Controlled paths and applications
A minicourse in Berlin

M. Gubinelli

December 2012.

Abstract. The aim of this series of talks is to give a perspective on the use of the idea of "controlled path" in various situations involving stochastic analysis and non-linear systems.

1. Introduction

A series of three lectures at Berlin the week of December, 3rd 2012:

- Tue 13:15–15:15 Room 1023.
- Wed 15:00–17:00. TU RTG Lounge 748.
- Thu HU 13:00–15:00 Room 1023 (same as Tue).

Topics I would like to touch (maybe briefly):

1. Regularization by noise of ODEs
   - Catellier, R., and M. Gubinelli. 2012. "On the Regularization of ODEs via Irregular Pertubations." arXiv:1205.1735.
   - Davie, A. M. 2007. "Uniqueness of Solutions of Stochastic Differential Equations." International Mathematics Research Notices. IMRN (24): Art. ID rnm124, 26.
   - Davie, A. M. 2011. "Individual Path Uniqueness of Solutions of Stochastic Differential Equations." In Stochastic Analysis 2010, 213–225. Heidelberg: Springer. http://www.ams.org/mathscinet-getitem?mr=2789085.
2. Regularization by noise in stochastic dispersive equations
   - M. Gubinelli and K. Chouk, "Nonlinear PDEs with stochastic dispersion", work in progress.
3. Controlled paths and rough paths (of course, or maybe not)
4. Stochastic partial differential equations driven by convolutional rough paths
   - Deya, A., M. Gubinelli, and S. Tindel. 2012. "Non-linear Rough Heat Equations." Probability Theory and Related Fields 153 (1): 97–147. doi:10.1007/s00440-011-0341-z.
5. Controlled distributions (work P. Imkeller and N. Perkowski)
   - Gubinelli, M., P. Imkeller, and N. Perkowski. 2012. "Paraproducts, Rough Paths and Controlled Distributions." arXiv:1210.2684.
6. Burgers like equations (Hairer's approach)
   - Friz, Peter K., Benjamin Gess, Archil Gulisashvili, and Sebastian Riedel. 2012. "Spatial Rough Path Lifts of Stochastic Convolutions." arXiv:1211.0046.
   - Hairer, M. 2011a. "Rough Stochastic PDEs". Communications on Pure and Applied Mathematics 64 (11): 1547–1585. doi:10.1002/cpa.20383.
   - Hairer, Martin, and Hendrik Weber. 2010. "Rough Burgers-like Equations with Multiplicative Noise." arXiv:1012.1236.

7. Energy (weak) solutions of KPZ equation (joint work with M. Jara)
   - Gubinelli, M., and M. Jara. 2012. "Regularization by Noise and Stochastic Burgers Equations." arXiv:1208.6551.
8. Hairer's solutions for the KPZ equation
   - Hairer, M. 2011. "Solving the KPZ Equation.". arXiv:1109.6811.
9. Tzvetkov and Burq's approach to super-critical solutions of non-linear wave equations and extensions of A. Poiret to the nonlinear Schrödinger equation
   - Burq, Nicolas. 2010. "Random Data Cauchy Theory for Dispersive Partial Differential Equations." In Proceedings of the International Congress of Mathematicians. Volume III, 1862–1883. New Delhi: Hindustan Book Agency. [mathscinet].
   - Burq, Nicolas, and Nikolay Tzvetkov. 2008a. "Random Data Cauchy Theory for Supercritical Wave Equations. II. A Global Existence Result." Inventiones Mathematicae 173 (3): 477–496. doi:10.1007/s00222-008-0123-0.
   - Burq, Nicolas, and Nikolay Tzvetkov. 2008b. "Random Data Cauchy Theory for Supercritical Wave Equations. I. Local Theory." Inventiones Mathematicae 173 (3): 449–475. doi:10.1007/s00222-008-0124-z.
   - Poiret, Aurélien. 2012. "Solutions Globales Pour Des Équations De Schrödinger Surcritiques En Toutes Dimensions", arXiv:1207.3519.

The first and most important example of controlled path are Itô's processes:

$$\mathrm{d}X_t = a_t \mathrm{d}W_t + b_t \mathrm{d}t$$

for them it is possible to give a meaning to integrals of the form

$$\int_0^t f_s \mathrm{d}X_s$$

for very general random functions f and to solve differential equations by interpreting them as stochastic differential equations, that is as equalities of certain random processes

$$X_t = x + \int_0^t b(X_s)\mathrm{d}W_s$$

that is X has to be an Itô process whose Brownian component is $b(X_s)$ and without component of bounded variation. In this way it has been possible to costruct a very successful theory for differential equations driven by Brownian motion and more general martingales.

Objects indexed by one parameter (time) are very special: they have a past and a future and the present is very simple. Here we are looking to alternative ways to expose a basic structure (in Itô approach *the independence of the increments* of the Brownian motion, or more generally the underlying martingale structure of the probability space) and put it in use to define interesting quantities (e.g. *stochastic integrals*) and use them to *describe and solve* differential equations or other similar non-linear relations among different quantities.

2. Regularization by noise in ODEs

(based on joint work with R. Catellier and on the paper: A.M. Davie "Uniqueness of solutions of stochastic differential equations", arXiv:0709.4147)

Consider the ODE

$$x_t = x_0 + \int_0^t b(x_s)\mathrm{d}s + w_t \tag{1}$$

where $w \in C([0,1],\mathbb{R}^d)$ is an irregular perturbation, $x \in C([0,1],\mathbb{R}^d)$. The space dimension d is arbitrary. The vector field b: $\mathbb{R}^d \to \mathbb{R}^d$ will be irregular. For example you can think of b as a distribution (e.g. $b(x) = v\delta(x)$ with $v \in \mathbb{R}^d$). The aim is first to give a meaning to this equation and then to study its properties (existence, uniqueness).

2.1. Basic setup.

Let us return to the ODE (1). Suppose that $b \in B^\alpha_{\infty,\infty}(\mathbb{R}^d)$ with $\alpha < 0$ and w a fractional Brownian motion (fBM) of Hurst index H sufficiently small $H \ll 1$. Recall that a fBM is a Gaussian process with zero mean and stationary increments with covariance

$$\mathbb{E}[|w_t - w_s|^2] = |t-s|^{2H}$$

A fBm of Hurst index H is Hölder continuous for any index $<H$. Moreover Brownian motion corresponds to the case $H = 1/2$.

Recall the definition of the Besov space $C^\alpha(\mathbb{R}^d) = B^\alpha_{\infty,\infty}(\mathbb{R}^d)$: $f \in C^\alpha(\mathbb{R}^d)$ means that f is a distribution satisfying

$$\|f\|_{B^\alpha_{\infty,\infty}} = \sup_{i \geqslant -1} 2^{-i\alpha} \|\Delta_i f\|_{L^\infty(\mathbb{R}^d)} < +\infty$$

where for $i \geqslant 0$, the Littlewood-Paley operators $\Delta_i f$ are defined as Fourier multipliers:

$$\widehat{\Delta_i f}(\xi) = \varphi(2^{-i}|\xi|)\hat{f}(\xi)$$

where φ: $\mathbb{R}_+ \to \mathbb{R}_+$ is a smooth function with support in $[1/2, 2+1/2]$, such that $\varphi(\xi) = 1$ for $\xi \in [1,2]$. It is defined as $\varphi(\xi) = \psi(2^{-1}\xi) - \psi(\xi)$ where the function ψ: $\mathbb{R}_+ \to \mathbb{R}_+$ is smooth and with support on $[0, 1+1/2]$ and $\psi(\xi) = 1$ on $[0,1]$. In this way

$$\psi(\xi) + \sum_{i \geqslant 0} \varphi(2^{-i}\xi) = \psi(\xi) + \sum_{i \geqslant 1} \psi(2^{-i}\xi) - \sum_{i \geqslant 0} \psi(2^{-i}\xi) = 1$$

for all $\xi > 0$. We let $\mathcal{F}(\Delta_{-1} f)(\xi) = \psi(|\xi|)\hat{f}(\xi)$ in such a way that

$$f = \sum_{i \geqslant -1} \Delta_i f$$

as distributions. It is possible to show that this decomposition of f as a sum of smooth (actually holomorphic) functions converges also in $B^\alpha_{\infty,\infty}$. Another useful property of the Littlewood-Paley operators is that they correponds to convolutions with kernels $K_i(x) = 2^{id}K(2^i x)$

$$\Delta_i f(x) = \int_{\mathbb{R}^d} K_i(x-y) f(y) \mathrm{d}y$$

and K is a $L^1(\mathbb{R}^d)$ smooth kernel which decay at infinity faster than any polynomial and such that

$$\int K(x)\mathrm{d}x = 0$$

since its Fourier transform is φ and so it is supported away from 0.

For example $\delta \in C^{-d}(\mathbb{R}^d)$: indeed $\Delta_i\delta(x) = K_i * \delta(x) = K_i(x) = 2^{id}K(2^i x)$ and

$$\|\Delta_i \delta\|_{L^\infty(\mathbb{R}^d)} \lesssim 2^{id}.$$

When $\alpha \in (0,1)$ these are the usual Hölder spaces, in particular when $\alpha > 0$ these are spaces of continuous functions. Indeed by estimating the LP decomposition in two different ways and using the estimate

$$\|D\,\Delta_i f\|_{L^\infty} \simeq 2^i \|\Delta_i f\|_{L^\infty}$$

(easy to prove using the representation of $\Delta_i f$ as a convolution) we have

$$|f(x)-f(y)| \leqslant \sum_{i\geqslant -1} |\Delta_i f(x)-\Delta_i f(y)|$$

$$\leqslant \sum_{i\geqslant -1} (\mathbb{I}_{|x-y|\leqslant 2^{-i}} 2\|\Delta_i f\|_{L^\infty} + \mathbb{I}_{|x-y|>2^{-i}} \|\mathrm{D}\Delta_i f\|_{L^\infty}|x-y|)$$

$$\lesssim \|f\|_{C^\alpha} \sum_{i\geqslant -1} 2^{-i\alpha}(\mathbb{I}_{|x-y|\leqslant 2^{-i}} + \mathbb{I}_{|x-y|>2^{-i}} 2^i|x-y|)$$

$$\lesssim \|f\|_{C^\alpha}|x-y|^\alpha.$$

On the other side if f is α-Hölder continuous we have

$$\Delta_i f(x) = \int_{\mathbb{R}^d} K_i(x-y)f(y)\mathrm{d}y = \int_{\mathbb{R}^d} K_i(x-y)(f(y)-f(x))\mathrm{d}y$$

(recall $\int K(x)\mathrm{d}x = 0$) and thus

$$|\Delta_i f(x)| \lesssim C_f \int_{\mathbb{R}^d} |K_i(x-y)||x-y|^\alpha \mathrm{d}y \lesssim C_f 2^{-i\alpha} \int_{\mathbb{R}^d} |K(y)||y|^\alpha \mathrm{d}y \lesssim C_f 2^{-i\alpha}$$

where

$$C_f = \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|^\alpha}.$$

2.2. Averaging properties of fBm.

The key idea in approaching the problem of giving a meaning to (1) is that $b(x_s)$ does not necessarily make sense for any $s\in[0,1]$ but the avereraged version provided by the integral over the time parameter is well defined

$$\int_0^t b(x_s)\mathrm{d}s.$$

So we will try to proceed by regularizing b by convolution $b_n \to b$ and look at the convergence as $n\to+\infty$ of

$$\int_0^t b_n(x_s)\mathrm{d}s$$

in some suitable sense. For definiteness take

$$b_n = \sum_{i\leqslant n} \Delta_i b = Q_n * b.$$

where $Q_n(x) = 2^{nd}Q(2^n x)$ and Q is a smooth kernel. The first observation is the following: take $x_s = w_s + x$ for some fixed $x\in\mathbb{R}^d$ and look at

$$\sigma_t^w b_n(x) = \int_0^t b_n(x+w_s)\mathrm{d}s = L_t^w * b_n$$

where $L_t(y)$ is the local time of w at y and is formally given by the distribution

$$L_t(y) = \int_0^t \delta(y-w_s)\mathrm{d}s.$$

Now it is possible to show that if w is fbm of Hurst index H and $b\in C^\alpha = B^\alpha_{\infty,\infty}$ with $\alpha > -1/(2H)$ tha averaged vector field $\sigma_t^w b_n$ has a limit as $n\to\infty$ for almost every sample of w. Moreover the limit does not depend on the type of regularization we use, meaning that taking another smooth kernel ρ and letting $b_\varepsilon = \rho_e * b$ with $\rho_\varepsilon = \varepsilon^{-d}\rho(x/\varepsilon)$ we have

$$\lim_{\varepsilon\to 0} \sigma_t^w b_\varepsilon(x) = \lim_{n\to+\infty} \sigma_t^w b_n(x).$$

In that case we can meaningfully define

$$\int_0^t b(x+w_s)\mathrm{d}s = \lim_{n\to\infty} \sigma_t^w b_n(x)$$

for any $b \in C^\alpha(\mathbb{R}^d)$ with $\alpha > -1/2H$.

Note that the the local time L_t has been studied a lot in the literature, both in the case of Gaussian processes (Rosen) and in the case of Markov processes, also in relation with the theory of Dirichlet forms and of formulas of Itô-Tanaka type. Here however we are not much interested in the path-wise properties of the local time L itself rather than in its properties as a convolution operator and it turns out that the existing literature is not very useful in this respect.

A first computation to understand the behavior of the local time as a convolution operator is given by estimation of its Fourier transform Y:

$$Y_t(\xi) = \widehat{L_t}(\xi) = \int_0^t e^{i\langle \xi, w_s\rangle}\mathrm{d}s$$

By explicit Gaussian computations or with martingale arguments it is possible to show that for any small $\varepsilon > 0$ there exists $\gamma > 1/2$ such that almost surely in W we have

$$|Y_t(\xi) - Y_s(\xi)| \lesssim C_w \frac{|t-s|^\gamma}{1+|\xi|^{(1/2H)-\varepsilon}} \qquad \xi \in \mathbb{R}^d, t, s \in [0,1]. \tag{2}$$

The basic observation leading to this result is that, by scale and rotational invariance of the fBm we have

$$Y_t(\xi) \sim_{\text{law}} |\xi|^{-1/H} Y_{|\xi|^{1/H}t}(e_1)$$

where $e_1 = (1,0,...,0) \in \mathbb{R}^d$ is the unit vector in the first direction of $\mathbb{R}^d$. Then the above result is a consequence of the fact that the asymptotic variance of $Y_t(e_1)$ for large t scales like $t^{1/2}$ and that $t^{-1/2}Y_t(e_1)$ has uniform gaussian tails when $t \to +\infty$. In particular note that the key estimate (2) is a consequence of a CLT-like result and thus the exponent γ is near $1/2$ irrespective of the time regularity H of the fBM, which can be arbitrarily small.

In the Brownian case the regularity of Y can be exposed using the Itô-Tanaka trick, that is writing an Itô formula for $e^{i\langle \xi, W_t\rangle}$:

$$e^{i\langle \xi, W_t\rangle} = e^{i\langle \xi, W_0\rangle} + i\xi \int_0^t e^{i\langle \xi, W_s\rangle}\mathrm{d}W_s - \frac{|\xi|^2}{2}\int_0^t e^{i\langle \xi, W_t\rangle}\mathrm{d}s$$

so that

$$Y(\xi) = \frac{i2\xi}{|\xi|^2}\int_0^t e^{i\langle \xi, W_s\rangle}\mathrm{d}W_s + 2\frac{e^{i\langle \xi, W_t\rangle} - e^{i\langle \xi, W_0\rangle}}{|\xi|^2}$$

from which it follows very easily that

$$\sup_{\xi,t} \mathbb{E}\big[e^{\lambda|\xi|^2|Y_t(\xi)|^2/t}\big] < +\infty$$

for some small $\lambda > 0$. Using this estimate, the trivial one $|Y_t(\xi)| \leqslant |t|$ and a Kolmogorov like argument you can easily get that almost surely

$$|Y_t(\xi) - Y_s(\xi)| \lesssim C(W)|\xi|^{-1+\varepsilon}|t-s|^{1/2+\varepsilon}$$

for any small $\varepsilon > 0$.

With some care these observations on the Fourier transform of Y can be cast into an explicit estimation of the regularizing properties of σ^w in the Hölder-Besov spaces C^α, in particular we have, almost surely

$$\|\sigma_t^w b - \sigma_s^w b\|_{C^{\alpha+\rho}} \lesssim \|b\|_{C^\alpha}|t-s|^\gamma$$

for any $\rho < 1/2H$ and some $\gamma > 1/2$ which depends only on ρ. This means that the averaging achieve a gain of $1/2H$ derivatives in the Hölder-Besov spaces and the avergain of a distribution can result into a function if the regularity of the distribution is greater that $-1/2H$.

Phenomena related to this stochastic averaging:

a) The regularization obtained by the transport term in kinetic theory: the probability distribution $f_t(x,v)$ in the phase space (x,v) (position and velocity of the particles) is transported according to the first order operator $\partial_t + v\cdot\partial_x$, this means that

$$f_t(x,v) = f_0(x - v\,t, v)$$

then the particle density $f_t(x)$ at time $t>0$ and position $x\in\mathbb{R}^d$ satisfies the bound

$$f_t(x) = \int_{\mathbb{R}^d} f_t(x,v)\mathrm{d}v = \int_{\mathbb{R}^d} f_0(x-v\,t,v)\mathrm{d}v \leqslant \int_{\mathbb{R}^d} \sup_{w\in\mathbb{R}^d} f_0(x-v\,t,w)\mathrm{d}v$$
$$\leqslant \frac{1}{|t|^d}\int_{\mathbb{R}^d} \sup_{w\in\mathbb{R}^d} f_0(x,w)\mathrm{d}x = \frac{1}{|t|^d}\left\| \sup_{w\in\mathbb{R}^d} f_0(\cdot,w)\right\|_{L^1(\mathbb{R}^d)}$$

uniformly in $x\in\mathbb{R}^d$. The stochastic nature of the velocity of the particles and the action of the first-order transport term result in a regularization of the initial condition which from being $L^1_x(\mathbb{R}^d)$ in the space variable at the initial time become $L^\infty_x(\mathbb{R}^d)$ at any other positive time. See the book of Villani for more information about the use of this bound in the analysis of kinetic equations.

b) The work of T. Tao and J. Wright (Tao, Terence, and James Wright. 2003. "L^p Improving Bounds for Averages Along Curves." Journal of the American Mathematical Society 16 (03) (January 28): 605–638.) on integrability gain of averaging over smooth curves

$$\sigma^\gamma_t f(x) = \int_0^t f(x+\gamma_s)\mathrm{d}s$$

and they show that $f\in L^q(\mathbb{R}^d)\to\sigma^\gamma_t f\in L^{q'}(\mathbb{R}^d)$ with $q'>q$.

2.3. The space of controlled paths and definition of the averaged vectorfield.

As a function of time, the averaged vector field $\sigma^w_t b$ is Holder continuous of oder $\gamma > 1/2$. So there is some hope that the solution of the ODE (1) will look like a sum of a fBm w and a Hölder continuous perturbation θ:

$$x_t = x_0 + w_t + \theta_t \tag{3}$$

where $\theta\in C^{1/2}([0,1];\mathbb{R}^d)$ at least. This decomposition will constitute our notion of controlled path in this particular problem. That is we will say that a path $x\in C([0,1],\mathbb{R}^d)$ is controlled by w if $\theta = x - w\in C^{1/2}([0,1],\mathbb{R}^d)$.

To take advantage of this decomposition we try to approximate the averaging over a controlled path x by the averaging over w which we analyzed just now. To do so we fix a partition $\Pi = \{t_0=0,\dots,t_n=t\}$ of $[0,t]$ and approximate the integral over the partition by freezing the parameter of θ at the initial point of each subinterval:

$$\int_0^t b(w_s+\theta_s)\mathrm{d}s \simeq \sum_i \int_{t_i}^{t_{i+1}} b(w_s+\theta_{t_i})\mathrm{d}s = \sum_i \sigma^w_{[t_i,t_{i+1}]} b(\theta_{t_i})$$

In this way we try to define the l.h.s. as the limit of the r.h.s when the size of the partition goes to zero:

$$\int_0^t \sigma^w_{\mathrm{d}s} b(\theta_s) = \lim_{|\Pi|\searrow 0}\sum_i \sigma^w_{[t_i,t_{i+1}]} b(\theta_{t_i}) = \int_0^t b(w_s+\theta_s)\mathrm{d}s$$

The limit exists due to the fact that $\theta \in C^{1/2}$ and that

$$|\sigma^w_{[s,t]} b(x) - \sigma^w_{[s,t]} b(y)| \lesssim |t-s|^\gamma |x-y|$$

since in that case you have

$$|\sigma^w_{[s,s']} b(\theta_{s''}) - \sigma^w_{[s,s']} b(\theta_{s'})| \lesssim |s-s'|^\gamma |\theta_{s''} - \theta_{s'}| \lesssim |s-s'|^\gamma |s''-s'|^{1/2}$$

and this allow to use the classic approach of Young integrals to take the limit over the partition. For example we can show that the maximal inequality for the Riemman sums approximations holds and the pass to the limit. The limit procedure needs Lipshitz regularity in space of $\sigma^w_t b$ which means that it works only for $b \in C^\alpha(\mathbb{R}^d)$ with $\alpha > 1 - 1/2H$ (taking into account the averaging properties of w).

$$\sigma^{w+\theta}_t b(x) = \int_0^t b(x + w_s + \theta_s) \mathrm{d}s = \int_0^t \sigma^w_{\mathrm{d}s} b(x + \theta_s)$$

and $\sigma^{w+\theta} b \in C^{\alpha + 1/2H - 1}$ if $b \in C^\alpha$.

$$x^n_t = x_0 + \int_0^t b_n(x^n_s) \mathrm{d}s + w_t$$

rewrite it as Young equation $x^n = w + \theta^n$

$$\theta^n_t = x_0 + \int_0^t \sigma^w_{\mathrm{d}s} b_n(\theta^n_s)$$

with $\theta^n \in C^{1/2}(\mathbb{R}^d)$. And when $b_n \to b$ we have that $\theta^n \to \theta$ in $C^{1/2}(\mathbb{R}^d)$ with θ unique solution to

$$\theta_t = x_0 + \int_0^t \sigma^w_{\mathrm{d}s} b(\theta_s).$$

Now define

$$x_t = x_0 + w_t + \theta_t$$

and you can show that x (since it is controlled by w and θ) satisfy

$$\int_0^t b_n(x_s) \mathrm{d}s = \int_0^t \sigma^w_{\mathrm{d}s} b_n(\theta_s) \to \int_0^t \sigma^w_{\mathrm{d}s} b(\theta_s) =: \int_0^t b(x_s) \mathrm{d}s$$

and x solve the original equation where the nonlinearity is interpreted in this sense

$$x_t = x_0 + \int_0^t b(x_s) \mathrm{d}s + w_t.$$

3. Regularization by noise in stochastic dispersive equations

(based on joint work with K. Chouk, almost finished)

Nonlinear dispersive equations with random modulation (on $\mathbb{T}$)

$$\mathrm{d}_t \varphi_t = A\varphi_t \circ \mathrm{d}W_t + \mathcal{N}(\varphi_t)\mathrm{d}t$$

where W is a Brownian motion and $\circ \mathrm{d}W$ is Stratonovich integration and :

- Nonlinear Schrödinger equation (NLSE): $A = i\Delta$, $\mathcal{N}(\varphi) = i|\varphi|^2\varphi$, $\varphi \colon [0,1] \times \mathbb{T} \to \mathbb{C}$
- Korteweg–de Vries equation (KdV): $A = \partial^3$ and $\mathcal{N}(\varphi) = \partial(\varphi^2)$, $\varphi \colon [0,1] \times \mathbb{T} \to \mathbb{R}$
- Linear stochastic transport equation: $A = \partial$, $\mathcal{N}(\varphi) = b \cdot \partial\varphi$, $\varphi \colon [0,1] \times \mathbb{T} \to \mathbb{R}$

where the initial conditions φ_0 is taken in $H^\alpha(\mathbb{T})$, the usual Sobolev spaces of periodic function on $\mathbb{T}$ with zero mean and norm given by

$$\|\varphi\|^2_\alpha = \sum_{k \neq 0} |k|^{2\alpha} |\hat\varphi(k)|^2 < +\infty.$$

where $\alpha \geqslant 0$ in the case of NLSE and the transport equation and $\alpha < 0$ (but not too negative) in the KdV case.

The difficulties in these equations lie again in the fact that non-linearity usually do not live in the same space where we would like to set up the equation.

Formally the linear transport equation belongs to this class of problems but the line of attach which we describe below seems to be uneffective.

To understand whether the interaction of the linear part with the non-linear one can give rise to some regularization phenomena we look at the mild formulation

$$\varphi_t = U_t^W \varphi_0 + U_t^W \int_0^t (U_s^W)^{-1} \mathcal{N}(\varphi_s) \mathrm{d}s$$

where $U_t^W = \exp(A W_t) = V_{W_t}$ and V is just the group generated by the linear operator A on H^α:

$$\frac{\mathrm{d}}{\mathrm{d}t} V_t \varphi = A V_t \varphi.$$

In Fourier space we have

$$\widehat{U_t^W \varphi}(\xi) = e^{a(\xi) W_t} \hat{\varphi}(\xi)$$

with $a(\xi)$ the symbol of the operator A. So U^W corresponds to a modulation with a time-varying phase factor.

One advantage of the mild formulation is that of being meaningful for any continuous path W and not only for Brownian motion. More surprisingly the information about the path W which is relevant in studying the mild formulation is again contained in the local time of W or better in its Fourier transform.

As in the ODE problem we can then consider W sampled according to the law of a fBm, for which we have detailed information on the pathwise behaviour of the local time.

A reasonable solution of the above equation will have the form

$$\varphi_t = U_t^W \theta_t \tag{4}$$

with θ some continous path in H^α. If we believe that the non-linearity constiture a perturbation of the linear flow we are led to postulate that θ should have a regular time dependence and we will see that assuming $\theta \in C^{1/2}(H^\alpha)$ will be enough. The factorization (4) will be our controlled structure. Here the controlled structure is multiplicative and this has to be compared with the additive controlled structure of the ODE case $x_t = w_t + \theta_t$. The equation for θ now reads:

$$\theta_t = \theta_0 + \int_0^t (U_s^W)^{-1} \mathcal{N}(U_s^W \theta_s) \mathrm{d}s$$

So proceeding similarly we introduce the averaging of the non-linear term along the conjugations of H^α given by the linear part:

$$X_t^W(\psi) = \int_0^t (U_s^W)^{-1} \mathcal{N}(U_s^W \psi) \mathrm{d}s.$$

In the above situations we can prove that

$$|X_t^W(\psi) - X_s^W(\psi)|_{H^\alpha} \lesssim F(|\psi|_{H^\alpha}) |t-s|^\gamma$$

where $\gamma > 1/2$ for suitable values of α and $F: \mathbb{R}_+ \to \mathbb{R}_+$ is some polynomial function. In particular:

- $\alpha \geqslant 0$ in the case of NLSE.
- $\alpha \geqslant -1/2H$ in the KdV case.

The nonlinear term for a general controlled path can then be represented as a Young integral according to the formula

$$\int_0^t (U_s^W)^{-1} \mathcal{N}(\varphi_s) \mathrm{d}s = \int_0^t (U_s^W)^{-1} \mathcal{N}(U_s^W \theta_s) \mathrm{d}s = \int_0^t X_{\mathrm{d}s}^W(\theta_s)$$

and the original mild equation can be cast in the form of a Young equation for the perturbation θ

$$\theta_t = \theta_0 + \int_0^t X^W_{\mathrm{d}s}(\theta_s).$$

4. Stochastic Burgers equation

In this section I would like to study the stochastic Burgers equation, again using a controlled structure but in the context of stochastic calculus instead of using Young integrals as in the previous two examples. The problem is to give a meaning to solutions to the following stochastic Burgers equation (obtained as the space derivative of the Kardar-Parisi-Zhang equation)

$$u_t = u_0 + \int_0^t \Delta u_s \mathrm{d}s + \int_0^t \partial(u_s^2)\mathrm{d}s + \partial \int_0^t \xi_s \mathrm{d}s \tag{5}$$

for $u\colon [0,T]\times\mathbb{T}\to\mathbb{R}$ ($\mathbb{T}=[-\pi,\pi]$ is the one-dimensional torus) where $\xi\colon [0,1]\times\mathbb{T}\to\mathbb{R}$ is space-time white noise (with zero space average for simplicity). In this situation, for fixed time $t\in[0,1]$ the solution u_t is expected not to be better than white-noise in space. In this situation the definition of the non-linear (in particular the square of u) term poses a problem.

Here we will see how despite of the fact that (u_s^2) is not a well defined function (in space for fixed time), is nonetheless a well define space-time distribution and the time integral of the non-linear term is well defined in a space of Hölder continuous functions of time with some space Sobolev (or Besov) regularity (not enough to be a function but enough to give a well-defined equation).

Consider again the mild formulation of the above equation

$$u_t = P_t u_0 + \int_0^t \partial P_{t-s}(u_s^2)\mathrm{d}s + \int_0^t \partial P_{t-s}\xi_s \mathrm{d}s$$

where we take the initial condition u_0 to be random and distributed like a space-white noise (with zero mean) independent of $(\xi_s)_{s\geqslant 0}$. It is expected that this initial condition give rise to stationary solutions to this equation. In this case we can represent it as

$$u_0 = \int_{-\infty}^0 \partial P_{-s}\xi_s \mathrm{d}s$$

where $(\xi_s)_{s\leqslant 0}$ is another space-time white noise independent of $(\xi_s)_{s\geqslant 0}$. With this representation we have a convenient formulation for the solution to the above equation

$$u_t = X_t + \int_0^t \partial P_{t-s}(u_s^2)\mathrm{d}s$$

where $(X_t)_{t\in\mathbb{R}}$ is a real stationary OU process obtained as

$$X_t = \int_{-\infty}^t \partial P_{t-s}\xi_s \mathrm{d}s$$

and has zero mean and covariance (in Fourier space)

$$\mathbb{E}\big[\hat{X}_t(k)\hat{X}_s(l)\big] = \mathbb{I}_{k+l=0,k\neq 0} e^{-k^2|t-s|}$$

showing for example that for any fixed time $t\in\mathbb{R}$ X_t is a space white noise (of zero space average $\hat{X}_t(0)=0$).

This problem has the same structure of the ODE problem: the stochastic perturbation enters linearly in the equation and the non-linear part is a time integral of a not well defined quantity (due to spatial irregularities). To have an intuition about what is going on we proceed to compute the non-linearity on the OU process, that is we look at the random distribution

$$A = \int_0^t \partial P_{t-s}(X_s^2)\mathrm{d}s$$

in Fourier space it reads

$$\hat{A}(k) = i\,k \int_0^t e^{-k^2(t-s)} \sum_{k_1+k_2=k} \hat{X}_s(k_1)\hat{X}_s(k_2)\mathrm{d}s$$

where the sum is performed over all couples $(k_1, k_2) \in \mathbb{Z}^2$ such that $k_1 + k_2 = k$ and $k_1, k_2 \neq 0$. Note that this quantity has a zero mean ($\mathbb{E}\hat{X}_s(k_1)\hat{X}_s(k_2) = \mathbb{I}_{k_1+k_2=0, k_1\neq 0}$)

$$\mathbb{E}\big[\,\hat{A}(k)\big] = i\,k \int_0^t e^{-k^2(t-s)} \sum_{k_1+k_2=k} \mathbb{I}_{k_1+k_2=0}\mathrm{d}s = 0$$

due to the total derivative which give a factor of k in front of the expression. This and other computation we will do below should be understood to be performed with a regularization of the non-linearity, for example we could take

$$A_N = \int_0^t \partial P_{t-s}((\Pi_N X_s)^2)\mathrm{d}s$$

where Π_N is the projection on Fourier modes of size less than N. In this exposition we will avoid to do it explicitly. As for the variance of $\hat{A}(k)$ a simple computation gives

$$\begin{aligned}
\mathbb{E}\big[\,\hat{A}(k)\hat{A}(-k)\big] &= k^2 \int_0^t \mathrm{d}s \int_0^t \mathrm{d}s'\, e^{-k^2(t-s)-k^2(t-s')} \sum_{k_1+k_2=k} e^{-(k_1^2+k_2^2)|s-s'|} \\
&\leqslant k^2 \int_0^t \mathrm{d}s\, e^{-2k^2(t-s)} \int_0^\infty \mathrm{d}\sigma \sum_{k_1+k_2=k} e^{-(k_1^2+k_2^2)\sigma} \\
&\lesssim \big(1-e^{-k^2t}\big) \sum_{k_1+k_2=k} \frac{1}{k_1^2+k_2^2} \\
&\lesssim \big(1-e^{-k^2t}\big) \int_{\mathbb{R}} \frac{\mathrm{d}x}{x^2+(k-x)^2} \\
&\lesssim \frac{1-e^{-k^2t}}{|k|}
\end{aligned}$$

which shows that the variance is finite so that A is a well defined random distribution of certain regularity. Actually this estimate shows that $A \in H^{-\varepsilon}(\mathbb{T})$ almost surely for any $\varepsilon > 0$:

$$\mathbb{E}[\|A\|^2_{H^{-\varepsilon}}] = \mathbb{E}\left[\sum_{k\neq 0} |k|^{-2\varepsilon}\big|\hat{A}(k)\big|^2\right] \lesssim \sum_{k\neq 0} |k|^{-2\varepsilon-1}\big(1-e^{-k^2t}\big) < +\infty.$$

So the point is that the quantity X_s^2 should be considered a space-time distribution to be tested against functions of zero mean and, in particular, $|X_s^2|$ does not make apriori make sense.

We expect similar properties from the square of the full solution $(u_s)^2$. This means for example that we aim to get uniform estimates on the distribution $(u_s^N)^2$ obtained by Galerkin approximations of the equation (or by approximating the noise – to handle the invariant measure it is easier to approximate the non-linearity and retain the full OU process as driving noise).

However here we cannot just resort to explicit computations and we will rely on the Itô–Tanaka trick as explained before in the analysis of the local time of the Brownian motion.

To this purpose we consider the following class of controlled distributions: we say that u is controlled if it is a continuous function of time with values in space distributions ($u \in C([0,1], \mathcal{S}'(\mathbb{T}))$) with the following properties:

a) Martingale decomposition:

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

where $\mathcal{A}_t(\varphi)$ is a process of zero quadratic variation.

b) For any $0 \leqslant t \leqslant 1$, u_t is a white noise on $\mathbb{T}$.

c) The reversed process $\tilde{u}_t = u_{T-t}$ on $t \in [0,T]$ has the same decomposition with a drift $\tilde{\mathcal{A}}_t$ which satisfy $\tilde{\mathcal{A}}_t = -\mathcal{A}_{T-t}$.

These property express the fact that u behaves essentially like an OU process. For any process with these properties we are able to have an Itô formula for smooth cylindric functions (i.e. smooth functions depending only on a finite number of Fourier modes), which reads

$$F(u_T) = F(u_0) + \int_0^T L_0 F(u_s)\mathrm{d}s + \int_0^T D\,F(u_s)\mathrm{d}\mathcal{A}_s + M_T^+(F)$$

where L_0 is the generator of the OU process and a similar formula for $\tilde{u}$:

$$F(\tilde{u}_T) = F(\tilde{u}_0) + \int_0^T L_0 F(\tilde{u}_s)\mathrm{d}s + \int_0^T D\,F(u_s)\mathrm{d}\tilde{\mathcal{A}}_s + M_T^-(F)$$

summing these two formulas we get rid of the drift term and obtain that

$$2\int_0^T L_0 F(u_s)\mathrm{d}s = M_T^+(F) + M_T^-(F)$$

The quadratic variation of the martingales $M^\pm$ is controlled by

$$\langle M^{\pm}(F)\rangle_T \propto \int_0^T \mathcal{E}(F)(u_s)\mathrm{d}s$$

where

$$\mathcal{E}(F)(u) = \frac{1}{2}\sum_k |k|^2 |D_k F(u)|$$

and $D_k = \delta/\delta\hat{u}(k)$ is the derivative with respect to the k-th Fourier mode of u.

In particular it is not difficult to find a (time-dependent) function H_s such that

$$L_0 H_s(u_s) = \langle \varphi, \partial P_{T-s} u_s\rangle^2$$

and for which $\mathcal{E}(H_s)(u_s)$ is finite if u_s has the law of the white noise on $\mathbb{T}$. In this way we have an explicit representation of the drift term for any controlled path and this is given by

$$\int_0^T \langle \varphi, \partial P_{T-s} u_s^2\rangle \mathrm{d}s = M^+(H) + M_T^-(H)$$

whose moments are controlled by the quantity

$$\int_0^T \mathcal{E}(H_s)(u_s)\mathrm{d}s$$

which is well defined and finite. What we have achieved is a definition of the non-linear term

$$\int_0^T \langle \varphi, \partial P_{T-s} u_s^2\rangle \mathrm{d}s$$

for the class of controlled paths. Then the equation (5)

$$u_t = u_0 + \int_0^t \Delta u_s \mathrm{d}s + \int_0^t \partial(u_s^2)\mathrm{d}s + \partial\int_0^t \xi_s \mathrm{d}s$$

make sense in this space of controlled paths. In other terms we can say that a (stationary) solution to the Burgers equation is a controlled path for which the drift $\mathcal{A}$ satisfy

$$\mathcal{A}_t = \lim_{\varepsilon\to 0}\int_0^t \partial((u_s)_\varepsilon^2)\mathrm{d}s$$

where $(u_s)_\varepsilon$ is some regularization of u, for example by convolution with a smooth kernel $\rho_\varepsilon \to \delta$ as $\varepsilon \to 0$.

5. Controlled distributions

Based on joint work with P. Imkeller and N. Perkowski.

Here we work in the spaces $C^\alpha(\mathbb{R}^d)$ or sometimes $C^\alpha_{\mathrm{loc}}(\mathbb{R}^d)$.

5.1. Rough Differential Equations.

Take as $x \in C(\mathbb{R}_+, \mathbb{R})$ a Brownian motion. Then $D_t x$ is well defined random distribution (it is white noise). We would like to solve an equation of the form

$$D_t y(t) = V(y(t)) D_t x(t)$$

in the space of distributions $\mathcal{S}'(\mathbb{R})$, subject for example to the condition that y is a function and $y(0) = y_0$ is fixed.

The first observation is that we cannot hope that y will be smoother than x. This means that we have to consider $y \in C^\gamma_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{R})$ with some $\gamma < 1/2$. We have also $\mathrm{D}_t x \in C^{\gamma-1}$. If V is smooth and bounded it is not difficult to show that $y \in C^\gamma_{\mathrm{loc}} \Rightarrow V(y) \in C^\gamma_{\mathrm{loc}}$. Then we are trying to multiply a distribution in $C^{\gamma-1}$ with a function in C^γ.

A basic result is that if $f \in C^\alpha$ and $g \in C^\beta$ then the product fg is a well defined distribution if $\alpha + \beta > 0$. In our case $\alpha = \gamma$ and $\beta = \gamma - 1$ so $\alpha + \beta = \gamma + \gamma - 1 = 2\gamma - 1 < 0$ and the product is not guaranteed to be well defined.

5.2. Paraproduct decomposition.

To undestand better the origin of the problem we decompose the product fg according to the LP blocks and introduce the notion of paraproduct which will be a useful tool below.

Take $f \in C^\alpha$ and $g \in C^\beta$ and decompose the product according to the following formula:

$$fg = \sum_{i,j \geqslant -1} \Delta_i f \, \Delta_j g = \pi_<(f,g) + \pi_\circ(f,g) + \pi_>(f,g)$$

where

$$\pi_<(f,g) = \pi_>(g,f) = \sum_{i,j : i < j-1} \Delta_i f \, \Delta_j g$$

$$\pi_\circ(f,g) = \sum_{i,j ; |i-j| \leqslant 1} \Delta_i f \, \Delta_j g$$

The paraproduct $\pi_<(f,g)$ is always well defined. If $\alpha > 0$

$$\|\pi_<(f,g)\|_\beta \lesssim \|f\|_\alpha \|g\|_\beta \tag{6}$$

while if $\alpha < 0$ we have

$$\|\pi_<(f,g)\|_{\alpha+\beta} \lesssim \|f\|_\alpha \|g\|_\beta \tag{7}$$

The diagonal term $\pi_\circ(f,g)$ is well defined only if $\alpha + \beta > 0$ and in this case it satisfy

$$\|\pi_\circ(f,g)\|_{\alpha+\beta} \lesssim \|f\|_\alpha \|g\|_\beta \tag{8}$$

The proof of these bounds are easy. A key observation is that if $i < j - 1$ then by the definition of the LP blocks we have that

$$\Delta_q(\Delta_i f \, \Delta_j g) = 0$$

if $|q-j|>1$. This is due to the fact that the typical oscillations in $\Delta_i f$ there are of lenght $\sim 2^{-i}$ while in $\Delta_j f$ are of lenght 2^{-j} so, since $2\,2^{-i}<2^{-j}$ the product $\Delta_i f \Delta_j g$ oscillate on scales which are still of order 2^{j-1} at least and not smaller. And of course the oscillation of the product cannot happen on scales much smaller than 2^{-j}. Projecting on LP blocks at scale 2^{-q} then give zero if q is far away from j. The quantification of this distance depends on the specific form of the LP blocks but it is unessential in the estimates.

For example let us consider the case of $\pi_<(f,g)$ with $\alpha<0$:

$$\Delta_q \pi_<(f,g)=\sum_{i,j:} \mathbb{I}_{q\sim j, i<j-1}\Delta_q(\Delta_i f\, \Delta_j g)$$

at this point

$$\|\Delta_q \pi_<(f,g)\|_{L^\infty}\leqslant \sum_{i,j:} \mathbb{I}_{q\sim j, i<j-1}\|\Delta_q(\Delta_i f\, \Delta_j g)\|_{L^\infty}$$

$$\lesssim \sum_{i,j:} \mathbb{I}_{q\sim j, i<j-1}\|\Delta_i f\, \Delta_j g\|_{L^\infty}\lesssim \|f\|_\alpha \|g\|_\beta \sum_{i,j:} \mathbb{I}_{q\sim j, i<j-1} 2^{-i\alpha-j\beta}$$

$$\lesssim \|f\|_\alpha \|g\|_\beta 2^{-q\beta}\sum_{i=-1}^{q} 2^{-i\alpha}\lesssim \|f\|_\alpha \|g\|_\beta 2^{-q(\alpha+\beta)}$$

which shows that $\pi_<(f,g)$ is well defined and that $\pi_<(f,g)\in C^{\alpha+\beta}$ as claimed.

5.3. Definition of the product.

Assume that V is smooth so that $V(y)\in C^\gamma$ if $y\in C^\gamma$ then

$$V(y_t)D_t x=\underbrace{\pi_<(V(y_t),D_t x)}_{\gamma-1}+\underbrace{\pi_\circ(V(y_t),D_t x)}_{2\gamma-1}+\underbrace{\pi_>(V(y_t),D_t x)}_{2\gamma-1}$$

Then, if I is the integral map such that

$$\mathrm{D}I(f)=f,\qquad I(f)(0)=0$$

we have

$$I\colon C^\alpha\to C^{\alpha+1}_{\mathrm{loc}}$$

and we can set

$$y=y_0+I(\pi_<(V(y),D_t x))+I(\pi_\circ(V(y_t),D_t x))+I(\pi_>(V(y_t),D_t x)).$$

One advantage of the paraproduct $\pi_<$ is that many operations behaves (modulo smooth corrections) like if the low frequency term is constant.

For example we have

$$I(\pi_<(V(y),Dx))\simeq \underbrace{\pi_<(V(y),x)}_{\gamma}+C^{2\gamma}$$

which shows explicitly that y behaves like x.

This computation motivates the following definition.

We say that y is controlled by x if

$$y=\pi_<(y',x)+y^\sharp$$

where $y'\in C^\gamma$ and $y^\sharp\in C^{2\gamma}$.

This defines the controlled structure.

The controlled structure is stable under non-linear map. It can be proved that $V(y)$ is also controlled by x with the following structure

$$V(y) = \pi_<\big(V^{(1)}(y)y', x\big) + (V(y))^\sharp$$

since you have the paralinearization theorem (Bony, Meyer)

$$V(y) = \pi_<\big(V^{(1)}(y), y\big) + C^{2\gamma}$$

where $V^{(1)}$ is the derivative of V and V at least C^2.

The assumption that z is controlled by x:

$$z = \pi_<(z', x) + z^\sharp$$

can be used to decompose the diagonal part of the product $z\mathrm{D}x$:

$$\pi_\circ(z, \mathrm{D}x) = \pi_\circ(\pi_<(z', x), \mathrm{D}x) + \pi_\circ(z^\sharp, \mathrm{D}x)$$

Now $\pi_\circ(z^\sharp, \mathrm{D}x)$ is well defined since $2\gamma + \gamma - 1 > 0$ if $\gamma > 1/3$.

Commutator lemma

$$\|\pi_\circ(\pi_<(f, g), h) - f\pi_\circ(g, h)\|_{3\gamma-1} \lesssim \|f\|_\gamma \|g\|_\gamma \|h\|_{\gamma-1}$$

so

$$\pi_\circ(\pi_<(z', x), \mathrm{D}x) = z'\,\pi_\circ(x, \mathrm{D}x) + C^{3\gamma-1}$$

The quantity $\pi_\circ(x, \mathrm{D}x)$ plays the role of the Levy area in the classic approach.

Suppose that $\pi_\circ(x, \mathrm{D}x)$ exists and belongs to $C^{2\gamma-1}$ (its natural regularity). We can define the product in the r.h.s. of our RDE as

$$V(y)\mathrm{D}x = \pi_<(V(y), \mathrm{D}x) + \underbrace{V^{(1)}(y)y'\pi_\circ(x, \mathrm{D}x)}_{2\gamma-1} + \underbrace{\pi_>(V(y_t), \mathrm{D}x)}_{2\gamma-1} + C^{3\gamma-1}$$

where the remainder term is an explicit function of $x, y', y^\sharp$.

And then after integration:

$$z = I(V(y)\mathrm{D}x) = \pi_<(V(y), x) + C^{2\gamma} + C^{3\gamma}.$$

we still retain the controlled structure with $z' = V(y)$, $z^\sharp \in C^{2\gamma}$.

In the usual controlled setting of rough paths you will have

$$\delta y = y'\delta x + y^\sharp$$

and if $z = \int V(y)\mathrm{d}x$ then

$$\delta z = V(y)\delta x + V^{(1)}(y)y'\mathbb{X}^2 + z^\flat$$

with $z^\flat$ of order 3γ.

6. A parabolic Anderson model

Consider $\mathbb{T}^2$ and the the following equation, $V\colon \mathbb{R} \to \mathbb{R}$

$$\partial_t u = \Delta u + V(u)\xi$$

ξ is a space white noise constant in time.

$$\mathbb{E}(\xi(x)\,\xi(x')) = \delta(x - x')$$

with $x, x' \in \mathbb{T}^2$ or otherwise

$$\xi(x) = \sum_{k\in\mathbb{Z}^2, k\neq 0} e^{i\langle k,x\rangle} g_k$$

with $(g_k)_{k\in\mathbb{Z}^2}$ a family of complex standard Gaussians such that $g_k = g^*_{-k}$ and g_k and $g_{k'}$ are independent if $|k| \neq |k'|$

$$\mathbb{E}[g_k g_l] = \delta_{k+l=0}.$$

Regularity of $\xi \in C^{-d/2-}$.

Analysis of the equation. Mild form

$$u_t = P_t u_0 + \int_0^t P_{t-s}(V(u_s)\xi)\mathrm{d}s$$

Even if u is smooth and so $V(u)$ we have $V(u)\xi \in C^{-1-}$ and

$$\int_0^t P_{t-s}(V(u_s)\xi)\mathrm{d}s \in C^{1-}$$

since the heat semigroup will allow you to gain almost two derivatives. The natural regularity of u is $1-$.

$$\|e^{\Delta t}\Delta_i f\|_{L^\infty} \simeq e^{-2^{2i}t}\|\Delta_i f\|_{L^\infty}$$

$$\left\|\Delta_i \int_0^t P_{t-s}(f_s)\mathrm{d}s\right\|_{L^\infty} \leqslant \int_0^t \|P_{t-s}(\Delta_i f_s)\|_{L^\infty}\mathrm{d}s \lesssim \int_0^t e^{-2^{2i}(t-s)}\mathrm{d}s \|\Delta_i f_s\|_{L^\infty}$$

$$\lesssim (1-e^{-2^{2i}t})2^{-2i} \sup_{0\leqslant s\leqslant t}\|\Delta_i f_s\|_{L^\infty}.$$

So now the problem is that $V(u) \in C^{1-}$ and $V(u)\xi$ is not well defined since the sum of the regularities fails (just for ε) to be positive.

Let $\xi \in C^{\gamma-2}$ (recall $\gamma = 1-\varepsilon$) and $u \in C^\gamma$ and let
What is the controlled structure.

$$V(u)\xi = \underbrace{\pi_<(V(u),\xi)}_{\gamma-2} + \underbrace{\pi_\circ(V(u),\xi)}_{2\gamma-2<0} + \underbrace{\pi_>(V(u),\xi)}_{2\gamma-2}$$

$$\int_0^t P_{t-s}(\pi_<(V(u_s),\xi))\mathrm{d}s \simeq \int_0^t \pi_<(V(u_s), P_{t-s}\xi)\mathrm{d}s + \text{corrections}$$

so let us try to fix the following controlled structure.

We say that u is controlled if

$$u_t = \int_0^t \pi_<(u'_s, P_{t-s}\xi)\mathrm{d}s + u^\sharp_t$$

where $u^\sharp \in C^{2\gamma}$ and $u' \in C^\gamma$.

Paralinearization (Taylor expansion)

$$V(u_t) = \pi_<(V'(u_t), u_t) + C^{2\gamma}$$

The difficult term is

$$\pi_\circ(V(u_s),\xi) = \pi_\circ(\pi_<(V'(u_t),u_t),\xi) + \pi_\circ(C^{2\gamma},\xi)$$

use the commutator lemma

$$= V'(u_t)\pi_\circ(u_t,\xi) + \underbrace{R(V'(u_t),u_t,\xi)}_{\text{from the commutator}} + \pi_\circ(C^{2\gamma},\xi)$$

and the controlled structure to write

$$\pi_\circ(u_t,\xi)=\pi_\circ\Big(\int_0^t \pi_<(u'_s,P_{t-s}\xi)\mathrm{d}s,\xi\Big)+\pi_\circ(u^\sharp,\xi)$$
$$=\int_0^t \pi_\circ(\pi_<(u'_s,P_{t-s}\xi),\xi)\mathrm{d}s+\pi_\circ(u^\sharp,\xi)$$

again commutator lemma

$$=\int_0^t u'_s\pi_\circ(P_{t-s}\xi,\xi)\mathrm{d}s+\int_0^t R(u'_s,P_{t-s}\xi,\xi)\mathrm{d}s+\pi_\circ(u^\sharp,\xi)$$
$$=u'_t\int_0^t \pi_\circ(P_{t-s}\xi,\xi)\mathrm{d}s+\int_0^t (u'_t-u'_s)\pi_\circ(P_{t-s}\xi,\xi)\mathrm{d}s+\int_0^t R(u'_s,P_{t-s}\xi,\xi)\mathrm{d}s+\pi_\circ(u^\sharp,\xi)$$

and the problem is now concentrated in the definition of

$$G_t=\pi_\circ(P_t\xi,\xi)=\sum_{|i-j|\leqslant 1} P_t(\Delta_i\xi)\Delta_j\xi$$

$$\Delta_j\xi(x)=\sum_{k\in\mathbb{Z}^2,k\neq 0}\varphi(2^{-j}k)e^{\iota\langle k,x\rangle}g_k$$

$$P_t(\Delta_i\xi)(x)=\sum_{k\in\mathbb{Z}^2,k\neq 0}e^{-k^2t}\varphi(2^{-i}k)e^{\iota\langle k,x\rangle}g_k$$

The mean

$$g_t=\mathbb{E}[G_t(x)]\simeq\sum_{|i-j|\leqslant 1}\sum_{k\in\mathbb{Z}^2,k\neq 0}\varphi(2^{-j}k)\varphi(2^{-i}k)e^{-k^2t}\simeq\sum_{k\in\mathbb{Z}^2,k\neq 0}e^{-k^2t}\simeq\frac{1}{t}$$

so

$$\mathbb{E}\int_0^t \pi_\circ(P_{t-s}\xi,\xi)\mathrm{d}s\simeq\int_0^t\frac{1}{t-s}\mathrm{d}s=+\infty$$

we define

$$\Xi_t=\int_0^t[\pi_\circ(P_{t-s}\xi,\xi)-\mathbb{E}(\pi_\circ(P_{t-s}\xi,\xi))]\mathrm{d}s$$

which is equivalent to take a Wick product to define $\Delta_iP_{t-s}\xi\,\Delta_j\xi$. The only thing that has some hope to be defined is the modified product

$$V(u_t)\diamondsuit\xi=V(u_t)\xi-V'(u_t)u'_t\int_0^t g_{t-s}\mathrm{d}s$$
$$=\pi_<(V(u_t),\xi)+V'(u_t)u'_t\,\Xi_t+\text{smoother terms}$$

Now it is not difficult to show that

$$\mathbb{E}\|\Xi_t\|_{-\varepsilon}<+\infty.$$

and letting

$$z=\int_0^t P_{t-s}(V(u_s)\diamondsuit\xi)\mathrm{d}s=\int_0^t\pi_<(V(u_t),P_{t-s}\xi)\mathrm{d}s+\text{smoother terms}$$

then z is again a controlled structure with $z'=V(u_t)$. Thus the operation preserves the controlled structure.