $\langle \cdot, \cdot \rangle : \mathcal{H}^* \times \mathcal{H} \to \mathbb{R}$ for the pairing between $\mathcal{H}$ and $\mathcal{H}^*$. We write $f \star g$ for the product on $\mathcal{H}^*$ and $\mathbf{1}^*$ for its unit. We use pairing to turn $\star$ into a "coproduct" on $\mathcal{H}$.

Also, inversion in $\mathcal{H}^*$ can be translated into a suitable notion on $\mathcal{H}$. When this is done successfully, we have a **Hopf algebra**.

Here is the idea: if $f, g \in \mathcal{H}^*$ and $h \in \mathcal{H}$, then

$$\underbrace{\langle f \star g, h\rangle}_{\text{pairing of } \mathcal{H}^*, \mathcal{H}} \quad = \quad \underbrace{\langle f \otimes g, C(h)\rangle}_{\text{pairing of } \mathcal{H}^* \otimes \mathcal{H}^*, \mathcal{H} \otimes \mathcal{H}} \quad .$$

If we find such a C, then we have a bialgebra. Here, C is our coproduct $C : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$.

We now try the idea of the inverse: If $f \in \mathcal{H}^*$ is invertible, then $f \star f^{-1} = \mathbf{1}^*$. We wish to find an operator $\mathcal{S} : \mathcal{H} \to \mathcal{H}$ so that $\mathcal{S}^* : \mathcal{H}^* \to \mathcal{H}^*$ is exactly $\mathcal{S}^*(f) = f^{-1}$. If such an operator exists, then we should have

$$\begin{aligned}
\langle \mathbf{1}^*, h\rangle &= \langle f \star \mathcal{S}^* f, h\rangle \\
&= \langle f \otimes \mathcal{S}^* f, C(h)\rangle \\
&= \langle ((\mathrm{id} \otimes \mathcal{S}^*)(f \otimes f), C(h)\rangle \\
&= \langle f \otimes f, (\mathrm{id} \otimes \mathcal{S}) C(h)\rangle.
\end{aligned}$$

Assume $\langle f, \mathbf{1}\rangle = 1$. Then if we require $(\mathrm{id} \otimes \mathcal{S})C(h) = \langle \mathbf{1}^*, h\rangle \mathbf{1}$, then we have our $\mathcal{S}$. And if such a C and $\mathcal{S}$ exist, we have a Hopf algebra.

Definition 27.1. A **Hopf algebra** is an algebra $\mathcal{H}$ with a coproduct $C : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ and an operator $\mathcal{S} : \mathcal{H} \to \mathcal{H}$ such that

$$(\mathrm{id} \otimes \mathcal{S})C(h) = \langle \mathbf{1}^*, h\rangle \mathbf{1}.$$

Example 27.4. Let $\mathcal{H}$ be the algebra generated from $(\partial_i = \frac{\partial}{\partial x_i} : i = 1, \dots, d)$ with the product given by the composition: $D = \partial_{i_1, \dots, i_k} = \partial_{i_1} \cdots \partial_{i_k}$. $\mathcal{H}^*$ is the space of smooth functions with pointwise multiplication. The pairing is $\langle f, D\rangle = (Df)(0)$.

We claim that this is a Hopf algebra. One can show by integration by parts that

$$C(\partial_i) = \mathrm{id} \otimes \partial_i + \partial_i \otimes \mathrm{id},$$

$$C(\partial_i \partial_j) = \mathrm{id} \otimes \partial_{i,j} + \partial_i \otimes \partial_j + \partial_j \otimes \partial_i + \partial_{i,j} \otimes \mathrm{id}\,.$$

28 Hopf Algebras for Constructing Regularity Structures

28.1 Building up to Hopf algebras

Here is the algebraic part of the story: We learn how to build an important group of transformations $\{\Gamma_g : g \in G_0\} = G$ for a Hopf algebra. It is this group that yields our group G in our regularity structure. Here are the first few steps:

Definition 28.1.

1. **Algebra**: Given a field k and a k-vector space A, by a **product**, we mean a linear map $m : A \otimes A \to A$ that is associative. We also have a unit $\mathbf{1} \in A$, which we also write as $\mathbf{1} : k \to A$ by $\mathbf{1}(\lambda) = \lambda \mathbf{1}$.

2. **Coalgebra**: With A as above, we now have a **coproduct**, a linear $\Delta : A \to A \otimes A$ (we can think of this as $\Delta a = \sum_i a^i \otimes \widehat{a}^i$ with $a^u, \widehat{a}^i \in A$). This is **coassociative**, which means
$$(\mathrm{id} \otimes \Delta) \circ \Delta = (\Delta \otimes \mathrm{id}) \circ \Delta : A \to A \otimes A \otimes A.$$
We also have a **counit** $\mathbf{1}' : A \to k$ such that
$$\Delta a = \sum_i a^i \otimes \widehat{a}^i \implies \sum_i \mathbf{1}'(a^i)\widehat{a}^i = \sum_i \mathbf{1}'(\widehat{a}^i) a^i = a.$$
(In fact, Δ being a coproduct is equivalent to $\Delta^* : A^* \otimes A^* \to A^*$ is a product.)

3. **Bialgebra**: This is $(A; m, \mathbf{1}; \Delta, \mathbf{1}')$ with $(A; m, \mathbf{1})$ an algebra, $(A, \Delta, \mathbf{1}')$ a coalgebra, and compatibility between these two structures: First, define a product $m_2 : A \otimes A \otimes A \otimes A \to A \otimes A$ by extending the following bilinear map:
$$m_2(a \otimes b, a' \otimes b') = m(a, a') \otimes m(b, b').$$
The compatibility is that $\Delta : (A, m) \to (A \otimes A, m_2)$ is a morphism with respect to the algebra structures:
$$\Delta(m(a, b)) = m_2(\Delta(a), \Delta(b)).$$
We may also write this as $\Delta(a \cdot b) = (\Delta a) \cdot_2 (\Delta b)$.

4. **Convolution product**: If $(A; m, \mathbf{1})$ is an algebra, $(C; \Delta, \mathbf{1}')$ is a coalgebra, let $\mathcal{L}_k(C, A)$ be the set of linear maps $C \to A$ (for example $\mathcal{L}(C, k) = C^*$). Then we can turn $\mathcal{L}_k(C, A)$ into an algebra by
$$(f \star g)(c) = (m \circ (f \otimes g) \circ \Delta)(c).$$
Indeed,
$$\Delta c = \sum_i c^i \otimes \widehat{c}^i \implies (f \star g)(c) = \sum_i f(c^i) \cdot_m g(\widehat{c}^i),$$
where $a \cdot_m b = m(a, b)$. Here is our unit element: $\mathbf{1}_A \circ \mathbf{1}'_C : C \to A$. This is nothing other than $(\mathbf{1}_A \circ \mathbf{1}'_C)(c) = \mathbf{1}_C(c) \mathbf{1}_A$.

Remark 28.1. If $A = k$, then $\mathcal{L}(C, k) = C^*$, and $f \star g = f \star_\Delta g = \Delta^*(f, g)$. In particular,

$$\Delta c = \sum_i c^i \otimes \widehat{c}^i \implies (f \star g)(c) = \sum_i f(c^i) g(\widehat{c}^i).$$

Example 28.1. Let $(G, \cdot, 1)$ be a group, and let k be a field. Then $A = kG = \mathrm{span}_k\{g : g \in G\}$ have multiplication $m(g_1, g_2) = g_1 \cdot g_2$, unit $\mathbf{1} = 1$, coproduct $\Delta(g) = g \otimes g$, and counit $\mathbf{1}'(\sum_i \lambda_i g_i) = \sum_i \lambda_i$. Then $(kG; m, 1; \Delta, \mathbf{1}')$ is a bialgebra.

Definition 28.2.

5. **Hopf algebra**: By a **Hopf algebra**, we mean a bialgebra $(H; m, \mathbf{1}; \Delta, \mathbf{1}')$ for which we can find a linear $S : H \to H$ such that if $\star$ is the convolution product for $\mathcal{L}_k(H, H)$, then
$$(S \star \mathrm{id}_H)(h) = (\mathrm{id}_H \star S)(h) = \mathbf{1}'(h)1,$$
where $\mathrm{id}_H, S : H \to H$. Equivalently,
$$\Delta h = \sum_i h^i \otimes \widehat{h}^i \implies \sum_i h^i \cdot_m S(\widehat{h}^i) = \sum_i S(h^i) \cdot_m \widehat{h}^i = \mathbf{1}'(h)\mathbf{1}.$$

Example 28.2. Continuing our previous example, our kG is a Hopf algebra with $S(g) = g^{-1}$ for $g \in G$.

28.2 Constructing a group of transformations from a Hopf algebra

Here is the next step:

6. Let $(H; m, \mathbf{1}; \Delta, \mathbf{1}'; S)$ be a Hopf algebra, and assume that we have a pairing of H^* and H (not necessarily the dual space pairing). Then $(H^*; \Delta*, (\mathbf{1}')^*; m^*, \mathbf{1}^*; S^*)$ is again a Hopf algebra. Given $g \in H^*$, set $\Lambda_g : H^* \to H^*$ by $\Lambda_g(f) = f \cdot_{\Delta^*} g$. We write $\Gamma_g : H \to H$ for Λ_g^*:
$$\langle \Delta^*(f, g), h \rangle = \langle f \otimes g, \Delta h \rangle = \langle \Lambda_g(f), h \rangle = \langle f, \Gamma_g(h) \rangle = f(\Gamma_g(h)).$$
Then
$$\Delta h = \sum_i h^i \otimes \widehat{h}^i \implies \sum_i f(h^i) g(\widehat{h}^i) = f\left(\sum_i g(\widehat{h}^i) h^i\right),$$
so
$$\begin{aligned}\Delta h = \sum_i h^i \otimes \widehat{h}^i \implies \Gamma_g(h) &= \sum_i g(\widehat{h}^i) h^i = \sum_i (\mathrm{id} \otimes g)(h^i \otimes \widehat{h}^i) \\ &= (\mathrm{id} \otimes g)\left(\sum_i h^i \otimes \widehat{h}^i\right) = (\mathrm{id} \otimes g)\Delta h.\end{aligned}$$

Thus, we have shown that

$$\Gamma_g = (\mathrm{id} \otimes g) \circ \Delta.$$

In summary, we have a map $\Gamma : H^* \to \mathcal{L}(H)$.

Remark 28.2. The map $g \mapsto \Gamma_g$ does not act nicely with respect to the product structure on H^*.

7. Define $G_0 = \{g \in H^* : g(h_1 \cdot_m h_2) = g(h_1)g(h_2)$. We claim that G_0 is a group and $\Gamma_{g_1 \cdot_{\Delta^*} g_2} = \Gamma_{g_1} \circ \Gamma_{g_2}$ for $g_1, g_2 \in G_0$ (so $G = \{\Gamma_g L g \in G_0\}$ is a group). In the interest of time, we will not show this now.

Definition 28.3.

8. We say that our bialgebra $\mathcal{H}$ is **graded** if $\mathcal{H} = \oplus_{n \geq 0} \mathcal{H}_n$ with $m : \mathcal{H}_n \otimes \mathcal{H}_m \to \mathcal{H}_{n+m}$ and $\Delta : \mathcal{H}_n \to \oplus_{i+j=n} \mathcal{H}_i \otimes \mathcal{H}_j$ and **connected** if $\mathcal{H}_0 = \mathrm{span}\{\mathbf{1}\}$.

Theorem 28.1. *For a graded and connected bialgebra, an antipode S exists is unique, and $S : \mathcal{H}_n \to \mathcal{H}_n$. In fact,*

$$S = \sum_{k \geq 0} (\mathbf{1} \circ \mathbf{1}' - \mathrm{id})^{\cdot_m k}.$$

29 The Final Ingredients in Our Regularity Structure

29.1 Constructing the group of transformations from a Hopf algebra

Consider a Hopf algebra $(H;\cdot,\mathbf{1};\Delta,\mathbf{1}';S)$ with dual $(H^*;\Delta^*,(\mathbf{1}')^*;\cdot^*,\mathbf{1}^*;S^*)$. Recall that we also have an algebra $(\mathcal{L}(H),\star,\mathbf{1}\circ\mathbf{1}')$, and recall that $S=(\mathrm{id}_H)^{-1}$ where the inverse is with respect to $\star$. Finally, we defined a map $\Gamma : H^*\to\mathcal{L}(H)$. (The example we should keep in mind is $H=T$ for the KPZ equation or for another PDE and G is a group of $\Gamma : H\to H$.) We defined $\Lambda : H^*\to\mathcal{L}(H^*)$ given by $\Lambda_g(f)=f\cdot_{\Delta^*}g$, which allowed us to define $\Gamma_g=\Lambda_g^*$.

Observe that $\Lambda_{g_1\cdot_{\Delta^*}g_2}=\Lambda_{g_2}\circ\Lambda_{g_1}$. From this, we can readily deduce that $\Gamma_{g_1\cdot_{\Delta^*}g_2}=\Gamma_{g_1}\circ\Gamma_{g_2}$. In other words, we have $\Gamma : (H^*,\cdot_{\Delta^*})\to(\mathcal{L}(H),\circ)$ as a homomorphism with respect to these algebra structures. For our purposes, we need a group. Namely, define the group of characters

$$G_0=\{g\in H^* : g : H\to\mathbb{R}\text{ linear}, g(h_1\cdot h_2)=g(h_1)g(h_2), g(\mathbf{1})=1\}.$$

We can see[15] that if $g_1,g_2\in G_0$, then $g_1\cdot_{\Delta^*}g_2\in G_0$. It turns out that if $g\in G_0$ and $\widehat{g}=g\circ S\in G_0$ then $\widehat{g}\cdot_{\Delta^*}g=g\cdot_{\Delta^*}\widehat{g}=(\mathbf{1}')^*$.

Recall that a **graded** bialgebra has $\mathcal{H}=\bigoplus_n\mathcal{H}_n$ with $\cdot : \mathcal{H}_m\otimes\mathcal{H}_m\to\mathcal{H}_{n+m}$ and $\Delta : \mathcal{H}_n\to\bigoplus_{i+j=n}\mathcal{H}_i\otimes\mathcal{H}_j$, and recall that a **connected** bialgebra has $\mathcal{H}_0=\{\lambda\mathbf{1} : \lambda\in k\}$.

Theorem 29.1. *Any connected, graded bialgebra has a unique antipode S.*

Proof. Here is the idea: Let $u_0=\mathbf{1}\circ\mathbf{1}'$ denote the unit for $(\mathcal{L}(H),\star)$. We want to say something like

$$\begin{aligned}(\mathrm{id})^{-1}&=(u_0-(u_0-\mathrm{id}))^{-1}\\&=\sum_{k\geq 0}(u_0-\mathrm{id})^{\star k}.\end{aligned}$$

This is algebra; we can't have an infinite sum! All we need to verify is that the if $h\in\mathcal{H}_n$, then $(\sum_{k\geq 0}(u_0-\mathrm{id})^{\star k})(h)=\sum_{k=0}^n(u_0-\mathrm{id})^{\star k}$ for some n. This is where the graded condition comes in. ☐

Example 29.1. Let $H=T(\mathbb{R}^\ell)=\bigoplus_{n\geq 0}H_n$, with $H_n=\mathrm{span}\{e_{i_1,\cdots,i_n} : i_1,\dots,i_n\in\{1,\dots,\ell\}\}$. The product on H is the **shuffle product**, $e_I\sqcup\!\sqcup e_J$. The coproduct is $\Delta(v_1\otimes\cdots\otimes v_n)=\sum_i(v_0\otimes\cdots\otimes v_i)\otimes(v_{i+1}\otimes\cdots\otimes v_n)$ with $v_0=1$. We are interested in $\mathbf{x} : [0,T]^2\to H^*$ with

$$\langle\mathbf{x},e_{i_1,\dots,i_n}\rangle=\int_0^T\cdots\int_0^T dx^{i_1}\cdots dx^{i_n}.$$

[15] Professor Rezakhanlou uses the phrase “It is not hard to see.” He has gotten comments from referees on his papers saying that he should prove more things. I like to avoid that phrase because sometimes it is hard to see.

As we discussed before, we require two properties:

1. $\mathbf{x}$ is a character:

$$\mathbf{x}(a \sqcup\!\sqcup b) = \langle \mathbf{x}, a \sqcup\!\sqcup b\rangle = \mathbf{x}(a)\mathbf{x}(b) = \langle \mathbf{x}, a\rangle\langle \mathbf{x}, b\rangle.$$

2. Chen's relation:

$$\mathbf{x}(s,u) \cdot_{\Delta^*} \mathbf{x}(u,t) = \mathbf{x}(s,t), \quad \text{so} \quad \mathbf{x}(s,t) = (\mathbf{x}(s))^{-1} \cdot_{\Delta^*} \mathbf{x}(t).$$

Now we want to use the same idea for the KPZ equation, but it doesn't work exactly the same way. We need a small variation of what we have done so far so we can deal with functions and distributions separately. Namely, we have two spaces $(T, \cdot)$ (algebra) and $(T^*, \cdot, \widetilde{\Delta}^+)$, where $\widetilde{\Delta}^+$ is a suitable coproduct. Moreover, we need $\Delta^+ : T \to T \otimes T^+$.

Recall that $\Gamma_g = (\mathrm{id} \otimes g) \circ \Delta$. Now given $g \in (T^+)^*$, we define $\Gamma_g(h) = (\mathrm{id} \otimes g) \circ \Delta^+(h)$ (so $\Gamma_g : T \to T$). Again, we may define

$$G^+ = \{g \in (T^+)^* : g(h_1 \cdot h_2) = g(h_1)g(h_2), g(\mathbf{1}) = 1\},$$

which is a group. As before, we have

$$\Gamma_{g_1 \cdot_{\Delta^+} g_2} = \Gamma_{g_1} \circ \Gamma_{g_2}.$$

Given the pair (T, T^+) with Δ^+, we are ready to build our regularity structure (not just for KPZ but for all the examples that have been worked out in this context). We use the following scheme:

First, build a linear $\mathbf{\Pi} : T \to \mathcal{D}'$, and imagine we have a map $F : \mathbb{R}^d \to G = \{\Gamma_g : G \in G^+\}$, so $F(x) = \Gamma_{f(x)}$. Then we set $\Pi_x = \mathbf{\Pi} \circ F_x^{-1}$. Then the requirement that $\Pi_x \Gamma_{x,y} = \Pi_y$ or $\mathbf{\Pi} F_x^{-1} \Gamma_{x,y} = \mathbf{\Pi} F_y^{-1}$ leads to $\Gamma_{x,y} = F_y^{-1} \circ F_x$. In fact, our T is the algebra generated by $\mathbf{1}, X_1, X_2, \partial_\ell \mathscr{I}(\tau), \mathscr{I}(\tau), \ldots$. T^+ is the algebra freely generated by $\mathbf{1}, X_1, X_2, ((\partial_\ell \mathscr{I})(\tau) : 2 - \deg \tau - |\ell| > 0)$. Set

$$\Delta^+(\mathbf{1}) = \mathbf{1} \otimes \mathbf{1}, \qquad \Delta^+(X_i) = \mathbf{1} \otimes X_i + X_i \otimes \mathbf{1}, \qquad \Delta^+(\Xi) = \Xi \otimes \mathbf{1},$$

$$\Delta^+(\underbrace{\tau \cdot \tau'}_{\text{product in } T}) = \underbrace{\Delta^+(\tau) \cdot \Delta^+(\tau')}_{\text{product in } T^+},$$

$$\Delta^+ \mathscr{I}(\tau) = (\mathscr{I} \otimes \mathrm{id})\Delta^+ \tau + \sum_{\ell : |\ell| < \deg \tau + 2} \frac{X^\ell}{\ell!} \otimes \partial_\ell \mathscr{I}(\tau).$$

29.2 Renormalization and the Wick product

We carry our all these operations to build a suitable operator $H : \mathbb{R}_x \times [0,T) \to R$ which depends on the model we have for ξ. Now replace ξ with the smoothized version $|xi * \chi^\varepsilon$ and denote the correspoding solution by H^ε. However, H^ε does not converge as $\varepsilon \to 0$. For example, replace Ξ by ξ^ε and consider $\partial \mathscr{I}(\xi^\varepsilon)$; this does not converge as $\varepsilon \to 0$. The issue is $(K_x * \xi^\varepsilon)^2 \to \infty$, where K is the heat kernel.

However, a miracle happens. If we look at $(K_x - \xi^\varepsilon)^2 - \mathbb{E}[(K_x * \xi^\varepsilon)^2]$, this has a limit as $\varepsilon \to 0$. We have

$$\int f(z_1, z_2)\xi(z_1)\xi(z_2)\, dz_1\, dz_2,$$

which causes a problem, and we replace it by

$$\int f(z_1, z_2)\xi(z_1) \diamond \xi(z_2)\, dz_1\, dz_2,$$

where $\diamond$ is the **Wick product**:

$$\xi(z_1) \diamond \xi(z_2) = \xi(z_1)\xi(z_2) - \delta_0(z_1 - z_2).$$

It turns out that all that we need to do is subtract a constant. These constants lie in a 4 or 6 dimensional group, but in the original problem, we only see 1 dimension of this.