approaching regularity structures

- what are spdes? just pdes with noise term on the right
- linear, such as SHE: $u_t = \Delta u + \xi$, where ξ is space-time white noise, or non-linear, such as PAM: $u_t = \Delta u + u\xi$ or Φ^4_3 theory $u_t = \Delta u - u^3 + \xi$
- notice that ξ is a distribution, not a function, it has regularity $-\frac{d}{2} - 1 - \varepsilon$ under parabolic scaling (explained later)
- linear theory is fairly easy: $u_t = \Delta u + \xi \implies u = G * \xi$, where G is the kernel of $(\partial_t - \Delta)^{-1}$ (Green function)
- even though G adds +2 to the regularity, $-\frac{d}{2} - 1 - \varepsilon + 2 \geq 0 \iff d < 3$, but u being a distribution won't bother us! within this lecture, we're not aiming for function-solutions, distribution-solutions are fine as well;
- in case of non-linear equations, such as PAM, we do the following:

$$u_t = \Delta u + u\xi \implies u = G * (u\xi).$$

- only now this isn't a solution, but a new integral equation. we treat this as a fixed point problem and build a ladder $u_0 = 1$, $u_1 = G * (u_0\xi)$, $u_2 = G * (u_1\xi)$, ... and expect it to converge in some sense (Picard's iterations)
- here's a problem: $u_1 = G * \xi$ is only a function if $d < 3$. if $d = 3$, its regularity is $\frac{-3}{2} - 1 - \varepsilon + 2 = -\frac{1}{2} - \varepsilon < 0$, hence the formula $u_2 = G * (u_1\xi)$ contains a product of two distributions u_1 and ξ of orders $-\frac{1}{2}$ and $-\frac{5}{2}$ correspondingly
- hence, the issue is not that u is a distribution, but that u is not even defined!
- even if we smooth the noise $\xi \rightsquigarrow \xi_\varepsilon \in C^\infty$ and the problem becomes well-defined, the solutions u_ε do not converge anywhere when we let $\varepsilon \to 0$
- **natural question:** why bother with such equations, are they even "real" is some sense, being so ill-posed? **answer:** turns out, they appear as scaling limits in a large number of microscopic non-linear random dynamics. similarly to how heat equation is a continuous proxy for studying discrete models, spdes and singular spdes are also such "universal" objects. another reason: stochastic quantization program in constructive QFT
- the class of spdes we're going to discuss are parabolic semilinear in noise:

$$u_t = \Delta u + F(u, \partial u, \zeta), \quad \text{where} \quad F(u, \partial u, \zeta) = f(u, \partial u)\,\zeta + g(u, \partial_x u).$$

- the function F is called a *nonlinearity* (even though it is allowed to be linear, such as $F(u, \partial u, \zeta) = u\,\zeta$ in PAM), the space of non-linearities will be denoted by $\mathfrak{F}$
- one natural thing to do is to mollify the noise: $\zeta_\varepsilon = \zeta * \rho_\varepsilon$, where ρ_ε is a deterministic mollifier. now the equation

$$(u_\varepsilon)_t = \Delta u_\varepsilon + F(u_\varepsilon, \partial u_\varepsilon, \zeta_\varepsilon)$$

- is well-posed and we can denote its solution by $u_\varepsilon = \mathrm{Sol}(\zeta_\varepsilon; F)$.
- **goal:**
  - with each *subcritical* (defined later) SPDE we are going to associate a **finite dimensional unbounded Lie group** called the **renormalization group**, its elements are going to be denoted by k
  - this group acts on the $\mathfrak{F}$: $(k, F) \mapsto F^{(k)} \in \mathfrak{F}$
  - **meta-theorem 1:** there exist some deterministic (typically diverging) families $(k_\varepsilon)_{\varepsilon\in(0,1]}$ such that

    $$\overline{u}^{(k)}_\varepsilon = \mathrm{Sol}(\zeta_\varepsilon; (F^{(k_\varepsilon)})^{(k)}) \to \overline{u}^{(k)} \quad \text{as} \quad \varepsilon \to 0$$

    for every k

- **def:** a **solution** to an SPDE $u_t = \Delta u + F$ is not a single function, but a family $(\overline{u}^{(k)})_k$ of functions/distributions indeded by the renormalization group
- i.e., no given k is a priori better than another from the pov of making sense of this equation

- we shall also see that $F^{(k_\varepsilon)}$ is typically given by

$$F = f\zeta + g \rightsquigarrow F^{(k_\varepsilon)} = f\zeta + g + C_\varepsilon(u, \partial u).$$

- which means that $\overline{u}_\varepsilon$ will be a solution of

$$(\partial_t - \Delta)\,\overline{u}_\varepsilon = f(\overline{u}_\varepsilon, \partial\overline{u}_\varepsilon)\,\zeta_\varepsilon + g(\overline{u}_\varepsilon, \partial_x\overline{u}_\varepsilon) + C_\varepsilon(\overline{u}_\varepsilon, \partial\overline{u}_\varepsilon).$$

- moreover, any solution theory for differential equations which aims to be **robust** should somehow be continuous with respect to the parameters. unfortunately, the continuity with respect to noise is **too much to expect** as it is false already for Ito SDEs, but the following will turn true:
- **meta-theorem 2:** there exist a *measurable* functional M^ζ of the noise ζ such that solutions are continuous with respect to M^ζ.
- i.e., we'll decompose the solution process into $\zeta \to M^\zeta \to$ Sol, where the second arrow is continuous and the first is merely measurable. M^ζ should be thought of as a very lean enhancement of ζ, necessary to make $M^\zeta \mapsto$ Sol continuous.
- "solution map is factorized"
- "wonderful decoupling of probability and analysis"
- **quick intro to rough path theory**
  - consider a deterministic ode $\dot{x}(t) = f(x(t))$ with $f \in C^\infty$. it may be *equivalently* rewritten as

$$x(t) - x(s) = f(x(s))(t-s) + O(|t-s|^2).$$

  within rough path theory and theory of regularity structures we prefer this second way of writing the equations
  - consider a similar looking equation

$$x(t) - x(s) = f(x(s))(t-s) + g(x(s))(t-s)^2 + O(|t-s|^3)$$

  is it solvable? only in the case when g is not really needed $g(x) = f(x)f'(x)$, otherwise g prescribes incompatible second order coefficient
  - now, consider another deterministic ode $dx_t = v(x_t)\,dh_t$ with $h \in C^\alpha$. what does it mean? actually, for $\alpha > \frac{1}{2}$ is makes sense as integral equation with Young integral, but let's ignore this question for now and interpret it as above:

$$x_t - x_s = v(x_s)(h_t - h_s) + O(|t-s|^{2\alpha})$$

  the question is: is it solvable?
  - to answer this, let's take the simplest case $v(x) = 1$. then the equation becomes

$$x_t - x_s = h_t - h_s + O(|t-s|^{2\alpha})$$

  every x of the form

$$x_t = x_0 + h_t + z_t, \quad |z_t - z_s| \lesssim |t-s|^{2\alpha}$$

  works. if $\alpha > \frac{1}{2}$, there is no such z (no Hölder continuous functions of order bigger than 1), hence the solution is unique. if $\alpha < \frac{1}{2}$, the equation is **underdefined**! we still need an equation to fix z_t

- we can try to add more terms to the equation, along the lines of

$$x_t - x_s = v(x_s)(h_t - h_s) + q(v_s)g_{s,t} + O(|t-s|^{3\alpha})$$

and try to see if this fixes the issue. if $\alpha \in \left(\frac{1}{3}, \frac{1}{2}\right]$. the answer is **yes, but g cannot be any function**. there's definitely *some freedom* in prescribing g, but g cannot be completely unrelated to h, it has to satisfy certain algebraic properties
- one of Lyons' deep insights was to realize $g_{s,t}$ must behave *algebraically* **as if** it was given by an integral of the form $\int_s^t (h_r - h_s)\, dh_r$, the precise notion is **Chen relations**

$$g_{s,t} = g_{s,u} + g_{u,t} + (h_u - h_s)(h_t - h_u)$$

(try manipulating $\int_s^t (h_r - h_s)\, dh_s$ into this form *formally*)
- here's a heuristic argument to see that $g_{s,t}$ is of this form:
  * assume that $x_t - x_s = \int_s^t v(x_r)\, dh_r$ with some notion of $\int (\dots)\, dh_r$
  * expand $v(x_r)$ around x_s: $v(x_r) = v(x_s) + v'(x_s)(x_r - x_s) + \dots$
  * plug this expansion in: $x_t - x_s \approx v(x_s)(h_t - h_s) + v'(x_s) \int_s^t (x_r - x_s)\, dh_r$
  * plug this expansion into itself: $x_t - x_s \approx v(x_s)(h_r - h_s) + v'(x_s)v(x_s) \int_s^t (h_r - h_s)\, dh_r$
  * hence, knowing x_s, to find x_t up to first order we need to know $h_r - h_s$, to second order we also need $\int_s^t (h_r - h_s)\, dh_r$
  * if we proceed further, we'll find that we need $\int_{s<u_1<u_2<u_3<t} dh_{u_1}\, dh_{u_2}\, dh_{u_3}$ and so on
  * how many integrals we need to keep track of depends on how rough h is, typically it's just one iterated integral $\int_{s<u_1<u_2<t} dh_{u_1}\, dh_{u_2}$
- the freedom in choosing g lies in the fact that Chen relations don't define the integral uniquely, there's still a lot of options. for instance, if h is the Brownian motion with regularity slightly less than $\frac{1}{2}$, the available notions include Ito integral, Stratonovich integral, backward integral among others
- there's always at least one choice $\int_s^t (h_r - h_s)\, dh_r := \frac{1}{2}(h_t - h_s)^2$ — the *geometric* lift. this is exactly the deterministic version of Stratonovich integral. in dimensions higher than one geometric lift also isn't unique. Föllmer's theory of pathwise quadratic variation allows to define deterministic Itô lift.
- assume we want to solve an SDE $dY_t = a\, dt + b\, dX_t$ with $X \in C^\alpha$, $\alpha \in \left(\frac{1}{3}, \frac{1}{2}\right)$, we can build an $\mathbb{R}^{d+d^2}$ valued process $(X, \mathbb{X}) = \left(X, \int X \otimes dX\right)$ called the rough path over X
- then $X \mapsto Y$ factorizes through $X \mapsto (X, \mathbb{X})$, which is measurable, and $(X, \mathbb{X}) \mapsto Y$ is continuous
- in other words, repeated integrals $\int X \otimes dX$ are the only extra information needed to make $(X, \text{other info about } X) \mapsto Y$ continuous and also the only extra information needed to define a rough integral $\int F(X)\, d(X, \mathbb{X})$
- note that if we enhance X with too much information, such as $X \mapsto (X, Y) \mapsto Y$, the second arrow is trivially continuous. the point is to find a lean extension, i.e., to add as little to X as possible
- moreover, $(X, \mathbb{X}) \mapsto Y$ is a *deterministic* map, built purely analytically, with no probability, Ito integrals and all
- $\int X \otimes dX$ is the notion of stochastic integration, and it may be defined in a bunch of different non-equivalent ways, such as Ito or Stratonovich, as well as others. the corresponding lifted object $\mathbb{X}$ *differ* as well, so in rough path theory we speak of Ito lift, Stratonovich lift, backward integral lift, etc.

- in other words, what we're saying is that to build an SDE solution theory we don't need to define all possible stochastic integrals (although they're very useful for other purposes), we just need to define integrals of X with respect to itself.

- so, Π shall serve a similar purpose as $X \mapsto \mathbb{X}$ in rough path theory. actually, theory of regularity structures is *a direct generalization of rough path theory*.
- before we proceed, let's take the notion of **Taylor polynomial** and look at it from a very specific angle:

$$u(\cdot) \approx \sum_{k=1}^{N} u_k(x)\,(\cdot - x)^k \quad \text{near } x$$

let's think about this formula as follows: we have a family of **abstract labels** k which serve as names of monomial objects $(-)^k$, the function u may locally be described by a jet $(u_k(x))_{k\in\mathbb{N},x\in\mathbb{R}^d}$ with respect to these monomial objects. notice that the **labels themselves** interact with different operations such as derivative, integral, applying smooth nonlinear function.
- polynomials, however, are not enough to describe how SPDE solutions look locally. the solution seems clear: replace monomials by some "basis" family $\{e_k\}_{k=1}^N \in S'$ and assume

$$u(\cdot) \approx \sum_{k=1}^{N} u_k(x)\, e_k(\cdot) \quad \text{near } x$$

- this is much better, but still not good enough, since we cannot really apply nonlinear maps to distributions. the theory of regularity structures takes this one step further: instead of concrete basis in S', let's introduce **formal symbols** $\tau \in T$ and endow them with some **natural algebraic structure**, and then define a map Π_x which takes a symbol τ and interprets it as function/distribution. we're going to do so in such a way that all objects relevant for our study can locally be described by finite expansions with respect to $\Pi_x\tau$:

$$u(\cdot) \approx \sum_{\tau} u_\tau(x)(\Pi_x\tau)(\cdot) \quad \text{with finite sum}$$

- i.e., $\Pi_x\tau(\cdot)$ plays similar role to $(\cdot - x)^k = \Pi_x k(\cdot)$, but specific spaces of τ we'll introduce later will be richer than $\mathbb{N}_0$, namely, *trees*
- what do we gain by this? we **separate the form** of the equation from how we make sense of its symbols. the *form* of the equation shall be represented faithfully in the deterministic symbolic τ-realm, but the *flesh*, including probabilistic construction of integrals with respect to noise and so on, shall be subsumed into the construction of Π
- since we now have a local description of u near x, we can use it to write $\mathcal{L}u = F(u,\zeta)$ in the first order as

$$F\left(\sum_{\tau} u_\tau(x)\,\Pi_x\tau(\cdot), \zeta\right) = \sum_{k} \frac{\partial_u^k F(u_{\mathbf{1}}(x),\zeta)}{k!} \left(\sum_{\tau\neq\mathbf{1}} u_\tau(x)\,\Pi_x\tau(\cdot)\right)^k$$

$$\mathcal{L}u = F(u,\zeta) \implies \sum_{\tau} u_\tau(x)\,(\mathcal{L}\Pi_x\tau)(\cdot) \approx \sum \frac{\partial_u^k F(u_{\mathbf{1}}(x),\zeta)}{k!} \sum_{\tau_1,\ldots,\tau_k\neq\mathbf{1}} \prod_{i=1}^{k} u_{\tau_i}(x)(\Pi_x\tau_i)(\cdot)$$

- this is a triangular system on $\boldsymbol{u} = (u_\tau(x))_{\tau,x}$
- nothing is resolved yet, of course, because products $\zeta\prod_i \Pi_x\tau_i$ are undefined, but we have achieved something: the problem of making sense of the equation is now isolated in making sense of these elementary products

- slogans:
  - keep local Taylor expansion device philosophy replacing monomials by abstract symbols
  - separate the combinatorics of non-linear expansions/iterations from analytic realizations
- having built T and Π, we shall endow the space of T-valued functions $\boldsymbol{u} : \mathbb{R} \times \mathbb{R}^d \to T$ with some topology and learn how to transition between $u \leftrightarrow \boldsymbol{u}$. the ultimate goal is to lift an SPDE into the $\boldsymbol{u}$ realm, use the fixed point theorem there, and lower it back to u

algebra of local expansion devices and regularity structures

- consider Taylor expansion device. let's answer the following question: given a family $(f_n(x))_{n\in\mathbb{N}_0, x\in\mathbb{R}^d}$, can we find a single function f such that

$$f(\cdot) \approx \sum_n f_n(x)(\cdot - x)^n \quad \text{near } x \text{ ?}$$

- well, clearly we need

$$f(\cdot) \approx \sum_n f_n(x)(\cdot - x)^n = \sum_\ell \left(\sum_{n\geq\ell} f_n(x) \binom{n}{\ell} (y-x)^{n-\ell} \right) (\cdot - y)^\ell.$$

- in other words, any family $(f_n(x))_{n,x}$ doesn't glue, we need to assume

$$f_\ell(y) = \sum_{n\geq\ell} f_n(x) \binom{n}{\ell} (y-x)^{n-\ell}.$$

- this is a *consistency condition*, and we should think of $(f_n(x))_{n,x}$ as *jet*
- we're going to generalize this "expansion device" idea now
- let $\mathcal{B}$ be any finite set, $\tau \in \mathcal{B}$ its element, $T = \operatorname{span} \mathcal{B}$
- we aim to introduce a family of maps Π_x which take $\tau \in T$ and produce a distribution $\Pi_x \tau(\cdot) \in S'(\mathbb{R}^d)$, which should be thought of as "centered at x", and then use those to describe functions/distributions as

$$f(\cdot) \approx \sum_\tau f_\tau(x)(\Pi_x \tau)(\cdot) \quad \text{near } x.$$

- to mimick the reexpansion operation above, we need to relate $f_\tau(x)$ to $f_\sigma(y)$ for $y \approx x$, so we want some linear consistency of the form

$$f_\sigma(y) \approx \sum_\tau f_\tau(x) \cdot (\dots)$$

- however, there's a certain assymetry with the Taylor expansion of functions here: there, we described *functions in terms of functions*, but here we describe **distributions** f in terms of function-coefficients $f_\tau(x)$ and distribution-basis $\Pi_x \tau(\cdot)$. hence, since relating $f_\sigma(y)$ to $f_\tau(x)$ should only involve **functions**, we cannot just use the same $\Pi_x \tau(\cdot)$ to reexpand $f_\sigma(y)$ with respect to $f_\tau(x)$.
- one way out of this situation is to split the set of τ's into those with $\Pi_x \tau(\cdot)$ being a distribution and those with $\Pi_x \tau(\cdot)$ being a function, but this is impercise and a tad ugly
- better approach is to just take another finite set $\mathcal{B}^+$ and to each $\mu \in \mathcal{B}^+$ associate a true function (not distribution) $g_{yx}(\mu)$
- you may, of course, think that $\mathcal{B}^+$ are just elements of $\mathcal{B}$ with "positive degree", but it's also fine and from certain perspective better to think that those are their own independent objects

- **rem:** elements of $\mathcal{B}^+$ are associated with functions, elements of $\mathcal{B}$ are associated with distributions. **define** $T^+ = \operatorname{span}\mathcal{B}^+$.
- then, we assume that
$$f_\sigma(y) \approx \sum_{\tau\in\mathcal{B},\mu\in\mathcal{B}^+} c_\mu^{\tau\sigma} f_\tau(x) g_{yx}(\mu)$$
with some constants $c_\mu^{\tau\sigma}$ independent of f
- we shall furthermore ask that the family $g_{yx}(\mu)$, $\mu \in \mathcal{B}^+$ is sufficiently rich to be an *algebra* of functions. the cheapest way to ensure this is to assume that the symbols T^+ themselves form an algebra and $g_{yx} : T^+ \to \mathbb{R}$ are characters of this algebra
- **note that** we don't require T to be an algebra, because T encodes distributions
- **define** formal ratio
$$\tau/\sigma := \sum_{\mu\in\mathcal{B}^+} c_\mu^{\tau\sigma} \mu \in T^+.$$
- then,
$$f_\sigma(y) \approx \sum_{\tau\in\mathcal{B}} f_\tau(x)\, g_{yx}(\tau/\sigma)$$
- now, the consistency assumption we want to impose becomes
$$\sum_{\sigma\in\mathcal{B}} g_{zy}(\sigma/\eta)\, g_{yx}(\tau/\sigma) = g_{zx}(\tau/\eta).$$
- this is the analogue of the binomial expansion consistency relation
- now, let's add more algebra to it. why? well. . .
  - the presence of so much algebra as we're about to see is due to the choice of working with jets
  - but also algebraic structures will turn useful for renormalization purposes
  - finally, it allows to strip the meat completely and obtain a bare skeleton of reexpansion relations without referencing functions or distributions
- to encode the reexpansion algebraically, we introduce the operator whith splits a symbol τ into data for reexpansion: $\Delta : T \to T \otimes T^+ : \Delta\tau = \sum_\sigma \sigma \otimes (\tau/\sigma)$,
  - for example, $\Delta F = F \otimes \mathbf{1} + X \otimes F' + X^2 \otimes \frac{F''}{2}$, then symbols on the left are interpreted by Π_x and objects on the right by g_x, so that, for example, $\Pi_x F = f(\cdot - x)$, $\Pi_x X^k = (\cdot - x)^k$ and $g_x(F^{(n)}) = f^{(n)}(x)$
  - by swapping the model, we can also handle $Y_t - Y_s \approx Y_s'(X_t - X_s)$ (Gubinelli derivative) in the same way
  - we shall call this a comodule map (proper definition below)
- we also need to reexpand the reexpansion coefficients, so we introduce a second operator which prepares $\tau/\eta \in T^+$ for further reexpansion $\Delta^+ : T^+ \to T^+ \otimes T^+ : \Delta^+(\tau/\eta) = \sum_\sigma (\sigma/\eta) \otimes (\tau/\sigma)$
  - note that elements of T^+ are not uniquely described as τ/σ, so the definition is questionable, but we'll see how it works and what it means in *concrete* regularity structures
  - we shall call this a coproduct (proper definition below)
- a few remarks on coproducts
  - coproduct is coassociative, but it doesn't need to be and is not in general cocommutative in the sense that $\Delta^+ = S \circ \Delta^+$ where $S(a \otimes b) = b \otimes a$
  - there's always at least some coproduct on an algebra. for example, we can define how it splits generators, the simplest example being $\Delta^+ x = x \otimes 1 + 1 \otimes x$ and extend Δ^+ multiplicatively, then $\Delta^+(xy) = xy \otimes 1 + x \otimes y + y \otimes x + 1 \otimes xy$ and so on. this gives polynomial coproduct
  - there's also a group-like coproduct $\Delta^+ x = x \otimes x$ for all x

- there's also a deconcatenation coproduct on words cutting words in two $\Delta(a_1 a_2 \dots a_n) = \sum_i (a_1 \dots a_i) \otimes (a_{i+1} \dots a_n)$, i.e., collect all possible splits of a word in two
- the motivation is over now, we're ready to give a proper abstract definition:
- **def.** a concrete regularity structure $\mathcal{T} = (T^+, T)$ is a
  - pair of *graded* vector spaces $T^+ = \bigoplus_{\alpha \in A^+} T^+_\alpha$ and $T = \bigoplus_{\beta \in A} T_\beta$
  - T^+_α and T_β are finite dimensional
  - T^+ is a connected graded **bialgebra** with unit $\mathbf{1}_+$, counit $\mathbf{1}'_+$, coproduct $\Delta^+ : T^+ \to T^+ \otimes T^+$ and grading $A^+ \subset [0, \infty)$
  - A locally finite subset of $\mathbb{R}$
  - T is a right comodule over T^+, i.e., $\Delta : T \to T \otimes T^+$, which satisfies

    $$(\Delta \otimes \mathrm{Id})\,\Delta = (\mathrm{Id} \otimes \Delta^+)\,\Delta \quad \text{and} \quad (\mathrm{Id} \otimes \mathbf{1}'_+)\,\Delta = \mathrm{Id}$$

    (since Δ encodes preparing for reexpansion, these conditions mean that reexpanding twice behaves as one expects and how to get the object back if we changed our mind about reexpanding it)
  - moreover, $\Delta T_\beta \subset \bigoplus_{\alpha \geq 0} T_{\beta - \alpha} \otimes T^+_\alpha$
- **rem.** recall that **bialgebra** is a an algebra B with
  - product $M : B \otimes B \to B$ and unit $\mathbf{1}$
  - coproduct $\Delta : B \to B \otimes B$ which is coassociative $(\Delta \otimes \mathrm{Id})\,\Delta = (\mathrm{Id} \otimes \Delta)\,\Delta$
  - counit $\theta : B \to \mathbb{R}$ such that $(\theta \otimes \mathrm{Id})\,\Delta = (\mathrm{Id} \otimes \theta)\,\Delta = \mathrm{Id}$ (identify $\mathbb{R} \otimes B \cong B \otimes \mathbb{R} \cong B$)
    * actually, if $\tau = \sum_{\sigma \in \mathcal{B}^+} c_\sigma\, \sigma$, then $\theta(\tau) = c_{\mathbf{1}}$, i.e., θ is projection
  - if, furthermore, there's an **antipode** $S : B \to B$ such that $M(\mathrm{Id} \otimes S)\,\Delta = M(S \otimes \mathrm{Id})\,\Delta = \theta(\cdot)\,\mathbf{1}$, then this is a **Hopf algebra**
    * graded bialgebra with $B_0 = \langle \mathbf{1} \rangle$ is said to be connected
    * connected graded bialgebra always has an antipode
- we should always keep in mind
  - the meaning of T and T^+: they **index expansion devices** for a given SPDE
  - we do not assume any relation between T_α and T^+_α at this point, they are just formal symbols
  - that each equation has **its own** regularity structure
  - that Δ and Δ^+ encode coherence relations, reexpansion; they should be interpreted as *chopping elements into pieces*
- **define**
  - τ_β by $\tau = \sum_{\beta \in A} \tau_\beta$ with $\tau_\beta \in T_\beta$
  - **homogeniety** $|\tau| = \alpha$ for $\tau \in T^+_\alpha$
  - $\|\cdot\|_\beta$ norm on T_β or T^+_β (since both are findim, all norms are equivalent; we need them to have at least some notion of size, for instance, to say things like "a function is bounded")
  - for $\tau \in T$, $\|\tau\|_\beta := \|\tau_\beta\|_\beta$
  - $\mathcal{B}^+_\alpha$ and $\mathcal{B}_\beta$ bases of T^+_α and T_β
  - $\mathcal{B}^+ = \bigcup_{\alpha \in A^+} \mathcal{B}^+_\alpha$
  - $\mathcal{B} = \bigcup_{\beta \in A} \mathcal{B}_\beta$
- **example of a regularity structure.**
  - take *symbols* $X_1, \dots, X_d$
  - for $n \in \mathbb{N}^d$, define $X^d = \prod X_i^{n_i}$
  - define $T_X = \mathrm{span}\{X^n : n \in \mathbb{N}^d\}$
  - this is a bialgebra with free product and coproduct $\Delta X^n = \sum_{\ell \leq n} \binom{n}{\ell} X^\ell \otimes X^{n-\ell}$
  - let $\mathfrak{s} = (\mathfrak{s}_i) \in (\mathbb{N} \setminus \{0\})^d$ be called *scaling*
  - define the scaled degree $|n|_{\mathfrak{s}} = \sum \mathfrak{s}_i n_i$

- define grading $T_\alpha = \operatorname{span}\{X^n : |n|_\mathfrak{s} = \alpha\}$, then T_X is a graded bialgebra
- note that $T_0 = \operatorname{span}\{\mathbf{1}\}$, hence T_X is a *connected* graded bialgebra, hence a Hopf algebra
- antipode is given by $S_+ X^n = (-X)^n$
- this is called **polynomial regularity structure**

- **another example.**
  - let T and T^+ be sets of rooted trees with some "homogeniety" assined to each tree and $T^+ \subset T$ be trees with positive homogenieties
  - $\Delta\tau = \sum_{\sigma \text{ substree of } \tau} \sigma \otimes \tau/\sigma$ where τ/σ is τ with σ contracted to its root (quotient tree)
- **def.** a chatacter g on a Hopf algebra T^+ is a linear map $g : T^+ \to \mathbb{R}$ such that
  - $g(\mathbf{1}_+) = 1$
  - $g(\tau_1\tau_2) = g(\tau_1)g(\tau_2)$
  - G^+ = the set of characters is a group with respect to the **convolution product** defined by $(g_1 * g_2)(\tau) = (g_1 \otimes g_2)\,\Delta^+\tau$
    - strictly speaking, the result is number $\otimes$ number $\in \mathbb{R} \otimes \mathbb{R}$, but we identify it with $\mathbb{R}$ as usual ($\mathbb{R}$ being the unit of $\otimes$)
  - the unit of G^+ is the counit $\mathbf{1}'_+$ of T^+
  - the inverse in G^+ is given by $g^{-1} = g \circ S_+$, where S_+ is the antipode of T^+
  - G^+ is called **the structure group**
- think of the usual convolution $(g_1 * g_2)(x) = \int g_1(y)g_2(x-y)\,dy$. notice that we need a coproduct to split x into $y \otimes (x-y)$ in this def
- define "Fourier transform" $\widehat{g} = (\operatorname{Id} \otimes g)\,\Delta : T \to T$. "$G^+$ acts on T from the left"
- it satisfies $\widehat{g_1 * g_2} = \widehat{g_1} \circ \widehat{g_2}$:

$$\begin{aligned}
\widehat{g_1 * g_2} &= (\operatorname{Id} \otimes (g_1 * g_2))\,\Delta && \text{def of Fourier transform}\\
&= (\operatorname{Id} \otimes g_1 \otimes g_2)(\operatorname{Id} \otimes \Delta^+)\,\Delta && \text{def of convolution}\\
&= (\operatorname{Id} \otimes g_1 \otimes g_2)(\Delta \otimes \operatorname{Id})\Delta && \text{comodule property}\\
&= (\operatorname{Id} \otimes g_1)\Delta \circ (\operatorname{Id} \otimes g_2)\Delta\\
&= \widehat{g_1} \circ \widehat{g_2}
\end{aligned}$$

- similarly, $\widehat{g}^+ = (\operatorname{Id} \otimes g)\Delta^+ : T^+ \to T^+$ acts on T^+ from the left
- we can show that $\widehat{g}(\tau) = \tau + \mu$ where $\tau \in T_\beta \implies \mu \in \bigoplus_{\beta' < \beta} T_{\beta'}$

Hölder spaces and models

- so far we've only described algebraic setting, let's turn to analytical. as we've said before, regularity structures are tailored to a specific SPDE, and, for instance, SPDEs involving $\partial_t - \Delta_x$ and $\partial_t + (-\Delta_x)^\alpha$ require different spaces. we'll only work with $\partial_t - \Delta_x$ from now on
- **def.** parabolic distance $d(x,y) = \sqrt{|x_0 - y_0|} + |x' - y'|$
- recall that we're building a device for describing a global distribution locally using building blocks $\Pi_x\tau$ and $g_{yx}(\tau)$, and recall that without consistency conditions just any family of coefficients $\boldsymbol{u} = (u_\tau(x))_{\tau \in T, x \in \mathbb{R}^{d+1}}$ doesn't glue to a global object
- the notion of modelled distribution provides such glueing conditions
- reconstruction theorem makes this precise
- first off, take parabolic scaling $\mathfrak{s} = (2, 1, \ldots, 1) \in \mathbb{N} \times \mathbb{N}^d$ and define scaled degree by $|n|_\mathfrak{s} = 2n_0 + \sum_{i \geq 1} n_i$
- now we're going to give a rather unusual definition of Hölder spaces
- **def.** Hölder spaces

- take a test function $\varphi \in C_c^\infty(\mathbb{R}\times\mathbb{R}^d)$ and rescale it parabolically

$$\varphi_x^\lambda(\cdot) = \lambda^{-(d+2)}\varphi\left(\frac{\cdot_0 - x_0}{\lambda^2}, \frac{\cdot' - x'}{\lambda}\right)$$

- a distribution $\Lambda \in S'(\mathbb{R}\times\mathbb{R}^d)$ is in $C^\alpha(\mathbb{R}\times\mathbb{R}^d)$ (modulo polynomials) if and only if the local averages of Λ scale as λ^α, i.e.,

  $$|\langle \Lambda, \varphi_x^\lambda\rangle| \lesssim \lambda^\alpha$$

  with constant depending on φ uniformly on $\varphi : \operatorname{supp}\varphi \subset B(0,1)$ and $\|\varphi\|_{C^m} \leq 1$ and uniform in x and $\lambda \in (0,1]$ for all φ such that φ kills polynomials up to order α

  $$\int x^k \varphi(x)\,dx = 0 \quad \text{for} \quad k : |k|_{\mathfrak{s}} < \alpha$$

- alternatively, if φ is any and the condition is replaced by

  $$\exists\, P_x \text{ polynomial of degree } < \alpha : |\langle \Lambda - P_x, \varphi_x^\lambda\rangle| \lesssim \lambda^\alpha$$

- technically, we defined so-called Hölder-Besov spaces $B^\alpha_{\infty,\infty}$. for positive non-integer α, these are *exactly* the usual Hölder spaces
- for positive integer k, these are so-called Zygmund spaces, which are slightly bigger than standard spaces of k times continuously differentiable functions C^k
- for instance, $f \in C^1$ means $|f(x+h) - 2f(x) + f(x-h)| \lesssim |h|$, which almost implies that f' is continuous, but not quite (note how this ties with the sewing lemma!)
- for instance, $\ln x \in C^0$, $x \ln x \in C^1$ and so on
- note also that $C^{-\alpha} = B^{-\alpha}_{\infty,\infty}$ is not the dual of C^α, the true dual is something bigger than $B^{-\alpha}_{1,1}$ (similarly to how the dual of L^∞ is bigger than L^1)

• the paper follows a more direct definition, but we won't need it
• there are some important distinctions/details in the definition which I ignored
• from now on x denotes the space-time point $x = (x_0, x') \in \mathbb{R}\times\mathbb{R}^d$. **we shall reserve t letter for a different purpose!**
• we shall replace all test functions φ by a simple parametrized family $p_t(x,y) = e^{t(\partial^2_{x_0} - \Delta^2_{x'})}(x,y)$
  - why replace? well, we want to have one regularity parameter s instead of many test functions
  - why this exact kernel? first off, the symbol of $\partial^2_{x_0} - \Delta^2_{x'}$ is $-|\xi_0|^2 - |\xi'|^4$, hence it's an elliptic operator on the full spacetime; second, it scales parabolically; a natural second order candidate $\partial^2_{x_0} + \Delta$ doesn't scale parabolically
  - I may ocasionally slip and refer to it as the *heat kernel*, because it is a heat kernel of $\partial^2_{x_0} - \Delta^2_{x'}$ in $\mathbb{R}\times\mathbb{R}^{d+1}$
  - also, this operator is quite nice analytically, has nice bounds

    $$\int |\partial_x^n p_t(x,y)|\, d^a(x,y)\,dy \lesssim t^{\frac{a-|n|_{\mathfrak{s}}}{4}}$$

    and enables the use of Littlewood-Paley theory

• **def: a model** over a regularity structure $\mathcal{T}$ is a pair of maps $M = (g, \Pi)$

  $$g : \mathbb{R}\times\mathbb{R}^d \to G^+ \quad \text{and} \quad \Pi : T \to S'(\mathbb{R}\times\mathbb{R}^d)$$

  such that, with $g_{yx} = g_y * g_x^{-1}$ and $\Pi_x^g = (\Pi \otimes g_x^{-1})\,\Delta$ we have

- $\|g\|_\gamma = \sup\left\{\frac{|g_{yx}(\tau)|}{d(y,x)^{|\tau|}} : \tau \in \mathcal{B}^+,\ |\tau| < \gamma,\ x, y \in \mathbb{R}\times\mathbb{R}^d\right\} < \infty$
  - i.e., reexpansion coefficient $g_{yx}(\tau)$ from x to y may be unbounded (as in the case of polynomials) but is controlled by the parabolic distance $d(y,x)^{|\tau|}$
  - for example, think of an element which splits as $\Delta^+\tau = \tau \otimes \mathbf{1} + \mathbf{1} \otimes \tau$. then

$$(g_y \otimes g_x^{-1})\Delta^+\tau = g_y(\tau) + g_x^{-1}(\tau) = g_y(\tau) + g_x \circ S_+(\tau) = g_y(\tau) - g_x(\tau)$$

    then this is the Hölder condition on $x \mapsto g_x(\tau)$
  - we used $S_+(\tau) = -\tau$ under this simple coproduct and $g_x(-\tau) = -g_x(\tau)$ because g is a character
    - proof: S is defined by $\mathcal{M}(S \otimes \mathrm{Id})\Delta\tau = \theta(\tau)$ and $\theta(\tau) = \mathbf{1}_{\tau=\mathbf{1}}$, therefore

$$\mathcal{M}(S \otimes \mathrm{Id})(\tau \otimes \mathbf{1} + \mathbf{1} \otimes \tau) = S(\tau) + \tau = 0$$

- $\|\Pi^\gamma\|_\gamma = \sup\{t^{-|\sigma|/4}|\langle \Pi_x^g \sigma, p_t(x,\cdot)\rangle| : \sigma \in \mathcal{B}, |\sigma| < \gamma, x \in \mathbb{R}\times\mathbb{R}^d, 0 < t \le 1\} < \infty$

- **def: distance between models**

$$d_\gamma(M, M') = \sup \frac{|g_{yx}(\tau) - g'_{yx}(\tau)|}{d(y,x)^{|\tau|}} + \sup t^{-|\sigma|/4}\left|\langle \Pi_x^g \sigma - \Pi_x'^{g'}\sigma, p_t(x,\cdot)\rangle\right|$$

- In Hairer's original work, a model is a pair of $\Pi = (\Pi_x : T \to S'(\mathbb{R}\times\mathbb{R}^d))_{x\in\mathbb{R}\times\mathbb{R}^d}$ and $\Gamma = (\Gamma_{yx} : T \to T)_{y,x\in\mathbb{R}\times\mathbb{R}^d}$, but framing with (g,Π) is considered more standard now. the relation between two notions is given by $\Gamma_{yx} = \widehat{g_{yx}}$
- **reminder:** g acts on T^+, Π acts on T, and g plays the same role as Π: g makes functions out of function-symbols similarly to how Π makes distributions out of distribution-symbols
  - $g_{yx}(\tau) = (g_y \otimes g_x^{-1})\Delta^+\tau$
  - $\Pi_x^g \sigma = (\Pi \otimes g_x^{-1})\Delta\sigma$
  - split, recenter by g_x^{-1}, interpret
  - Δ identifies within σ pieces which can be meaningfully recentered and recenters them to x
  - the condition $\|\Pi^\gamma\|_\gamma < \infty$ should be thought of as "$\Pi_x^g\tau$ behaves at x as an element of $C^{|\tau|}(\mathbb{R}\times\mathbb{R}^d)$"
  - for instance, we can build a model on (T^+, T^+) (everything is a function, no distributions) from g only by $(\Pi^{(g)}\tau)(x) = g_x(\tau)$
- in the class of problems we consider, it is sufficient to fix one large γ and work under $|\tau| < \gamma$ always
- in the notion of regularity structure, gradings A and A^+ are abstract, purely algebraic. choosing a model on regularity structure *forces* a grading, because it uses these gradings in analytical bounds
- restricting the set of test functions to just $p_t(x,\cdot)$ is enough, because it allows to bound $|\langle \Pi_x^g\tau, \partial_x^n p_t(x,\cdot)\rangle|$ as well
- **example: canonical model over a polynomial regularity structure**: $g_x(X_+^n) = x^n$ and $\Pi X^n(y) = y^n$
- **example:** let's say we want to represent the delta-function δ_0 by a symbol τ in our regularity structure. first, the analytic bound above *forces* $|\tau| < -(d+2)$. then, we have to explain how it reexpands from point to point. the reexpansion mechanism is *data* we put into the model, not something contained in δ_0 itself. this data is provided by defining $\Pi_x\tau \in S'$ and $\Gamma_{yx}\tau \in T$ in the Hairer's original notion of regularity structures or by defining how τ splits for reexpansion by Δ in the concrete reexpansion framework. we can *postulate* $\Delta\tau = \tau \otimes \mathbf{1}$, which means that τ is invariant under recentering, or something else. one bad option is to demand that δ_0 becomes

δ_x when interpreted at x: indeed, under such reexpansion δ_0 picks up terms of infinitely bad regularities

$$\delta_x = \sum \frac{x^k}{k!}\partial^k \delta_0 \quad \text{in the sense that} \quad \varphi(x) = \sum \frac{x^k}{k!} \underbrace{\partial^k \varphi(0)}_{=\langle \partial^k \delta_0, \varphi\rangle}$$

this means that if we want δ_x in our model, we'd better include it as a separate symbol, than as reexpansion of δ_0

modelled distributions

- so far we built a family of abstract symbols T and manually specified which S' distribution we want to assign to each $\tau \in T$. we have also manually defined what we mean by recentering a distribution to x. in the language of Taylor polynomials, τ are "monomials as such (i.e., X^k)" as opposed to "monomials based at x (i.e., $(\cdot - x)^k$)". in general, I'd call them "homogeniety stencils". both assignments are done by hand, we chose these two notions by ourselves. now, we want to paste τ's together and build distributions, we want to say things like "near x, my f looks as τ up to order γ". to this end, we take a T-valued function $\boldsymbol{f} : \mathbb{R}^{d+1} \to T : \boldsymbol{f}(x) = \sum f_\tau(x)\,\tau$ and treat it as a jet of coefficients. we can then interpret this jet by $(\Pi_x \boldsymbol{f}(x))(\cdot)$. the question is: do these pieces glue together into a globally defined object?
- if what we want is to describe the *values* of a function f in terms of its Taylor jet, zeroth order jet trivially suffices ($f(x)$ = coefficient of 1 in the expansion of $f(\cdot)$ near x), but we need richer jets
- why richer jets? think of jets as *values enriched:* we enhance a number $f(x)$ to $(f(x), \text{some other information about } f \text{ at } x)$. in programming terms, think of precomputing this "other info at x" and making it available for use of different operators such as differential operators: if the jet $\boldsymbol{f}(x)$ contains only the values $f(x)$ at every point, describing a first order differential operator requires actually computing derivatives, hence knowing f at points other than x. but if $\boldsymbol{f}(x) = (f(x), f'(x), f''(x), \dots)$ is available, computing $Df(x) = f'(x)$ requires only reading the second component $Df(x) = (\boldsymbol{f}(x))_2$, which is done algebraically and without leaving x
- we can also lift D on the level of jets by demanding that it sends k-jets into $(k-1)$-jets by $D(f(x), f'(x), \dots, f^{(k)}(x)) = (f'(x), \dots, f^{(k)}(x))$. note that it doesn't preserve the order of a jet because computing $f^{(k+1)}(x)$ from a k-jet requires actual computation necessarily involving points other than x
- for our purposes finite jets would suffice, so we define truncated T

$$T_{<\gamma} = \bigoplus_{\beta<\gamma} T_\beta, \quad T^+_{<\gamma} = \bigoplus_{\alpha<\gamma} T^+_\alpha$$

- **def: the space of modelled distributions on a regularity structure $\mathcal{T}$ with transition** g is the space of functions $\boldsymbol{f} : \mathbb{R} \times \mathbb{R}^d \to T_{<\gamma}$ such that

$$[\![\boldsymbol{f}]\!]_{\mathcal{D}^\gamma} = \max_{\beta<\gamma} \sup_x \|\boldsymbol{f}(x)\|_\beta < \infty \quad \text{and} \quad \|\boldsymbol{f}\|_{\mathcal{D}^\gamma} = \max_{\beta<\gamma} \sup_{x,y} \frac{\|\boldsymbol{f}(y) - \widehat{g_{yx}}\boldsymbol{f}(x)\|_\beta}{d(y,x)^{\gamma-\beta}} < \infty$$

- note that $[\![\boldsymbol{f}]\!]_{\mathcal{D}^\gamma} < \infty$ implies that all expansion coefficients are globally bounded. this seems too restrictive, but
  - first, we're usually working on the torus $\mathbb{T}$, hence everything is bounded by compactness

– second, in Hairer's original definition the choice is more local, we can replace this condition with local boundedness
– in the end, it's just a simplification

- $|||f|||_{\mathcal{D}^\gamma} = [\![f]\!]_{\mathcal{D}^\gamma} + \|f\|_{\mathcal{D}^\gamma}$
- there's also a pseudometric between $f \in \mathcal{D}^\gamma(T, g)$ and $f' \in \mathcal{D}^\gamma(T, g')$
- for a basis element $\sigma \in \mathcal{B} \subset T$ and an arbitrary $h \in T$, denote by h_σ its σ-th component

$$f(\cdot) = \sum_{\sigma \in \mathcal{B}} f_\sigma(\cdot)\, \sigma$$

- the consistency condition implies a kind of Taylor remainder estimate:

$$\left\| f_\sigma(y) - \sum_\tau f_\tau(x)\, g_{yx}(\tau) \right\| \lesssim d(y, x)^{\gamma - |\sigma|}$$

- **the archetype** of a modelled distribution is

$$f(x) = \sum_{|n|_\mathfrak{s} < \gamma} \frac{f^{(n)}(x)}{n!} X^n$$

- in the polynomial regularity structure. therefore,

$$f^{(n)}(y) - f^{(n)}(x) - \sum_{|\ell|_\mathfrak{s} < \gamma - |n|_\mathfrak{s}} \frac{1}{\ell!} f^{(n+\ell)}(x)(y-x)^\ell = O(d(y,x)^{\gamma - |n|_\mathfrak{s}})$$

- **reconstruction theorem:** let $\mathcal{T}$ be a concrete regularity structure and $M = (g, \Pi)$ a model over $\mathcal{T}$. fix $\gamma > 0$. there exist a unique linear continuous operator

  $$R^M : \mathcal{D}^\gamma(T, g) \to C^{\beta_0 \wedge 0}(\mathbb{R} \times \mathbb{R}^d), \quad \beta_0 = \min A,$$

  such that

  $$\left| \langle R^M f - \Pi_x^g f(x), p_t(x, \cdot) \rangle \right| \lesssim \|\Pi^g\| \cdot \|f\|_{\mathcal{D}^\gamma}\, t^{\gamma/4}$$

  uniformly in $f \in \mathcal{D}^\gamma(T, g)$, $x \in \mathbb{R} \times \mathbb{R}^d$ and $0 < t \leq 1$.
- recall the question we asked before: when a family of k-order jets can be glued into a global object? here's our answer
- the condition $R^M f \approx \Pi_x^g f(x)$ near x gives meaning to $f(x)$ as local approximation. before this condition, there's no reason to expect that $f(x)$ glues into a global object, since it's not even clear what kind of thing $f(x)$ is
- to understand the reconstruction theorem better, let's strip away the regularity structures context and frame the question behind this theorem as follows: given a family $\{F_x\}_{x \in \mathbb{R}^d} \subset S'(\mathbb{R}^d)$ (jet/germ of distributions), can we find $F \in S'(\mathbb{R}^d)$ such that $F(\cdot) \approx F_x(\cdot)$ for $\cdot \approx x$? the latter should be understood via test functions localized near x. reconstruction theorem can be completely extracted from theory of regularity structures and formulated analytically
- Gubinelli's sewing lemma in rough path threory is a precursor of Hairer's reconstruction theorem (and can even be derived from the latter!), both answer the same question: how much consistency we need to ask of local pieces to ensure they uniquely describe a global object? the answer is: $f \in \mathcal{D}^\gamma$ with $\gamma > 0$. similarly to sewing lemma, if $\gamma \leq 0$, neither existence, nor uniqueness can be ensured.

- proof consists in writing down an explicit formula for the reconstructed object and carefully showing that it makes sense:

$$R^M \boldsymbol{f}(x) = \lim_{t\to 0}\lim_{s\to 0} \int p_{t-s}(x,y)\langle \Pi_y^g \boldsymbol{f}(y), p_s(y,\cdot)\rangle\, dy$$

- here evaluation of $R^M\boldsymbol{f}$ at x should be understood in the distributional sense (by testing against a test function)
- the point, however, is not a specific formula for $R^M\boldsymbol{f}$. what we ultimately need is a guarantee that expansions determine some object uniquely, because it justifies working on the level of jets
- some nice facts about this reconstruction:
  - **corollary: reconstruction is local.** $\boldsymbol{f}|_U = \boldsymbol{0}$ for $U \subset \mathbb{R}\times\mathbb{R}^d$ open, then $R^M\boldsymbol{f}|_U = 0$
    - proof: $\boldsymbol{f}(x) = 0$ for $x \in U$ implies $\Pi_x^g\boldsymbol{f}(x) = 0$ on U (Π is a linear map), hence the bound defining R^M gives $|\langle R^M\boldsymbol{f}, p_t(x,\cdot)\rangle| \lesssim t^{\gamma/4}$. Since $\int \varphi(x)p_t(x,y)\,dx \to \varphi(y)$ as $t\to 0$ in an appropriate C^k space, we obtain

      $$\langle R^M\boldsymbol{f}, \varphi\rangle = \lim_{t\to 0}\int \varphi(x)\langle R^M\boldsymbol{f}, p_t(x,\cdot)\rangle\,dx = 0$$

      since $t^{\gamma/4}\to 0$
  - **corollary: if model is smooth,** then $R^M\boldsymbol{f}(x) = (\Pi_x^g\boldsymbol{f}(x))(x)$
    - i.e., if $\Pi : T \to C^\infty$
    - proof: by uniqueness
    - compare this with $f(x) = (f_n(x))_0$ in polynomial jets
- remark: natural question of "for which T-valued functions reconstruction theorem holds" (we have shown for $\mathcal{D}^\gamma$ with $\gamma > 0$) is solved, the corresponding iff-condition is γ-coherence given in terms of some simple analytic bound
- let's give one instructive example of how reconstruction theorem may be used to establish the following folklore fact: if $-\alpha+\beta > 0$, there exists a continuous product (bilinear map) $B : C^{-\alpha}\times C^\beta \to S'$ such that $B(f,g) = fg$ for all continuous f, g.
  - we'll start with the following observation: although products of functions and distributions are canonically defined by $\langle \xi f, \varphi\rangle = \langle \xi, f\varphi\rangle$ for smooth f (so that $f\varphi$ be a test function again), this defnition no longer makes sense if $f \in C^\beta$ (it would make sense if we defined $C^{-\alpha} = B^{-\alpha}_{\infty,\infty}$ as the dual of $C^\alpha = B^\alpha_{\infty,\infty}$, but the dual of C^α!). however, we can give *a local description* of ξf by replacing f with its Taylor polynomial $p_x(\cdot)$ up to $\lfloor\beta\rfloor$ and define $\xi\, p_x(\cdot)$ canonically. can we glue these pieces together into a globally defined distribution? reconstruction theorem says yes, provided that these local descriptions are *consistent enough*
  - the condition $\alpha+\beta > 0$ says exactly how many terms of Taylor expansion of f we need to take for a given $\alpha < 0$ to ensure that consistency conditions are strong enough to identify a unique object
  - take the following regularity structure: $A = \mathbb{N}\cup(\mathbb{N}-\alpha)$, $T = V\oplus W$ where V and W are copies of the polynomial regularity structure, *denote* the basis of V by X^k and the basis of W by ΞW^k (this is not product, just *names* for the objects) and postulate the grading: $X^k \in T_{|k|_\mathfrak{s}}$ and $\Xi X^k \in T_{|k|_\mathfrak{s}-\alpha}$
  - now, we build a model: $\Pi_x X^k(\cdot) = (\cdot - x)^k$ and $\Pi_x \Xi X^k(\cdot) = (\cdot - x)^k\xi(\cdot)$
  - both make sense because we're not multiplying $\xi \in C^{-\alpha}$ by $f \in C^\beta$ yet, only by a monomial, which is perfectly nice and smooth
  - now, let F be the lift of f into the polynomial structure: $F(x) = \sum_{|n|_\mathfrak{s}<\beta}\frac{\partial^n f(x)}{n!}X^k$
  - define $\Xi F(x)$ by $\Xi F(x) = \sum_{|n|_\mathfrak{s}<\beta}\frac{\partial^n f(x)}{n!}\Xi X^k$

- we can show now that $\Xi F \in \mathcal{D}^{\beta-\alpha}$ and hence we can define $B(\xi, f) = R^{\xi}(\Xi F)$
  - if ξ was a continuous function, $R^{\xi}(\Xi F)(x) = (\Pi_x \Xi F(x))(x) = \xi(x) F_{\mathbf{1}}(x) = \xi(x) f(x)$
- consider the case when $\boldsymbol{f}(x) = \tau$ for all x with $|\tau| < 0$. what does the reconstruction theorem say in this case depends on how τ behaves under reexpansion: $\|\tau - \widehat{g}_{yx}\tau\|_\beta$ may be bounded as follows:

$$\begin{aligned}\|\tau - \widehat{g}_{yx}\tau\|_\beta &= \left\| \tau - (\mathrm{Id} \otimes g_{yx}) \left(\tau \otimes \mathbf{1} + \sum_{|\sigma|<|\tau|} \sigma \otimes (\tau/\sigma) \right) \right\|_\beta \\ &= \left\| \tau - \tau - \sum_{|\sigma|<|\tau|} \sigma g_{yx}(\tau/\sigma) \right\|_\beta \\ &\le \sum_{|\sigma|<|\tau|} \|\sigma\|_\beta \cdot |g_{yx}(\tau/\sigma)| \\ &\le \sum_{|\sigma|<|\tau|} \|\sigma\|_\beta d(y,x)^{|\tau|-|\sigma|} \lesssim d(y,x)^{|\tau|-\beta}\end{aligned}$$

  and in definition of $\mathcal{D}^\gamma$ we need $\|\tau - \widehat{g}_{yx}\tau\|_\beta \lesssim d(y,x)^{\gamma-\beta}$. A priori, this says that $\boldsymbol{f}(x) = \tau$ belongs to $\mathcal{D}^\gamma$ for $\gamma < |\tau|$, but that's it. Note, however, that τ may actually reexpand better than assumed in the definition of the model, which gives an a priori worst case reexpansion behaviour. For example, if τ is constant under reexpansions, then $\|\tau - \widehat{g}_{yx}\tau\|_\beta = 0$ trivially and $\boldsymbol{f} \in \mathcal{D}^\gamma$ for all $\gamma > |\tau|$. Hence, constant jets invariant under reexpansion have unique reconstructions regardless of their degree!
- spaces $\mathcal{D}^\gamma$ can be thought of as Hölder spaces in the following sense: assume $\gamma \in (1,2)$ and take a pair of functions $F = (f, g)$. since pair has values in $\mathbb{R}^2$, we can choose any basis for $\mathbb{R}^2$, so I'll call the first basis vector $\mathbf{1}$ and the second X (why not?). Then,

$$F(x) = (f(x), g(x)) = f(x)\,\mathbf{1} + g(x)\,X.$$

  then, I assume that $F \in \mathcal{D}^\gamma$ with respect to polynomial regularity structure endowed with its canonical model. This means that $|f(x) - f(y)| \lesssim |x-y|^\gamma$ and $|f(x) - f(y) - g(x)(y-x)| \lesssim |x-y|^{\gamma-1}$. This actually *forces* $g(x) = f'(x)$ and $f \in C^\gamma$. Hence, $F \in \mathcal{D}^\gamma$ means here that F is a jet uniquely identifying an element of C^γ.
  - endowing this space of jets with product $\mathbf{1} \cdot \mathbf{1} = \mathbf{1}$, $X \cdot \mathbf{1} = \mathbf{1} \cdot X = X$, $X \cdot X = 0$, we obtain Leibniz rule: $F(x) \cdot G(x) = f(x)g(x)\,\mathbf{1} + (f'(x)g(x) + f(x)g'(x))\,X$

an assumption

- we henceforth assume that
  - our regularity structure (T^+, T) contains the polynomial regularity structure in the sense that there are symbols $X^n \in T$ and $X^n_+ \in T^+$ belonging to the correct grading level based on $|n|_\mathfrak{s}$ and such that Δ acts on them as usual $\Delta X^n = \sum_{\ell \le n} \binom{n}{\ell} X^l \otimes X^{n-\ell}$
- on the polynomial regularity structure, there's a canonical model
  - $g_x(X^n_+) = x^n$ and $\Pi X^n(y) = y^n$
- we assume that the model (g, Π) we're working on, when restricted to the polynomial regularity structure contained in (T^+, T), is canonical

products, derivatives, integrals

- **def.** $V \subset T$ is a subcomodule or a sector if it is a linear subspace and $\Delta V \subset V \otimes T^+$ (splits into itself + functions). define $V_\alpha = V \cap T_\alpha$.
- **def.** a product on $V \times W$ is a continuous bilinear map $\star : V \times W \to T$ such that

$$V_\alpha \star W_\beta \subset T_{\alpha+\beta}$$

- the product is said to be **regular** if $\Delta(\tau \star \sigma) = (\Delta\tau)(\Delta\sigma)$, where the product on the right is defined by

$$(V \otimes T^+) \times (W \otimes T^+) \to T \otimes T^+ : (\tau \otimes \mu)(\sigma \otimes \nu) = (\tau \star \sigma) \otimes (\mu\nu)$$

- recall that T^+ is an algebra, so $\mu\nu$ makes sense automatically
- regularity implies $\widehat{g}(\tau \star \sigma) = \widehat{g}(\tau) \star \widehat{g}(\sigma)$
- **product between T_X and T is always defined** and

$$1 \star \tau = \tau \star 1 = \tau \quad \text{and} \quad X^k \star X^\ell = X^{k+\ell}$$

- i.e., we can multiply distributions by polynomials
- define the canonical projection $Q_{<\gamma} : T \to T_{<\gamma}$.
  - this operator **truncates the jet**, i.e., forgets "higher derivatives"
  - in the polynomial rs, $Q_{<2}$ truncates the jet of a function by first two derivatives
- define D_α^γ the space of modelled distributions of the form

$$\boldsymbol{f} = \sum_{\alpha \le |\tau| < \gamma} f_\tau \, \tau.$$

- **prop: regularity of product.** let $\alpha_1, \alpha_2 \le 0 < \gamma_1, \gamma_2$. if $\star$ is a regular product on $V \times W$, then

$$Q_{<\gamma}(\boldsymbol{f}_1 \star \boldsymbol{f}_2) \in \mathcal{D}^\gamma_{\alpha_1+\alpha_2}(T, g) \quad \text{where} \quad \gamma = (\gamma_1 + \alpha_2) \wedge (\gamma_2 + \alpha_1)$$

  for $\boldsymbol{f}_1 \in \mathcal{D}^{\gamma_1}_{\alpha_1}(V, g)$ and $\boldsymbol{f}_2 \in \mathcal{D}^{\gamma_2}_{\alpha_2}(W, g)$. moreover, $(\boldsymbol{f}_1, \boldsymbol{f}_2) \mapsto Q_{<\gamma}(\boldsymbol{f}_1, \boldsymbol{f}_2)$ is continuous
- note that $\gamma \le \gamma_1 \wedge \gamma_2$. formally multiplying Taylor polynomial of f_1 up to order γ_1 and of f_2 up to order γ_2 doesn't magically give us information about $f_1 f_2$ up to order $\gamma_1 + \gamma_2$, only up to $\gamma_1 \wedge \gamma_2$, therefore we dispose those garbage coefficients
- **def.** a subcomodule V is **function-like** if it contains polynomial regularity structure and $V_\beta = 0$ for $\beta < 0$.
- on function-like comodules the reconstruction works the same as in the Taylor polynomial case:

$$\boldsymbol{f}(x) = (f(x), f'(x), f''(x), \ldots, f^{(N)}(x)) \rightsquigarrow R^M \boldsymbol{f}(x) = f(x)$$

- specifically,

$$R^M \boldsymbol{f}(x) = f_{\mathbf{1}}(x) = \text{projection of } \boldsymbol{f}(x) \text{ on } \mathbf{1} \in \mathcal{B}$$

- if V is a function-like subcomodule equipped with an associative product $\star : V \times V \to V$, then we can define for a smooth $F : \mathbb{R} \to \mathbb{R}$

$$F^\star(\boldsymbol{f}) = Q_{<\gamma}\left(\sum_{n=0}^\infty \frac{F^{(n)}(f_{\mathbf{1}})}{n!}(\boldsymbol{f} - f_{\mathbf{1}}\mathbf{1})^{\star n}\right)$$

- why does this definition make sense? supposing $\boldsymbol{f}$ is a jet of f, shouldn't $F(\boldsymbol{f})$ represent the jet of $F(f)$? it both should and does! as a matter of fact, we can either show this directly via Faa di Bruno formula, or by noting that applying non-linear function commutes with truncation: indeed, if $f(x+h) = P_x(h) + O(|h|^\gamma)$, then $F(f(x+h)) = F(P_x(h)) + O(|h|^\gamma)$ for any smooth F, the order of approximation is preserved. takeaway: composition with smooth function can be described on the level of jets
- **prop:** the mapping $\boldsymbol{f} \mapsto F^\star(\boldsymbol{f})$ is locally Lipschitz continuous
- note that we only defined $F^\star(\boldsymbol{f})$ if $\boldsymbol{f}$ is function-like. first off, this makes sense: $Q_{<\gamma}$ truncates from above, but if we plug in something of negative regularity Taylor series will immediately spit out distributions of all negative orders and we're doomed. but how do we handle u^3 in the Φ^4_3 then? well, notice that this issue doesn't happen as long as we plug it into polynomials: only finite number of terms appear and we never descend into stuff of regularity $-\infty$.
  - this still rules out sine-Gordon, which has $\sin(\beta u)$ in the equation
- **def.** a derivative is a continuous linear map $D : T \to T$ such that $DT_\alpha \subset T_{\alpha-1}$ and

$$\Delta D\tau = (D \otimes \mathrm{Id})\Delta\tau$$

- arguably, this is too general to be called a derivative since it doesn't satisfy any kind of Leibniz rule. but Leibniz rule specifies how derivative interacts with products, and we don't have products in T. instead, this definition says how derivatives interact with reexpansion (commutes) and homogeniety (lowers by one)
- note that this means that $\Delta\tau = \sum_\sigma \sigma \otimes \tau/\sigma$ hence $\Delta D\tau = \sum_\sigma (D\sigma) \otimes \tau/\sigma$
  - this means that we should only differentiate the actual symbol, not the transition data
  - example:
    * $\Delta X^n = \sum \binom{n}{\ell} X^{n-\ell} \otimes X^\ell$
    * $DX^n = nX^{n-1}$
    * $\Delta DX^n = n \sum \binom{n-1}{\ell} X^{n-1-\ell} \otimes X^\ell$
    * $(D \otimes \mathrm{Id})\Delta X^n = (D \otimes \mathrm{Id}) \sum \binom{n}{\ell} X^{n-\ell} \otimes X^\ell = \sum \binom{n}{\ell}(n-\ell) X^{n-1-\ell} \otimes X^\ell$
    * $\binom{n}{\ell}(n-\ell) = \binom{n-1}{\ell} n$
- most importantly,
  - the mapping $D : \mathcal{D}^\gamma \to \mathcal{D}^{\gamma-1} : \boldsymbol{f} \mapsto D\boldsymbol{f}$ is continuous
  - if $\mathscr{D}$ is a first order differential operator on S' and D is defined by $\Pi \circ D = \mathscr{D} \circ \Pi$, then **the reconstructions also agree**: $R^M D\boldsymbol{f} = \mathscr{D} R^M \boldsymbol{f}$ for $\boldsymbol{f}$ with $\gamma > 1$

lifting the Green function

- in this section we aim to build a model-dependent operator $\mathcal{K}^M$, which will serve us as the lifted version of $(\partial_{x_0} - \Delta_{x'})^{-1}$, intertwined with the latter via the reconstruction map
- from now on we assume that $\beta_0 = \min A > -2$. this condition ensures that heat kernel convolution (gaining $+2$ regularity by Schauder estimates) sends everything into functions. this is not a necessary assumption, it's only for the sake of simplicity. not all singular SPDEs satisfy it (KPZ does, Φ^4_3 doesn't, sine-Gordon doesn't, and so on)
- define $L = \partial_{x_0} - \Delta_{x'} + 1$
- within this chapter I'm going to sometimes lie and cut corners, because otherwise we won't have enough time
- **Schauder estimates:** $L^{-1} = K + K'$ where K maps $C^\gamma \to C^{\gamma+2}$ for all non-integer γ and K' maps $C^\gamma \to C^\infty$
  - the decomposition we have in mind is roughly similar to the decomposition of Green function into singular kernel and smooth remainder

- we need this split to separate the action of L^{-1}, which we're actually interested in, into something local-ish plus something smooth. the reason: we will ignore the smooth part K' altogether, because it will cause no harm, and approximate K further by something acting **on jets**, as if it was a local operator, up to some tractable corrections
- acting on jets means $u(\cdot) \approx (\Pi_x \boldsymbol{f}(x))(\cdot) \rightsquigarrow Ku \approx K(\Pi_x \boldsymbol{f}(x)((\cdot)))$
- in other words, K is designed to capture the jet-level action of K

- **example.** let's take polynomial regularity structure and try to add just one $f \in C^\gamma$ with $\gamma > 0$ to it
  - introduce its symbol F and define $T = \mathrm{span}(\mathcal{B}_X \cup \{F\})$, $\Pi F = f$
  - this is ok, but not enough for T^+, which needs to support recenterings

$$(\Pi_x^g F)(\cdot) = f(\cdot) - \sum_{|n|_\mathfrak{s} < \gamma} \frac{\partial^n f(x)}{n!} (\cdot - x)^n$$

  - to represent $\partial^n f$, we add new symbols $\{F_n\}_{|n|_\mathfrak{s} < \gamma}$ to T^+
  - we also need to exdend the definition of Δ:

$$\Delta F = F \otimes \mathbf{1} + \sum_{|n|_\mathfrak{s} < \gamma} \frac{X^n}{n!} \otimes F_n$$

  - everything on the left is acted on by Π, everything on the right is acted on by g_x
  - **remark:** recall that T encode types of local behaviours and T^+ encode expansion coefficients (something which g_x makes into a number); this is why we put F into T (as a new type of local behaviour in our expansion pallette) and F_n into T^+ (as describing how F reexpands from point to point). we could, of course, add F_n to T as well if we needed $\partial^n f$ type local behaviours, but here we're only talking about minimal extension by F
  - we also need to extend Δ^+:

$$\Delta^+ F_n = F_n \otimes \mathbf{1} + \sum_{|m|_\mathfrak{s} < \gamma - |n|_\mathfrak{s}} \frac{X^m}{m!} \otimes F_{n+m}$$

  - note that F_n is only defined for $|n|_\mathfrak{s} < \gamma$ where $\gamma : f \in C^\gamma$. in other words, we're not defining some weak derivatives here, only extrating as many classical derivatives as there is

- now, we introduce operator $\mathcal{I}_n$ which models the operator $\partial^n K$'s action **on T (monomials)**
  - assume that $\mathcal{B}^+$ is a commutative monoid freely generated by symbols $X_+^{e_i}$ and $\mathcal{I}_n^+ \tau$ with $\tau \in \mathcal{B}$, $n \in \mathbb{N} \times \mathbb{N}^d : |\tau| + 2 - |n|_\mathfrak{s} > 0$
  - i.e., applying K adds $+2$ to the regularity and using ∂^n reduces regularity by $-|n|_\mathfrak{s}$
  - homogenieties are defined by $|X_+^{e_i}| = \mathfrak{s}_i$, $|\mathcal{I}_n^+ \tau| = |\tau| + 2 - |n|_\mathfrak{s}$
  - $\Delta^+ \mathcal{I}_n^+ \tau = (\mathcal{I}_n^+ \otimes \mathrm{Id})\, \Delta\tau + \sum_{|\ell|_\mathfrak{s} < |\tau| + 2 - |n|_\mathfrak{s}} \frac{X_+^\ell}{\ell!} \otimes \mathcal{I}_{n+\ell}^+ \tau$
  - for $n : |n|_\mathfrak{s} \leq 1$, there is an operator $\mathcal{I}_n : T \to T$ such that $|\mathcal{I}_n \tau| = |\tau| + 2 - |n|_\mathfrak{s}$ and

$$\Delta \mathcal{I}_n \tau = (\mathcal{I}_n \otimes \mathrm{Id})\, \Delta\tau + \sum_{|\ell|_\mathfrak{s} < |\tau| + 2 - |n|_\mathfrak{s}} \frac{X^\ell}{\ell!} \otimes \mathcal{I}_{n+\ell}^+ \tau$$

  - for simplicity, we write $\mathcal{I} = \mathcal{I}_0$
  - notice that we're merely encoding relations between objects at this point, for instance, $\Delta \mathcal{I}_n \tau = \ldots$ encodes

$$\partial^n K f(y) \approx \sum \frac{1}{\ell!} \partial^{n+\ell} K f(x) (y - x)^k$$

    (second term) plus the fact that τ also has to be reexpanded from x to y

– recall that interpretation map $\Pi = (\Pi_x^g \otimes g_x)\,\Delta$ assigns values via Π_x^g on the left and via g_x on the right, so expressions for $\Delta\tau$ are automatically in the form amenable for interpretation

- **def.** Π is K-admissible iff $\Pi\mathcal{I}\tau = K\Pi\tau$ and $\Pi(X^n \star \tau)(x) = x^n(\Pi\tau)(x)$ for any $\tau \in \mathcal{B}$, $n \in \mathbb{N}^{1+d}$
  - remember that any distribution-symbol can be multiplied by monomial-symbol by our assumption
  - "at least the way we interpret homogenity templates τ themselves is compatible with K"
- our current goal is to upgrade $\Pi\mathcal{I}\tau = K\Pi\tau$ relation on T to $\mathcal{K}^M$ **on modelled distributions**, i.e., T-valued functions, such that $K \circ R^M = R^M \circ \mathcal{K}^M$
  - $\boldsymbol{f}(x) = \sum f_\tau(x)\,\tau$, our operator will act both in τ and in x
  - of course, $\mathcal{I}\boldsymbol{f}(x) = \sum f_\tau(x)\,\mathcal{I}\tau$, but this doesn't transform correctly under changes of basepoints: $\Delta\mathcal{I}\boldsymbol{f}(x) = \sum_\tau f_\tau(x)\,\Delta\mathcal{I}\tau$ and the dependence of f_τ on τ gets ignored
  - we're going to see though that $\mathcal{K}^M\boldsymbol{f} = \mathcal{I}\boldsymbol{f}$ + correction
  - recall that K isn't local (acting on jets), but it is local-ish
- so, what's exactly the issue?
  - up to now we only have $\Pi\mathcal{I}\tau = K\Pi\tau$, where Π is the interptretation map, but does this still hold when we recenter? $\Pi_x^g\mathcal{I}\tau \stackrel{?}{=} K\Pi_x^g\tau$? actually, no! only up to a polynomial...
- **prop.** $K(\Pi_x^g\tau) = \Pi_x^g(\mathcal{I}\tau + \mathcal{J}^M(x)\,\tau)$, where

$$\mathcal{J}^M(x)\,\tau = \sum_{|n|_\mathfrak{s} < |\tau|+2} \frac{X^n}{n!}\partial^n K(\Pi_x^g\tau)(x) \in T_X$$

- note that $K\Pi_x^g\tau$ has regularity $|\tau| + 2$, hence $n : |n|_\mathfrak{s} < |\tau| + 2$ derivatives of it are nice Hölder functions with values at points
- recall that we have assumed that all τ's are of homogeniety higher than -2. this was done so that one application of K be enough to make then functions. it wasn't strictly necessary, just helpful
- let's see this:

$$\begin{aligned}
\Pi_x^g\mathcal{I}\tau &= (\Pi \otimes g_x^{-1})\Delta\mathcal{I}\tau \\
&= \sum_\sigma (\Pi \otimes g_x^{-1})(\mathcal{I}\sigma \otimes \tau/\sigma) + \sum_{|\ell|_\mathfrak{s}<|\tau|+2} \frac{(\cdot)^\ell}{\ell!}\, g_x^{-1}\mathcal{I}_\ell^+\tau \\
&= \sum (\Pi\mathcal{I}\sigma)(g_x^{-1}(\tau/\sigma)) + \underbrace{P_x(\tau,\cdot)}_{\text{polynomial of deg at most } |\tau|+2} \\
&= \sum K(\Pi\sigma)\, g_x^{-1}(\tau/\sigma) + P_x(\tau,\cdot) \\
&= K(\Pi_x^g\tau) + P_x(\tau,\cdot)
\end{aligned}$$

- it remains to identify $P_x(\tau,\cdot) = -\Pi_x^g(\mathcal{J}^M(x)\,\tau)$
- okay, so K acts on $(\Pi_x^g\tau)(\cdot)$ as $\mathcal{I}\tau + \mathcal{J}^M(x)\,\tau$, is this all we need? well, not quite, because $(\Pi_x^g\boldsymbol{f}(x))(\cdot)$ is not exactly $R^M\boldsymbol{f}(\cdot)$ near x, it is so only up to a certain order, there's still some remainder

$$KR^M\boldsymbol{f}(\cdot) = K\Pi_x^g\boldsymbol{f}(x)(\cdot) + \underbrace{K(R^M\boldsymbol{f}(\cdot) - (\Pi_x^g\boldsymbol{f}(x))(\cdot))}_{=:(\mathcal{N}^M\boldsymbol{f})(x)}$$

- the idea up to now: f can be reconstructed from its jet $\boldsymbol{f} : \mathbb{R}\times\mathbb{R}^d \to T$ up to some remainder; hence, the action of K on f at x can be built from the jet $\boldsymbol{f}(x)$ up to some non-local remainder $(\mathcal{N}^M\boldsymbol{f})(x)$

- we can write down a formula for $\mathcal{N}^M$, although it's not highly useful (something akin Taylor with remainder)
- Now, we can define finally

$$(\mathcal{K}^M \boldsymbol{f})(x) = (\mathcal{I} + \mathcal{J}^M(x))\boldsymbol{f}(x) + (\mathcal{N}^M \boldsymbol{f})(x)$$

- zap! there is no a priori guarantee that $\mathcal{K}^M \boldsymbol{f} : \mathbb{R}\times\mathbb{R}^d \to T$ is a modelled distribution (i.e., satisfies reconstructability conditions), but this **turns out to be true**
- *formally,* although this is not rigorous and correct,

$$(\mathcal{K}^M \boldsymbol{f})(x) = \mathcal{I}\boldsymbol{f}(x) + \sum_{|\ell|_\mathfrak{s} < \gamma+2} \frac{X^\ell}{\ell!} \partial^\ell K(R^M \boldsymbol{f})(x)$$

- **theorem.** let the regularity structure $\mathcal{T}$ satisfy our assumptions (includes polynomials canonically + polynomials can be multiplied by distributions + includes integral operator symbols) and let the model (g, Π) be K-admissible on it. then $\mathcal{K}^M : \mathcal{D}^\gamma \to \mathcal{D}^{\gamma+2}$ is a continuous map satisfying $R^M \circ \mathcal{K}^M = K \circ R^M$

lifting the spde to its regularity structure

- there's an interesting question of building an admissible model which we're going to skip
- from now on we shall work specifically with generalized KPZ equation for illustration:

$$(\partial_{x_0} - \Delta_{x'})u = f(u)\zeta + \underbrace{g_2(u)\,(\partial_x u)^2 + g_1(u)\,\partial_x u + g_0(u)}_{=:g(u,\partial_x u)}, \quad u(0,x') = u_0(x')$$

  convert it into an integral equation

$$u = e^{x_0(\Delta_{x'}-1)}u_0 + L^{-1}(f\zeta + g)$$

  (recall that $L = \partial_{x_0} - \Delta_{x'} + 1$) and then lift this to

$$\boldsymbol{u} = \boldsymbol{h} + \mathcal{P}^M(f^\star(\boldsymbol{u})\,\Xi + g^\star(\boldsymbol{u}, D\boldsymbol{u})),$$

  where $\boldsymbol{h}$ is the lift of a smooth function $e^{x_0(\Delta_{x'}-1)}u_0$ and $\mathcal{P}^M = \mathcal{K}^M$ + something smooth plays the role of $L^{-1} = K$ + something smooth and Ξ is a special **noise symbol**
- to makes matters simpler, let's move from $\mathbb{R}^d$ to $\mathbb{T}^d$ (compactness)
- let P be the kernel of L^{-1}:

$$L^{-1}f(x) = \int_{\mathbb{R}\times\mathbb{T}^d} P(x-y)\,f(y)\,dy$$

- **prop.** under certain assumptions for a K-admissible model $M = (g, \Pi)$ there exists a continuous linear map $\mathcal{P}_t^M$ representing L^{-1} in some space of modelled distributions by

$$(R^M \mathcal{P}_t^M \boldsymbol{f})(x) = \int P(x-y)\,(R^M \boldsymbol{f})(y)\,dy$$

  and this map is appropriately bounded $\|\mathcal{P}_t^M \boldsymbol{f}\| \lesssim t^{\kappa/2}\|\boldsymbol{f}\|$
- I am being vague about specific spaces and norms here because we don't have enough time to go into such details, but it's important to understand why this proposition is called **multilevel Schauder estimates**: we bound L^{-1} acting on the jet $\boldsymbol{f}$ of f, hence the word *multilevel*

- i'm also ignoring some smooth cut-offs which are needed in some of the claims
- note that we have the intertwining property we wanted:

$$R^M \mathcal{P}_t^M \boldsymbol{f} = L^{-1} R^M \boldsymbol{f}$$

- we need one last **assumption** before we lift
- **def.** a regularity structure satisfying the assumptions we've already imposed is said to be **associated with equation** $(\partial_t - \Delta_x)u = f\zeta + g$ it contains subcomodules

$$S, \quad DS, \quad F, \quad N$$

  of T such that
  - each contains polynomial regularity structure T_X
  - S is for Solution, it is function-like and is used to represent u
  - DS is for Derivative of Solution, it represents objects like $\partial_{x'} u$
  - F is for Function, it is used to represent f and g nonlinearities
  - N is for Noise, it contains a special symbol Ξ representing ζ, it also contains singular objects $\partial u, (\partial u)^2, \ldots$
  - there are regular products

$$S \times \cdots \times S \to F, \quad DS \times DS \to N, \quad F \times N \to T,$$

    all three denoted by $\star$
  - there are abstract integration operators

$$\mathcal{I} : T \to S, \quad \mathcal{I}_{e_i} : T \to DS$$

    satisfying the assumptions we imposed on them above
  - partial derivatives act $D_i : S \to DS$ and satisfy

$$\Pi \circ D_i = \partial_{x_i} \circ \Pi, \quad D_i X^k = k_i X^{k-e_i} \mathbf{1}_{k \geq e_i}, \quad D_i \mathcal{I}\tau = \mathcal{I}_{e_i}\tau$$

- why is S function-like if we said initially that we're fine with distributional solutions? well, it turns out that our simplifying condition $\beta_0 > -2$ makes solutions functions.
- now, set up a fixed point version of the equation in the modelled distributions space:

$$\boldsymbol{u} = \Phi_t^{\boldsymbol{h},M}(\boldsymbol{u}) \quad \text{where} \quad \Phi_t^{\boldsymbol{h},M}(\boldsymbol{u}) = \boldsymbol{h} + \mathcal{P}_t^M(f^\star(\boldsymbol{u}) \star \Xi + g^\star(\boldsymbol{u}, D\boldsymbol{u}))$$

- let's think of this lifted expansion carefully:
  - $f^\star(\boldsymbol{u})$ is defined because $\boldsymbol{u}$ will be valued in a function-like sector S
  - then, $f^\star(\boldsymbol{u})\,\Xi$ is defined as $f^\star(\boldsymbol{u}) \star \Xi$ because we assumed that there's a product $F \times N \to T$
  - next, $g^\star(\boldsymbol{u}, D\boldsymbol{u})$ needs a proper definition, because $D\boldsymbol{u} \in DS$, which is not assumed to be function-valued. however, we have assumed g to be of the form

$$g(u, \partial_u x) = g_2(u)\,(\partial_x u)^2 + g_1(u)\,\partial_x u + g_0(u),$$

    hence $g^\star$ may be lifted by

$$g^\star(\boldsymbol{u}, D\boldsymbol{u}) = Q_{<\gamma}\left(g_2^\star(\boldsymbol{u}) \star (D\boldsymbol{u})^{\star 2} + g_1^\star(\boldsymbol{u}) \star (D\boldsymbol{u}) + g_0^\star(\boldsymbol{u})\right)$$

- we are **not** trying to define $F^\star(\boldsymbol{u})$ by Taylor expansion for non-function-like $\boldsymbol{u}$, since this would lead to negative regularities of infinite order! whenever we really need non-linear functions of distributions (as in sine-Gordon), we define them as separate symbols subject to their own renormalization, and all this goes beyond the scope of this mini-course
- **def.** a **solution** of the previous equation on $(0, t_0)$ is a fixed point of $\Phi_{t_0}^{\boldsymbol{h},M}$
- here are the two jewels of rs theory:
- **theorem.** assume that f and g are smooth functions, let $\mathcal{T}$ be a regularity structure associated with our equation and satisfying all assumptions we imposed. then for any K-admissible model $M = (g, \Pi)$ and any $\boldsymbol{h} \in \mathcal{D}^{\gamma,\eta}(T_X, g)$ there exists a positive time $t_0 = t_0(\boldsymbol{h}, M)$ such that the there exists a solution $\boldsymbol{u}$ on $(0, t_0)$.
- **theorem.** $u = R^M \boldsymbol{u}$ is a continuous function of $(\boldsymbol{h}, M)$
  - this is what we promised in the beginning! the solution map $(\zeta, \boldsymbol{h}) \mapsto u$ factorizes as $(\zeta, \boldsymbol{h}) \mapsto (M, \boldsymbol{h}) \mapsto u$ and the second arrow is continuous (assuming that the solution map $(M, \boldsymbol{h}) \mapsto \boldsymbol{u}$ is continuous, which it will be), whereas the first we're yet to build
- what does this theorem tells us exactly? it tells us that **once we have an admissible model**, the solution exists. we didn't need renormalization so far because renormalization is needed to **build an admissible model**
- if the noise ζ is smooth, we can build a *canonical* model M^ζ, which defines a solution $u = u^{M^\zeta}$. in this case u coincides with the classical solution of the well-posed equation $(\partial_t - \Delta_x)u = F(u, \zeta)$
- if the noise ζ is non-smooth and we mollify it ζ_ε, typically neither models M^{ζ_ε}, nor solutions $u_\varepsilon = u^{M^{\zeta_\varepsilon}}$ converge. defining $M^\zeta = \lim_{\varepsilon\to 0} M^{\zeta_\varepsilon}$ is in general impossible.
- our next goal is to build admissible models with $\Pi\Xi = \zeta$ and a non-smooth noise ζ. this is the subject of renormalization
- from now on $M^\varepsilon = (g^\varepsilon, \Pi^\varepsilon)$ stands for M^{ζ_ε} with $\zeta_\varepsilon = \zeta * \rho_\varepsilon$. the specific choice of mollifier ρ_ε won't matter

where do regularity structures come from?

- now's the right time to finally answer this question. let's consider Φ^4_3, for instance, which *formally* looks as follows: $\boldsymbol{u} = \mathcal{I}(\Xi - \boldsymbol{u}^3) + \text{smooth}$, and let's think how to build a regularity structure rich enough to represent such equations
- clearly, we need a symbol Ξ. since $d = 3$, we postulate $|\Xi| = -\frac{d+2}{2} = -\frac{5}{2}$
- to handle smooth part, we assume that the regularity structure contains T_X as before
- to save space, I'll denote all smooth parts by a single symbol p (for *polynomial*) from now
- we need both left and right hand sides of the equation to be represented by a modelled distribution in $\mathcal{D}^\gamma$ with $\gamma > 0$ for reconstruction theorem to work
- since $\mathcal{I}$ increases regularity, the goal seems plausible: we need to start with the worst types of local behaviour and add symbols until everything is described up to a positive order
- to guess the worst types of local behaviour, we use Picard iterations: $\boldsymbol{u}_0 = \mathcal{I}\Xi + p$, so we add $\mathcal{I}\Xi$ and postulate $|\mathcal{I}\Xi| = -\frac{5}{2} + 2 = -\frac{1}{2}$
- at the moment, we're postulating that $\mathcal{I}$ increases degrees by 2, so we'll need to check if this is consistent on the level of models. the consistency will be ensured by Schauder estimates.
- to write down $\boldsymbol{u}_1 = \mathcal{I}(\Xi - \boldsymbol{u}_0^3) + p$, we need to handle $\boldsymbol{u}_0^3$ first:

$$\boldsymbol{u}_0^3 = (\mathcal{I}\Xi + p)^3 = (\mathcal{I}\Xi)^3 + 3(\mathcal{I}\Xi)^2 p + 3(\mathcal{I}\Xi)p^2 + p^3,$$

  so we add
  - $(\mathcal{I}\Xi)^3$ with degree $|(\mathcal{I}\Xi)^3| = -\frac{3}{2}$

- $(\mathcal{I}\Xi)^2$ with degree $|(\mathcal{I}\Xi)^2| = -1$
- $(\mathcal{I}\Xi)^2 X_i$ with degree $-1 + 1 = 0$

- for each new symbol τ, we add $\tau \star X^k$ for all k to handle products with polynomials
- next, we continue adding symbols from Picard iterations until everything is of degree > 0
- the first iteration gives

$$\boldsymbol{u}_1 = \mathcal{I}(\Xi - \boldsymbol{u}_0^3) + p = \mathcal{I}\Xi - \mathcal{I}(\mathcal{I}\Xi)^3 - 3\mathcal{I}((\mathcal{I}\Xi)^2 p) - 3\mathcal{I}((\mathcal{I}\Xi)p^2)p - \mathcal{I}(p^3) + p,$$

  so we add
  - $\mathcal{I}(\mathcal{I}\Xi)^3$ with degree $-\frac{3}{2} + 2 = \frac{1}{2} > 0$
  - $\mathcal{I}((\mathcal{I}\Xi)^2)$ with degree $-1 + 2 = 1 > 0$
  - $\mathcal{I}((\mathcal{I}\Xi))$ with degree $-\frac{1}{2} + 2 = \frac{3}{2} > 0$
- next, we need $\boldsymbol{u}_1^3$, which involves (only terms we haven't seen before)
  - $(\mathcal{I}\Xi)^2\mathcal{I}(\mathcal{I}\Xi)^3$ of degree $-1 + \frac{1}{2} = -\frac{1}{2}$
  - $(\mathcal{I}\Xi)\mathcal{I}(\mathcal{I}\Xi)^3$ of degree $-\frac{1}{2} + \frac{1}{2} = 0$
  - $(\mathcal{I}\Xi)^2\mathcal{I}(\mathcal{I}\Xi)^2$ of degree $-1 + 1 = 0$
- one can check that everything higher than these is already of degree higher than 1, hence handled by polynomials
- now, we have a full list of basis elements we need the following ten symbols:

$$\Xi,\ \mathcal{I}\Xi,\ (\mathcal{I}\Xi)^2,\ (\mathcal{I}\Xi)^3,\ \mathcal{I}((\mathcal{I}\Xi)),\ \mathcal{I}((\mathcal{I}\Xi)^2),\ \mathcal{I}((\mathcal{I}\Xi)^3),$$

$$(\mathcal{I}\Xi)^2\mathcal{I}(\mathcal{I}\Xi)^3,\ (\mathcal{I}\Xi)\mathcal{I}(\mathcal{I}\Xi)^3,\ (\mathcal{I}\Xi)^2\mathcal{I}(\mathcal{I}\Xi)^2$$

- note that we never need symbols like Ξ^2 or $(\mathcal{I}\Xi)^4$, so we don't add them. we're not turning T into an algebra! the more symbols we add, the more interpretative work we'll need to do when defining a model
- we can add more, it will only increase the "resolution" γ of our jet. what we've added so far is enough to ensure uniqueness of reconstruction
- strictly speaking, we didn't even need to enlarge it with all polynomials. for instance, we don't actually need $(\mathcal{I}\Xi)^2 X^k$ for k other than $|k| = 1$, but this isn't an issue because products with polynomials are defined canonically, except for this case then $(\mathcal{I}\Xi)^2 X_i$ is of degree $-1 + 1 = 0$, hence may (and, in fact, does) diverge logarighmically and require renormalization
- since writing it like this quickly turns into a mess and also hides the algebraic structure of these trees, people turn to diagrammatics and encode them as rooted trees
- if we started with some noise other than white noise, we would need to postulate $|\Xi|$ as the regularity α of that noise. if $\alpha \in \left(-\frac{18}{7}, -\frac{5}{2}\right)$, the family of trees would be the same, whereas at $\alpha = -\frac{18}{7}$ we'd need to add $(\mathcal{I}\Xi)^2\mathcal{I}(\mathcal{I}(\Xi)^2\mathcal{I}(\mathcal{I}\Xi)^3)$ because its regularity

$$2(2 + \alpha) + 2 + 2(2 + \alpha) + 2 + 3(2 + \alpha) = 7\alpha + 18$$

  crosses zero at $\alpha = -\frac{18}{7}$. If α is even smaller, we'd need to add even more trees

renormalization structures

- recall the interpretation of a symbol $\tau \in T$ as one type of local behaviour, typically arising from Picard iterations. some of these symbols correspond to things which are problematic, i.e., represent undefined quantities, such as products of distributions. renormalization is a formal way to chop them off.

- within this section we set $U := T$ and $\mathcal{U} := \mathcal{B}$ (in principle, the notion of renormalization structure is independent of a regularity structure, but we'll only work with so-called *compatible* renormalization structures)
- however, when chopping the "bad" pieces off, we want to keep them in order to regularize them later. this idea chop-and-keep is formalised by introducing a splitting map

$$\delta : U \to U^- \otimes U,$$

  where U^- is a new set which stores these bad pieces. the expression $\delta\tau$ should be thought of as

$$\delta\tau = \mathbf{1}_- \otimes \tau + \sum_{\sigma \text{ divergent subforests of } \tau} \sigma \otimes \tau/\sigma$$

  with all divergent pieces removed and stored in U^-
- now it's time to give a formal definition:
- **def.** a renormalization structure $\mathcal{U}$ is a pair of graded vector spaces

$$U = \bigoplus_{\beta \in B} U_\beta, \quad U^- = \bigoplus_{\alpha \in B^-} U^-_\alpha,$$

  such that
  - U^-_α and U_β are finite dimensional
  - U^- is a connected bialgebra with unit $\mathbf{1}_-$, counit $\mathbf{1}'_-$ and coproduct

$$\delta^- : U^- \to U^- \otimes U^-$$

    and grading $0 \in B^- \subset (-\infty, 0]$
  - $B \subset \mathbb{R}$ is locally finite and bounded from below
  - U is a left comodule over U^-, i.e., there's a splitting map

$$\delta : U \to U^- \otimes U$$

    which satisfies

$$(\mathrm{Id} \otimes \delta)\delta = (\delta^- \otimes \mathrm{Id})\delta \quad \text{and} \quad (\mathbf{1}'_- \otimes \mathrm{Id})\delta = \mathrm{Id}$$

  - the renormalization structure is denoted by $\mathcal{U} = (U, U^-)$
- let's make a few comments on what all the objects are and how they are related in the case of Picard iteration trees
  - a tree is a connected loopless graph; it is rooted if there's a distinguished node called root; a decorated tree is a tree endowed with some functions on nodes and edges (assigning to a node its type or just some number)
  - let's say that a **decorated rooted tree** is *real* if it actually appears from Picard iterations
  - then $T = U$ (as sets) consists of
    - real trees
    - subtrees of real trees
    - trees obtained from real trees by contraction of *divergent* subtrees (of negative naive homogeneity)
  - the grading on U is the so-called *naive* homogeneity (is tree divergent or not?)
  - the grading on T is the so-called *extended* homogeneity: some nodes n are marked by an inscription "this node is a contraction of a tree whose homogeneity was $\mathfrak{o}(n)$" (is the tree divergent if we count with extracted pieces?)

  - T^+ consists of trees with positive *extended* homogeniety which are obtained from real ones by contracting divergent subforests, its endowed with the tree product (joining at the root — actual product) which makes it an algebra
  - U^- consists of real trees of negative *naive* homogeniety and enlarged to an algebra by including all forest products (disjoint union of trees — just keep the list of stuff to renormalize later) and unit $\mathbf{1}_- = \varnothing$
  - U^- should be thought of as the opposite of T^+, it stores divergent extracted pieces as opposed to storing what's left
  - **def.** the spde is **subcritical** if there are finitely many *real* trees of naive regularity less than γ
- now we have two Hopf algebras: (T^+, Δ^+) and (U^-, δ^-), and similarly to how Δ^+ induced a convolution product, so does δ^- on the set of characters G^- of U^- by

$$(k_1 * k_2)\,\tau = (k_1 \otimes k_2)\,\delta^- \tau$$

- the inverse of a character is given by $k \circ S_-$, where S_- is the antipode of Hopf algebra (U^-, δ^-) (connected bialgebra always has an antipode)
- leaping ahead, G^- (or rather its subgroup) is going to be the **renormalization group**, which makes sense, since $k \in G^-$ assigns numerical values to divergent pieces. these values will be our regularizations
- subcriticality makes G^- finite dimensional
- similarly to how each character $g \in G^+$ induced $\widehat{g} : T \to T$, a character $k \in G^-$ induces a linear map $\widetilde{k} : U \to U$ given by

$$\widetilde{k} = (k \otimes \mathrm{Id})\delta$$

- i.e., G^- acts on U from right. this is going to be (almost) the renormalization group action
- **def.** let's say that a regularity structure $\mathcal{T} = (T^+, T)$ and a renormalization structure $\mathcal{U} = (U, U^-)$ are **compatible** if
  - $T = U$ as vector spaces, $\mathcal{B} = \mathcal{U}$ (bases), gradings may be different ($\mathcal{B}_\beta \neq \mathcal{U}_\beta$)
  - compatibility with grading: $\delta T_\beta \subset U^- \otimes T_\beta$ for all β
    - i.e., δ chops off pieces into U^- and keeps remaining regularity at β
  - there exists an algebra morphism $\delta^+ : T^+ \to U^- \otimes T^+$ (encoding recentering the extraction of bad pieces) which is compatible with everything else (we skip precise conditions)
  - most importantly, the grading T_β shall be used to describe local regularity and scaling of U_β shall be used to describe divergence as $\varepsilon \to 0$ of regularized versions
- since $T = U$, $\widetilde{k} : T \to T$ and $\widetilde{k}^+ := (k \otimes \mathrm{Id})\delta^+ : T^+ \to T^+$
- **def.** a character $k \in G^-$ acts on the space of models by

$$M = (g, \Pi) \mapsto {}^k M = (g \circ \widetilde{k}^+, \Pi \circ \widetilde{k})$$

  and this action is compatible with g and Δ
- unfortunately, ${}^k M$ isn't always K-admissible! let's impose one more assumption to make it admissible:
- **assumption.** let $\mathcal{F} = \{\mathcal{I}_p\}_{|p|_\mathfrak{s} \le 1} \cup \{X^n \star\}_{n \in \mathbb{N}^{d+1} \setminus \{0\}}$
  - here $X^n \star : T \to T$ maps $\tau \mapsto X^n \star \tau$ and same for $\mathcal{I}_p$
  - assume that U^- is the algebra generated by $\mathcal{U}_{<0} = \bigcup_{\alpha<0} \mathcal{U}_\alpha$
    - previously we made no assumptions on how U and U^- are related at all!
  - let $\mathfrak{J}^-$ be the ideal of U^- generated by $\mathcal{F}(\mathcal{U}) \cap \mathcal{U}_{<0}$
    - $\mathcal{F}(U) \cap \mathcal{U}_{<0}$ = stuff which is still of negative degree after applying $\mathcal{I}_p$ or $X^n\star$

- assume that for every $F \in \mathcal{F}$ and $\tau \in \mathcal{U}$ holds

$$\delta(F\tau) - (\mathrm{Id} \otimes F)\delta\tau \in \mathfrak{J}^- \otimes U,$$

  which should be read as follows: δ prepares objects for renormalization by splitting them into those that need renormalization and those that are fine as is. the condition here demands that δ commutes with basic operations (integration and product by polynomials) up to stuff which we're not renormalizing anyway (see below)
  - define a projection operator $P_- : U \to U^-$ by $P_-\tau = \tau \mathbf{1}_{\tau \in \mathcal{U}_{<0}}$
  - the linear map $\delta^- : U^- \to U^- \otimes U^-$ is defined by $\delta^- = (\mathrm{Id} \otimes P_-)\delta$ on $\mathcal{U}$ and extended multiplicatively
- now, define

$$G^-_{\mathrm{ad}} = \{k \in G^- : k(F(\tau)) = 0 \quad \text{for all} \quad F(\tau) \in \mathcal{F}(\mathcal{U}) \cap \mathcal{U}_{<0}\}$$

- i.e., characters which don't touch the image of $\mathcal{F}$. in other words, allowed renormalizations are those that don't renormalize objects built from more elementary problematic pieces **independently**: renormalizing $F\tau$ should be the same as renormalizing τ and applying F later
- **prop.** $G^-_{\mathrm{ad}} \subset G^-$ is a subgroup and for $k \in G^-_{\mathrm{ad}}$ the model ${}^kM$ is K-admissible.
- **def.** G^-_{ad} is called the renormalization group
- now, let's ζ be a *smooth* noise. define a character h^ζ on $\mathbb{R}[U]$ by

$$h^\zeta(\mathbf{1}_-) = 1 \quad \text{and} \quad h^\zeta(\tau) = \mathbb{E}\{\Pi^\zeta \tau(0)\},$$

  where Π^ζ is the unique canonical model when noise is smooth (we only mentioned it, but didn't build).
- to make this into a character on U^-, we need some special algebra morphism

$$S'_- : U^- \to \mathbb{R}[U]$$

  called the *negative twisted antipode.* We don't have time to construct it, but it's meaning is purely combinatorial: it builds correct counterterms for every divergent piece, accounting for all nested sub-divergences
  - naive idea: see divergent $\Pi^\varepsilon\tau$ — subtract its expectation $\mathbb{E}\{\Pi^\varepsilon\tau\}$
  - but what if τ is divergent under Π^ε and contains divergent sub-piece φ, then the expectation $\mathbb{E}\{\Pi^\varepsilon\tau\}$ is itself divergent and in need of regularization
  - S'_- does this recursively
- **def: BHZ character**

$$k^\zeta := h^\zeta \circ S'_-$$

- **theorem.** $k^\zeta \in G^-_{\mathrm{ad}}$ and $\mathbb{E}\{({}^{k^\zeta}\Pi^\zeta\tau)(x)\} = 0$ for all x and $\tau \in \mathcal{U}_{<0}$. moreover, k^ζ is the **unique** character with this property
  - note that we've only centered at $x = 0$, but the thing is centered everywhere
- now, if k_1, k_2 are two characters on U^-, then

$${}^{k_1 * k_2}\Pi^\zeta = {}^{k_1}({}^{k_2}\Pi^\zeta)$$

- here's the punchline: let $k_\varepsilon = h^{\zeta_\varepsilon} \circ S'_-$. then ${}^{k_\varepsilon}\Pi^\varepsilon$ converges, and moreover ${}^{k * k_\varepsilon}\Pi^\varepsilon$ converges as $\varepsilon \to 0$ for every $k \in G^-_{\mathrm{ad}}$. this means that we've described the whole class of renormalization schemes as indexed by G^-_{ad} !
- **Chandra-Hairer convergence theorem:** renormalized models ${}^{k_\varepsilon}\Pi^\varepsilon$ converge in the d_γ-metric

- moreover,
  - $u^{(k)} \in C^\eta(\mathbb{R}\times\mathbb{R}^d)$ for $\eta \in (0, \beta_0 + 2]$ (provided that u_0 doesn't make regularity worse)
  - moreover, $\{u^{(k)}\}_{k\in G^-_{\mathrm{ad}}}$ is a finite dimensional submanifold of C^η
- **example:** PAM in $d = 2$, simplest model
  - noise Ξ has regularity $-\frac{d}{2} = -1$,
  - hence $\mathcal{I}\Xi$ has regularity $+1$,
  - hence $(\mathcal{I}\Xi)\Xi$ has regularity 0 (borderline),
  - hence $\mathcal{I}(\mathcal{I}(\Xi)\Xi)$ has regularity $+2$, but it contains a divergent subtree
  - in this case $U^- = \mathrm{span}\{\mathcal{I}(\Xi)\Xi\}$, hence one-dimensional, hence $G^- \cong \mathbb{R}$
  - BHZ character: $k_\varepsilon(\mathcal{I}(\Xi)\Xi) = \mathbb{E}\{\Pi^\varepsilon(\mathcal{I}(\Xi)\Xi)(0)\} = C_\varepsilon \sim c\ln\frac{1}{\varepsilon}$
  - renormalized model $(G * \xi_\varepsilon)\xi_\varepsilon \rightsquigarrow (G*\xi_\varepsilon)\xi_\varepsilon - C_\varepsilon$
  - renormalized equation $(\partial_t - \Delta)u_\varepsilon = u_\varepsilon(\xi_\varepsilon - C_\varepsilon)$ with $C_\varepsilon = c\ln\varepsilon^{-1}$
  - full family of renormalized equations:

$$(\partial_t - \Delta)u_\varepsilon = u_\varepsilon(\xi_\varepsilon - C_\varepsilon - k), \quad k \in \mathbb{R}$$

- renormalization procedure may thus be summarized as follows:
  - analytical step: construct the space of models (M, d_γ) and establish continuity of the solution map (Hairer 2014)
  - algebraic step: describe the group action on the space of models $M \mapsto {}^kM$ (Bruned-Hairer-Zambotti 2019)
  - probabilistic step: prove the convergence of renormalized models (Chandra-Hairer 2016)
  - second algebraic step: identify the renormalized equation satisfied by u_ε (Bruned-Chandra-Chevyrev-Hairer 2017)
- important things we didn't cover:
  - the dual action of G^- on the space of nonlinearities $\mathcal{F}$. in other words, how exactly does the new renormalized equation look? answer: some constants C_ε are subtracted in different places, but need to say which constants ($\mathbb{E}\{\dots\}$) and where
  - most importantly, how to build the rs from a given equation: trees, decorated trees, ... — and how S'_- acts; we almost never gave concrete interpretations of Δ, δ, ..., stayed abstract
  - construction of the canonical model when noise is smooth
  - a lot of analytic details

stochastic quantization, euclidean field theory

- Φ^4_3 model in QFT corresponds to the following action:

$$S[\varphi] = \int \left(\frac{1}{2}(\nabla\varphi)^2 + \frac{m}{2}\varphi^2 + \frac{\lambda}{4}\varphi^4\right) dx$$

- QFT, as far as I understood, is interested in the Gibbs measure (and even less, only correlation functions/expectations/moment generating function) corresponding to this action and given formally by

$$d\nu(\varphi) \propto e^{-S(\varphi)}\,\mathcal{D}\varphi,$$

  but no such $\mathcal{D}\varphi$ exists, so we need to build this measure somehow else
- one idea is to write this as

$$d\nu(\varphi) \propto e^{-\int\left(\frac{m}{2}\varphi^2 + \frac{\lambda}{4}\varphi^4\right)dx}\, d\mathcal{G}(\varphi) \quad \text{or} \quad d\nu(\varphi) \propto e^{-\frac{\lambda}{4}\int \varphi^4\,dx}\, d\mathcal{G}_m(\varphi),$$

where $d\mathcal{G}$ is the Gaussian free field measure and $d\mathcal{G}_m$ is the massive Gaussian free field measure, but neither $e^{-\int(\frac{m}{2}\varphi^2+\frac{\lambda}{4}\varphi^4)\,dx}$, nor $e^{-\frac{\lambda}{4}\int\varphi^4\,dx}$ are valid densities with respect to $d\mathcal{G}$ and $d\mathcal{G}_m$

- in finite dimensional world, Gibbs measures can be build from SDEs as follows:

$$\mu(dx) \propto e^{-S(x)}\,dx \quad \text{is the invariant measure of} \quad dx_t = -S'(x_t)\,dt + \sqrt{2}\,dB_t$$

- Parisi-Wu prescription: $d\nu(\varphi)$ should be built similarly:

$$\partial_t \Phi = -\frac{\delta S}{\delta \Phi} + \xi,$$

where ξ is the space-time white noise

- "to dress a classical path $\delta S = 0$ into quantum fluctuations", "stochastic quantization"
- it remains to note that

$$\frac{\delta S}{\delta \varphi} = -\Delta\varphi + m\varphi + \lambda\varphi^3$$

to obtain the Φ^4_3 spde

$$\partial_t \Phi = \Delta\Phi - m\Phi - \lambda\Phi^3 + \xi$$