Stochastic PDE

Fraydoun Rezakhanlou
Department of Mathematics, UC Berkeley

January 20, 2026

1 Introduction

Stochastic differential equations are ubiquitous in many disciplines of science. For an example, imagine that the velocity of a particle suspended in a fluid at position x and time t is $u(x,t)$. Then the ordinary differential equation (ODE in short) $\dot{x} = u(x,t)$ can be used to determine the position $x(t)$ of the particle at time t. To take into account the thermal fluctuations of fluid molecules, we may use a *stochastic differential equation* (in short SDE) for the dynamics of the particle, namely

$$\dot{x}(t) = u(x(t),t) + \sigma(x(t),t)\eta(t), \tag{1.1}$$

where η represents the *white noise*, and the matrix $\sigma(x,t)$ measures the size of fluctuations at (x,t). Now imagine that the fluid is incompressible that is subject to an external force f. Then a classical model for the evolution of the velocity field u is a partial differential equation (in short PDE) known as Navier-Stokes equation, namely

$$u_t + (u \cdot \nabla)u = \nu \Delta u - \nabla P + f, \qquad \nabla \cdot u = 0, \tag{1.2}$$

where P represents the pressure, and for simplicity we have assumed that $\sigma = \sqrt{2\nu} I$ is a constant multiple of the identity matrix. In many models of interest, we may take f to be random (for example, f is white in time and correlated in space), and the resulting equation is an example of a *stochastic partial differential equation* (in short SPDE).

A solution (1.1) is an example of a *diffusion process*. The white noise η on the right-hand side is a Gaussian process such that $\mathbb{E}\eta(t) = 0$, and $\mathbb{E}\eta(s)\eta(t) = \delta_0(t-s)$. Since η cannot be represented as a function, some care is needed for the precise meaning of (1.1). As it is well-known, we may realize η as the derivative of a standard Brownian motion B. In other words, when $u = 0$ and $\sigma = I$, then $x(t) = x(0) + B(t)$ is the desired solution to (1.1). It is known that B is nowhere differentiable and Hölder continuous. In fact if $\mathcal{C}^\alpha$ denotes the set of Hölder continuous functions of Hölder exponent α, then $B \in \mathcal{C}^{\frac{1}{2}-}$, where

$$\mathcal{C}^{\alpha-} = \bigcap_{\beta<\alpha} \mathcal{C}^\beta. \tag{1.3}$$

Because of this, we may expect $x(\cdot) \in \mathcal{C}^{\frac{1}{2}-}$ for almost sure realization of a diffusion. We may attempt to make sense of (1.1) by integrating both sides to write

$$x(t) = x(0) + \int_0^t u(x(s),s)\ ds + \int_0^t \sigma(x(s),s)\eta(s)\ ds. \tag{1.4}$$

For this formulation we need to make sense of

$$\int_0^t \sigma(x(s),s)\eta(s)\ ds = \int_0^t \sigma(x(s),s)\ dB(s).$$

Writing $f(s) = \sigma(x(s), s)$ and $g = B$, we may wonder how one can make sense of the integral $\int_0^t f\ dg$? This is a very classical question with a rich history that we now review (as a warm up, we may even try a function f that is of the form $f(t) = F(B(t))$ for a Lipschitz function F):

1. The classical Riemmann-Steiltjes integral gives meaning to the expression $\int f\ dg$ by the following approximation procedure:

$$h(t) = \int_0^t f\ dg = \lim_{n\to\infty} \sum_{i=1}^{2^n} f(s_{i-1})(g(t_i) - g(t_{i-1})), \tag{1.5}$$

where $t_i = it2^{-n}$ and s_i is chosen from the interval $[t_{i-1}, t_i]$. The limit exists provided that f is continuous and g is a function of bounded variation. By integration by parts, we can also make sense of $\int g\ df$.

2. The Lebesgue Theory allows us to relax the continuity requirement of the integrand to the mere integrability. More importantly, we may interpret the finiteness of the total variation as the weak differentiability of g and that as a Schwartz distribution g' is a measure. After all the language of distributions provides us with a mean of measuring the roughness of a nondifferentiable function. For example a function is of bounded variation if and only if the distribution g' is a finite measure. What (1.5) gives us is a recipe for finding a function h of bounded variation such that the *measure* h' is absolutely continuous with respect to the measure g' and its Radon-Nikodym derivative with respect to g' is the function f. In summary, if the distribution g' is a measure we can afford to multiply it by a continuous (even g'-integrable) function.

3. We learned from **1-2** that for a pair of continuous functions (f, g), the product fg' is well defined if at least one of the functions in the pair is of finite variation. Young [Y] observed that the approximation procedure used in (1.5) still works even when f and g share the burden of nondifferentiability among themselves. More precisely if the total p-variation of f and total q-variation of g are finite, and $1/p + 1/q > 1$, then the limit in (1.5) exists. In particular if f is Hölder continuous of exponent α (in short $f \in \mathcal{C}^\alpha$) and $g \in \mathcal{C}^\beta$, then the limit in (1.5) exists provided that $\alpha + \beta > 1$. Moreover, Young established an important bound for h that uniquely characterizes it:

$$|h(t) - h(s) - f(s)(g(t) - g(s))| \lesssim [f]_\alpha [g]_\beta |t - s|^{\alpha+\beta}. \tag{1.6}$$

Here

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

Modulo a constant, and whenever $\alpha + \beta > 1$, there is a unique function h for which (1.6) is true for every (s, t). From a modern perspective, what Young's theorem asserts is that we can afford to multiply the derivative g' of a function $g \in \mathcal{C}^\beta$ by a function f provided that $f \in \mathcal{C}^\alpha$ for some $\alpha > 1-\beta$. Moreover the bilinear operator $\mathcal{A} : \mathcal{C}^1 \times \mathcal{C}^1 \to \mathcal{C}^0$, defined by $\mathcal{A}(f,g) = fg'$ has a *continuous extension* $\mathcal{A} : \mathcal{C}^\alpha \times \mathcal{C}^{\beta-1} \to \mathcal{C}^{\beta-1}$ provided that $\alpha+\beta-1 > 0$. Here $\mathcal{C}^{\beta-1}$ is understood as the space of Schwartz distributions that can be represented as g' for some $g \in \mathcal{C}^\beta$.

4. When $\alpha + \beta \le 1$ the approximation scheme in (1.5) fails in two ways; the limit may not exist, and if it does exist, the limit may depend on the choice of $s_i \in [t_{i-1}, t_i]$. This is a well-known phenomena in stochastic calculus: To make sense of ODEs that are perturbed by white noise, we need to make sense of integrals of the form $\int_0^t F(B) \cdot dB$, where B is a standard multidimensional Brownian motion and F is a $\mathcal{C}^1$ vector field. Writing $g = B$ and $f = F(B)$, we realize that the Young's theorem is not applicable because $f, g \in \mathcal{C}^{\frac{1}{2}-}$. However, Itô managed to define $h(t) = \int_0^t F(B) \cdot dB$ by showing that the limit in (1.5) exists with probability 1 provided that we choose $s_i = t_{i-1}$. Probabilistically speaking, this is a preferred choice because the outcome h is a *martingale*. Though the other choice $s_i = t_i$ would lead to different notion of integrals known as backward, which is different from Itô's integral. The average of the Itô and the backward integrals is known as the Stratonovich integral.

5. There is a distinct difference between Young's integral and Itô's; whereas the former $\mathcal{Y}(f,g) := h$ is continuous in (f,g), the latter $\mathcal{I}(B) = \int_0^1 F(B) \cdot dB$ is not continuous with respect to its input, namely the Brownian trajectory. In other words, if we first smoothize B, take the integral and pass to the limit, then the limit may not exist or depend on our scheme of smoothization. This unsatisfactory feature of stochastic integral led Lyons to invent *rough path theory* [FH]. What Lyons discovered was that in order to determine continuously $\mathcal{I}$, the input B alone is not enough. In addition, we also need to decide what version of the integrals

$$\mathbb{B}(s,t) := \int_s^t (B(\theta) - B(s)) \otimes dB(\theta) := \left[\int_s^t (B_i(\theta) - B_i(s))\ dB_j(\theta) \right]_{i,j=1}^d,$$

we want to use. In other words, we may define $\int F(B) \cdot dB$ or rather

$$\mathcal{L}(B, \mathbb{B}) = \int F(B) \cdot d(B, \mathbb{B}),$$

as an operator that has both B and $\mathbb{B}$ as inputs and varies continuously with respect to its inputs. This also allows us to define $\mathcal{L}(B, \mathbb{B})$ for any $B \in \mathcal{C}^\alpha$ provided that $\alpha > 1/3$. To go below $1/3$, additional inputs are needed. Moreover, if we write $\Gamma(f,g)$ for the set of functions h, for which $h(0) = 0$ and something like (1.6) is true, namely

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

then a theorem of Lyons and Victoir [LV] guarantees that $\Gamma(f,g) \neq \emptyset$. Though $\Gamma(f,g)$ is singleton if and only if $\alpha + \beta > 1$. □

The study of PDEs perturbed by white noise calls for a generalization of the rough path theory to higher dimensions. This was achieved by Hairer after initiating the theory of regularity structures. To explain this theory, first observe that (1.6) may be understood as a way of characterizing the function h in terms its Taylor-like approximation $h(t) \approx h(s) + f(s)(g(t) - g(s))$. Though this approximation is based on the possibly rough function g, as opposed to the standard calculus in which $g(t) = t$. The approximation scheme of (1.5) may be regarded as a way to patch up all the local information we have in order to recover the function itself. In higher dimension the *Reconstruction Theorem* of Hairer guarantees that there is always at least one distribution that is compatible with the local approximation we specify, provided that such local data satisfies Hölder type regularity as we vary the base point. The Reconstruction Theorem gives a unifying treatment for the work of Young and Lyons. □

The theory of regularity structure of Martin Hairer [H] has successfully been employed for several important examples of nonlinear (singular) SPDEs. Alternatively, Gubinelli, Imkeller, and Perkowski [GIP15] have initiated the theory of *paracontrolled distributions* to treat such SPDEs; their approach relies on the notion of paraproducts of Bony and Littlewood-Paley decomposition. In these notes, we will introduce and study a few important classes of SPDEs. In these notes however, we will follow Hairer's approach to study some important examples of *subcritical SPDEs*. Here is a list of SPDEs that will be discussed in these notes:

(i) Stochastic Heat Equation (SHE). Many examples considered in these notes are nonlinear perturbations of Stochastic Heat Equation (SHE). It is a linear SPDE that is driven by a space-time white noise ξ:

$$u_t = \Delta u + \xi. \tag{1.7}$$

Again ξ is a Gaussian process with δ-correlation:

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

In the parabolic setting, it is more convenient to scale time and space differently. For example, a natural Hölder norm would look like

$$[[f]]_\alpha = \sup_{s \neq t} \sup_{x \neq y} \frac{|f(x,s) - f(y,t)|}{|x-y|^\alpha + |s-t|^{\alpha/2}},$$

for $\alpha > 0$. We also write $\widehat{\mathcal{C}}^\alpha$ for the set of continuous functions f with $[[f]]_\alpha < \infty$. As before, we may define the Hölder space $\widehat{\mathcal{C}}^\alpha$ for negative α (the precise definition will be given later). It is well known that in fact $\xi \in \widehat{\mathcal{C}}^{(-\frac{d+2}{2})-}$. Standard parabolic estimates suggest a gain of two derivatives. This means that a solution u of (1.7) belongs to $\widehat{\mathcal{C}}^{\frac{2-d}{2}-}$. Hence u is a function only when $d = 1$. In this case u is Hölder $\frac{1}{2}-$ in x and Hölder $\frac{1}{4}-$ in t. To explain this heuristically, let us study the scaling behavior of SHE.

Given a Schwartz (possibly random) distribution $w(x,t)$, define

$$S_\lambda^\gamma w(x,t) := \lambda^{-\gamma} w\left(\lambda x, \lambda^2 t\right).$$

Observe that $S_\lambda^{-d-2}\,\delta_0 = \delta_0$. Because of this, and the definition of the white noise, we deduce $S_\lambda^{-(d+2)/2}\xi \overset{D}{=} \xi$. Now if u is a solution to SHE, and $\hat{u} = S_\lambda^{1-\frac{d}{2}} u$, then

$$\hat{u} = \Delta \hat{u} + \hat{\xi},$$

where $\hat{\xi} = S_\lambda^{-(d+2)/2}\xi \overset{D}{=} \xi$. This means that SHE is invariant under the action of $S_\lambda^{1-\frac{d}{2}}$.

(ii) SHE with multiplicative noise. This SPDE is a nonlinear perturbation of SHE in the form

$$Z_t = \Delta Z + \sigma(Z)\xi. \tag{1.8}$$

When $\sigma(Z) = Z$, then we may interpret ξ as a random potential for a Brownian particle. This case is also related to stochastic growth models as we will see later. When $\sigma(Z) = \sqrt{Z}$, the equation (1.8) is related to the so-called *super-Brownian motion* and Le Gall's *Brownian snake*. To explain the former, imagine that a large number of particles evolve in $\mathbb{R}^d$ as independent Brownian motions. At random exponential times each particle dies and is replaced by a random number of identical particles, which behave as all other particles in the system. When the average number of descendants is 1, then we are in a *critical regime*. In this regime, when $d = 1$ the population density satisfies (1.8) for $\sigma(Z) = \sqrt{Z}$.

When $d = 1$, one possible strategy for making sense of (1.8) is by developing an Itô type calculus in infinite dimensional setting. More precisely, we may define a *cylindrical Brownian motion* $W(x,t)$ such that $W_t = \xi$ and write

$$Z(x,t) = \int p(x-y,t) Z(y,0)\ dy + \int_0^t \int \sigma(Z(y,s))\ p(x-y,t-s)\ dy\ W(dy,ds),$$

where p is the fundamental solution of the heat equation.

In higher dimension, it is not clear how to make sense of $\sigma(Z)\xi$ because both Z and ξ are distribution when $d \geq 2$. In fact the SPDE (1.8) is *critical* when $d = 2$, and *supercritical* when $d \geq 3$. To explain this, observe that if $\sigma(Z) = Z$ and Z is a solution of (1.8), then

$$\hat{Z}(x,t) = \lambda^{\frac{d}{2}-1} Z(\lambda x, \lambda^2 t),$$

satisfies

$$\hat{Z}_t = \Delta \hat{Z} + \lambda^{1-\frac{d}{2}} \hat{Z}\hat{\xi}.$$

This means that when $d = 1$, then nonlinear part can be ignored locally (near $(0,0)$, and hence near any (x_0, t_0) by translation); a property that is our requirement for a subcritical SPDE. When $d \geq 2$, we may replace ξ by a smooth approximation ξ^ε, and study the behavior of the corresponding solution as $\varepsilon \to 0$. As it turns out the nonlinear term blows up. To guarantee that our SPDE has an interesting solution, we multiplying the product $Z\xi$ by a small number that goes to 0 as $\varepsilon \to 0$. This idea so far has been rigorously worked out when $d = 2$. To explain this, let us take a smooth function of compact support $\chi : \mathbb{R}^d \to [0, \infty)$ with $\int \chi \, dx = 1$, and put $\xi^\varepsilon = \xi *_x \chi^\varepsilon$ (convolution in x), where $\chi^\varepsilon(x) = \varepsilon^{-d}\chi(x/\varepsilon)$. When $d = 2$, we formulate the SPDE

$$Z_t^\varepsilon = \Delta Z^\varepsilon + \beta |\log \varepsilon|^{-\frac{1}{2}} Z^\varepsilon \xi^\varepsilon. \tag{1.9}$$

A phase transition occurs as we vary β. When $\beta \in (0, 2\pi)$, then the small ε limit of the process Z^ε exists (see [DG] and [CSZ]). When $d = 3$, (1.8) is expected be renormalizable. For $d \geq 4$, no renormaliztion procedure is expected to be available.

(iii) Kardar-Parisi-Zhang equation (KPZ). This SPDE models the stochastic growth models. When the interface is given by the graph of a height function $h : \mathbb{R}^d \times [0, T] \to \mathbb{R}$, then as the first approximation, we may derive a Hamilton Jacobi PDE of the form $h_t = H(h_x)$ to model the evolution of h. To take into account the interface diffusion and thermal fluctuations, we may modify our PDE by adding a second order term and a stochastic term. After some manipulations involving a rescaling, a translation, and a Taylor expansion of the Hamiltonian function H, we end up with KPZ equation

$$h_t = \Delta h + |h_x|^2 + \xi. \tag{1.10}$$

Formally, we may apply Hopf-Cole transform $Z = e^h$ to derive (1.10) with $\sigma(Z) = Z$. Since we expect h to have the same regularity as solutions of (1.7), it is not clear how to make sense e^h when $d > 1$. However, when $d = 1$, we already know that (1.10) has a solution à la Itô, and we may try to use this solution to construct a solution $h = \log Z$ for (1.10). It turns out, $h = \log Z$ does not solve (1.10), because in our formal derivation we have not taken into account the Itô correction term as we apply chain rule. In fact the Itô correction in this case is infinite which has to do with the fact that the singularity of (1.10) is much worse than (1.8). In (1.8), we are multiplying the white noise ξ with the Hölder continuous function Z, in KPZ equation however we are squaring a distribution $h_x \in \mathcal{C}^{(-\frac{1}{2})-}$. To get a better feel for the infinite term that shows up in our calculation, we replace ξ with an approximation ξ^ε that is smooth in x, and white noise in t, and examine the type of divergence the corresponding solution Z^ε exhibits. More precisely we pick a smooth function $\chi : \mathbb{R} \to [0, \infty)$, define $\chi^\varepsilon(x) = \varepsilon^{-1}\chi(x/\varepsilon)$, and put $\chi^\varepsilon := \xi *_x \chi^\varepsilon$ (convolution in x), and consider the solution

$$Z_t^\varepsilon = Z_{xx}^\varepsilon + Z^\varepsilon \xi^\varepsilon. \tag{1.11}$$

It is not hard to see that for fixed x, the process $t \mapsto \xi^\varepsilon(x,t)$ is a constant multiple of the standard white noise:

$$\mathbb{E}\xi^\varepsilon(x,t)\xi^\varepsilon(x,s) = C^\varepsilon \delta_0(t-s),$$

where

$$C^\varepsilon = \int \chi^\varepsilon(x-y)\chi^\varepsilon(x-y)\ dy = \varepsilon^{-1}\int \chi^2(y)\ dy =: \varepsilon^{-1}\bar{C}(\chi).$$

Hence, we can write $\xi^\varepsilon(x,t) = \dot{B}^x(t)$, where $B^\varepsilon(\cdot;x)$ is a Brownian motion with

$$\mathbb{E}\left[B^\varepsilon(t;x)\right]^2 = C^\varepsilon t, \tag{1.12}$$

Treating Z^ε as a diffusion with the drift $u = Z^\varepsilon_{xx}$, and $\sigma = (C^\varepsilon)^{\frac{1}{2}} Z^\varepsilon$, we learn

$$d(\log Z^\varepsilon) = \left[\frac{Z^\varepsilon_{xx}}{Z^\varepsilon} - \frac{C^\varepsilon (Z^\varepsilon)^2}{2(Z^\varepsilon)^2}\right]\ dt + dB(t;x).$$

Now if $h^\varepsilon = \log Z^\varepsilon$, then h^ε satisfies

$$h^\varepsilon_t = h^\varepsilon_{xx} + (h^\varepsilon_x)^2 - C_\varepsilon + \xi^\varepsilon, \tag{1.13}$$

where $C_\varepsilon := C^\varepsilon/2$. What we learn from this calculation is that if Z is a solution of (1.8), for $\sigma(Z) = Z$, then $h = \log Z$, in some sense solves

$$h_t = h_{xx} + \left(h_x^2 - \infty\right) + \xi. \tag{1.14}$$

Put it differently, if we smoothize ξ as above, then there exists a sequence $C_\varepsilon = \varepsilon^{-1}\bar{C}/2$ (with $\bar{C}$ depending on χ) such that the solution h^ε converges to a Hölder continuous h. This h is our candidate for a solution to (1.14). This means that we may not be able to find a reasonable solution of the original equation (1.10). But after a suitable *renormalization*, we have a candidate solution.

The above reasoning was based on the Hopf-Cole transform that is only applicable for KPZ equation. In 2013, Martin Hairer [H1] initiated a new proof for this renomalization phenomenon that does not rely on Hope-Cole transform.

Theorem 1.1 *Given a nice initial data g, and for C_ε as above, let h^ε be the unique solution of* (1.13) *subject to the initial condition $h^\varepsilon(x,0) = g(x)$. Then the limit of h^ε exists as $\varepsilon \to 0$.*

More importantly Hairer's approach has been applied to a number of *subcritical* SPDEs. Observe that if h solves (1.10), and $\hat{h}(x,t) = S_\lambda^{1-\frac{d}{2}} h(x,t)$, then

$$\hat{h}_t = \hat{h}_{xx} + \lambda^{\frac{d}{2}-1}|\hat{h}_x|^2 + \hat{\xi}.$$

This means that KPZ is critical when $d = 2$, and supercritical when $d > 2$. It turns out that when $d = 2$, and $\tilde{h}^\varepsilon$ solves

$$\tilde{h}^\varepsilon = \Delta\tilde{h}^\varepsilon + C^1_\varepsilon|\tilde{h}^\varepsilon|^2 + \xi^\varepsilon - C^2_\varepsilon,$$

then $\tilde{h}^\varepsilon$ has a limit that solves SHE, provided $C^1_\varepsilon = C|\log\varepsilon|^{-1/2}$, $C^2_\varepsilon = \bar{C}C^1_\varepsilon\varepsilon^{-2}$, and C is sufficiently small (see [CSZ]). When $d > 2$, it is conjectured that C^1_ε should be of order $O(\varepsilon^{\frac{d}{2}-1})$ in order to have a Gaussian limit of the solution to KPZ equation.

(iv) Stochastic Quantization/Euclidean Quantum Field Theory. *Euclidean Quantum Field Theory* (in short EQFT) is a formulation of quantum field theory where the three (spatial) coordinates of a world point are real and the time coordinate is purely imaginary (this would turn the Minkowski metric to the Euclidean metric). In 1981 Parisi and Wu proposed an dynamical approach to the construction of probability measures which arise in EQFT, namely

$$Z^{-1}\ e^{-\mathcal{H}(\phi)}\ \mathcal{D}\phi. \tag{1.15}$$

Here the Hamiltonian (energy or actions depending on the interpretation) $\mathcal{H}$ is defined on a suitable (generalized) function space, Z is a normalizing constant, and $\mathcal{D}\phi$ is presumably a Lebesgue-like measure on our function space. The task of *constructive field theory* is to make sense of (1.15). One strategy to achieve this is to cook up a dynamics for which (1.15) is invariant. Formally, the gradient flow of $\mathcal{H}$ perturbed by ξ would do the job, namely

$$\phi_t = -\partial\mathcal{H}(\phi) + \xi, \tag{1.16}$$

where $\partial\mathcal{H}$ is a suitable (variational) derivative of $\mathcal{H}$ (for this, we need a metric on our function space, and the same metric is used as we define the white noise). Once we make sense of (1.15) (this is our *stochastic quantization*), then the law of $\phi(t)$ should converge to our desired measure as $t \to \infty$. The point is that the equation (1.16) has a better chance for a mathematical treatment than the measure (1.15).

As a classical example, consider

$$\mathcal{H}(\phi) = \int_{\mathbb{R}^d}\left[\frac{1}{2}|\nabla\phi|^2 + V(\phi)\right]\ dx, \tag{1.17}$$

for a potential function $V : \mathbb{R}^d \to \mathbb{R}$. If we use L^2 inner product for a metric, the corresponding SPDE takes the form

$$\phi_t = \Delta\phi - V'(\phi) + \xi. \tag{1.18}$$

As a simple example, consider the free case $V = 0$. Then the corresponding (1.16) is SHE. If we write $\langle\cdot,\cdot\rangle$ for L^2 inner product, then from

$$\frac{1}{2}\int|\nabla\phi|^2\ dx = -\int\phi\ \Delta\phi\ dx,$$

we deduce that the measure (1.15) in this case is a Gaussian field with correlation $(-\Delta)^{-1}$, which can be constructed by standard arguments. This measure is known as *Gaussian Free Field* (in short GFF). In summary, in large t limit, the law of the solution $(\partial_t - \Delta)^{-1}\xi$ converges to the GFF measure μ_{GFF} of correlation $(-\Delta)^{-1}$. From our discussion regarding the regularity of SHE, we expect that the support of GFF to be contained in $\mathcal{C}^{(1-d/2)-}$. A natural way to construct GFF is to construct it in a bounded domain D with a suitable boundary condition. The correlation $(-\Delta)^{-1}$ has a kernel G^D that is known as the *Green's function.* In other words

$$\int \phi(x)\phi(y)\ \mu^D_{GFF}(d\phi) = G^D(x,y).$$

In fact we can construct ϕ from the white noise directly. To explain this, let us write $\boldsymbol{\eta}$ for the d-dimensional white noise in dimension d. In other words, $\boldsymbol{\eta} = (\eta_1, \dots, \eta_d)$, where $\eta_1, \dots, \eta_d$ are independent white noise in $\mathbb{R}^d$. Now, the solution

$$-\Delta\phi = \nabla \cdot \boldsymbol{\eta},$$

(with say 0 Dirichlet boundary condition) is distributed as GFF. To see this, observe

$$\begin{aligned}
\mathbb{E}\phi(x)\phi(x') =& \mathbb{E}\left[\int G^D(x,y)(\nabla\cdot\boldsymbol{\eta})(y)\ dy\right]\left[\int G^D(x',y')(\nabla\cdot\boldsymbol{\eta})(y')\ dy'\right] \\
=& \mathbb{E}\left[\int G^D_y(x,y)\cdot\boldsymbol{\eta}(y)\ dy\right]\left[\int G^D_y(x',y')\cdot\boldsymbol{\eta}(y')\ dy'\right] \\
=& \int G^D_y(x,y)\cdot G^D_y(x',y)\ dy = \int G^D(x,y)\cdot(-\Delta G^D)(x',y)\ dy = G^D(x,x').
\end{aligned}$$

In summary $(-\Delta)^{-1}\nabla\cdot\boldsymbol{\eta}$ is distributed as μ^D_{GFF}.

In dimension one, our GFF is nothing other than a Brownian measure. For example, when $D = (0,\infty)$ with 0 Dirichlet boundary condition, then the Green's function is $G_D(x,y) = \min(x,y)$, and GFF $\phi(x) = B(x)$ is the Wiener measure. When $D = (0,\ell)$, then $G_D(x,y) = \min(x,y) - \ell^{-1}xy$ and GFF is the law of a Brownian bridge; $\phi(x) = B(x) - \ell^{-1}B(\ell)x$.

In dimension one, we can also make sense of the measure (1.15) for $\mathcal{H}$ as in (1.17), by rewriting it as

$$Z^{-1}\ e^{-\int V(\phi)\ dx}\ \mu_0(d\phi),$$

where μ_0 is the Wiener measure.

In dimension 2, GFF has been extensively investigated in recent years because of its connection with *Schramm-Loewner Evolution* (in short SLE), and *Liouville quantum gravity surface* (in short LQGS). Its relevance to conformal field theory has to do with the fact that of $f : D \to D'$ is a conformal map, then $G^D(z,z') = G^{f(D)}(f(z), f(z'))$. The level curves of GFF are distributed according to SLE_4. Moreover, GFF can be used to construct random

Riemannian LQGS. The metric in *isothermal* coordinates is given by $e^{\gamma\phi(z)}|dz|^2$, where ϕ is selected according to μ_{GFF}. Though a suitable renormalization must be performed to make sense of this metric because $\phi \in \mathcal{C}^{0-}$ is a distribution. In fact, if we replace ϕ with a smooth approximation ϕ^ε by a convolution with a standard approximation to identity, then after a renormalization

$$e^{\gamma\phi^\varepsilon(z)+\gamma\log\varepsilon/2}|dz|^2,$$

we have a sequence of random metrics that has a limit in some weak sense (for example the corresponding area form converges).

In many examples of interest the SPDE (1.16) is nonlinear and since for $d \geq 2$, we expect ϕ to be a distribution, we encounter the problem of making sense of $\partial\mathcal{H}(\phi)$.

(v) Dynamical Φ^4_d equation. This is the SPDE (1.18), when and $V(\phi) = 4^{-1}\beta\phi^4$, for $\beta > 0$:

$$\phi_t = \Delta\phi - \beta\phi^3 + \xi. \tag{1.19}$$

It is also expected to model the magnetization at the critical temperature when $d = 3$. When a magnet is heated up, it loses its magnetic property after a certain critical (Curie) temperature T_c. As a phenomenological model for magnetism, we associate spins $\pm$ to points of a d-dimensional lattice and let spins change according to some stochastic rule (Ising model/Glauber Dynamics). Near the critical temperature, the magnetic field fluctuations are conjectured to be governed by the SPDE

Note that if $\hat\phi(x,t) = \lambda^{\frac{\lambda}{2}-1}\phi(\lambda x, \lambda^2 t)$, then

$$\hat\phi_t = \Delta\hat\phi - \lambda^{4-d}(\hat\phi)^3 + \hat\xi,$$

which means that (1.17) is subcritical when $d \leq 3$, and critical when $d = 4$. We now give a brief historical account some of the work that is done for this model.

We have already discussed the mathematical treatment of the measure (1.15) when $\mathcal{H}$ is given by (1.17). We are tempted to write (1.15) as

$$Z^{-1}e^{\int V(\phi)\ dx}\ \mu_{GFF}(d\phi). \tag{1.20}$$

Since $\phi \in \mathcal{C}^{(1-d/2)-}$ in the support of μ_{GFF}, we encounter the problem of making sense $V(\phi)$ when $d \geq 2$. We may replace $\mathbb{R}^D$ with the lattice $\mathbb{Z}^d$, and replace ∇ with the discrete gradient to have a well-defined measure. Then a suitable scaling limit of this discrete approximation would serve as our candidate for the measure (1.20). This scaling limit does not exist for general V when $d \geq 2$. Though we expect that the issue of divergence would no so drastic in dimension two because $\phi \in \mathcal{C}^{0-}$. To explain this, let us first rewrite (1.20) with

$$Z^{-1}e^{\int_{\mathbb{T}^2} P(\phi)\ dx}\ \mu_0(d\phi), \tag{1.21}$$

where P is a polynomial, and μ_0 is a Gaussian field with the covariance operator $(1-\Delta)^{-1}$ (known as *massive Gaussian Free Field*) that is acting on periodic functions. One can make sense of (1.21) only after a *Wick ordering* renormalization. This means that given a polynomial $Q(\phi)$ of even degree, we renormalize/replace Q with $P =: Q :$, where $: \cdot :$ is a linear operation, and $: \phi^r :$ is the r-th Wick power of ϕ with respect to the measure μ_0. Si,milarly, the measure (1.21) is invariant for the SPDE,

$$\phi_t = \Delta\phi - : Q'(\phi) : + \xi.$$

In particular, we can make sense of (1.19) when $d = 2$ after replacing ϕ^3 with $: \phi^3 :$ (see the introduction of [AK] for an overview).

In 2014, Hairer [H2] succeeded to make sense of the SPDE (1.19) in dimension three by employing his theory of regularity structures. According to [H2], the SPDE (1.19) must be renormalized in the following sense: If ξ^ε is as before, and ϕ^ε solves

$$\phi_t^\varepsilon = \Delta\phi^\varepsilon - (\phi^\varepsilon)^3 + C_\varepsilon\phi^\varepsilon + \xi^\varepsilon, \tag{1.22}$$

for a suitably selected constant $C_\varepsilon = O(\varepsilon^{-1})$, then ϕ^ε has a limit as $\varepsilon \to 0$.

(vi) Dynamical sine-Gordon equation. This is the SPDE (1.16), when

$$\mathcal{H}(\phi) = \int_{\mathbb{R}^d} \left[\frac{1}{2}|\nabla\phi|^2 + C\beta^{-1}\cos(\beta\phi)\right] dx,$$

so that ϕ satisfies

$$\phi_t = \Delta\phi + C\sin(\beta\phi) + \xi. \tag{1.23}$$

The renomalized version of (1.23) reads as

$$\phi_t^\varepsilon = \Delta\phi^\varepsilon + C_\varepsilon\sin(\beta\phi^\varepsilon) + \xi^\varepsilon. \tag{1.24}$$

To have an interesting limit for ϕ^ε when $d = 2$ and $\beta \in (0, 4\pi)$, we need to choose $C_\varepsilon = O(\varepsilon^{-\frac{\beta^2}{4\pi}})$. For $\beta \geq 4\pi$ the limit is supposed to be Gaussian process. If we do not renormalize, leave C independent of ε, then the oscillatory term $\sin(\beta\phi^\varepsilon)$ averges out and converges to 0 as $\varepsilon \to 0$.

(vii) Parabolic Anderson Model (PAM) A Brownian particle $x(t)$ that is killed at rate $V(x)$ is associated with the parabolic PDE $u_t = \Delta u + Vu$. PAM describes a Brownian particle with a random killing rate that is given by spatial white noise $\eta(x)$:

$$u_t = \Delta u + \eta u.$$

(viii) Nonlinear Schrödinger equation. In **(iv)**, we learn that if we add noise to the gradient flow associated with the Hamiltonian function $\mathcal{H}$, then the Gibbs-like measure (1.15) is invariant for the dynamics. On the other hand, if instead of gradient flow, we consider the Hamiltonian ODE, then the enegry and the measure $\mathcal{D}\phi$ are both invariant for the dynamics, and as a consequence the measure (1.15) is also invariant. As an example, consider complex-valued field $\psi : \mathbb{R}^d \to \mathbb{C}$, and the Hamiltonian function

$$\mathcal{H}(\psi) = \int \left[\frac{1}{2} |\nabla \psi|^2 + \frac{\kappa}{r+1} |\psi|^{r+1} \right] \, dx,$$

where $\kappa \in \{\pm 1\}$. We also choose the (L^2) inner product and the *symplectic form*

$$\langle \psi, \varphi \rangle = \int Re(\psi \bar{\varphi}) \, dx, \qquad \omega(\psi, \varphi) = \int Im(\psi \bar{\varphi}) \, dx.$$

Note that $\omega(\psi, \varphi) = \langle J\psi, \varphi \rangle$, where the *complex structure* J is simply the multiplication by $-i$. If $\partial \mathcal{H}$ denotes the gradient of $\mathcal{H}$ with respect to L^2-inner product, then the Hamiltonian system associated with $\mathcal{H}$ and the complex structure J take the form $i\psi_t = \partial \mathcal{H}(\psi)$. More specifically,

$$i\psi_t = -\Delta \psi + \kappa |\psi| \psi^{r-1}.$$

1.1 Stochastic Growth Models

Various phenomena in physics and biology, such as the formation of crystals and the spread of infections are modeled by stochastic growth models. Many of such growth models are macroscopically described by Hamilton-Jacobi partial differential equations. In these models, a random interface separates regions associated with different phases and the interface can be locally approximated by the graph of a solution to a Hamilton–Jacobi equation. Such a solution gives us a macroscopic description of the interface. Microscopically though, the interface is rough and fluctuates about the macroscopic solution. A central limit theorem should provide us with a better description of the interface. This central limit theorem is expected to yield KPZ equation for many stochastic growth models.

As a concrete stochastic growth model, consider the diffusion $\mathbf{h}(t) = (h_i(t) : \; i \in \mathbb{Z})$ that satisfies the SDE

$$\begin{aligned} \frac{dh_i}{dt} &= \left(\sigma + \frac{\gamma \beta^2}{2} \right) V'(h_{i+1} - h_i) + \left(\sigma - \frac{\gamma \beta^2}{2} \right) V'(h_i - h_{i-1}) + \beta \frac{dB_i}{dt} \\ &= \sigma (V'(h_{i+1} - h_i)) + V'(h_i - h_{i-1})) \\ &\quad + \frac{\gamma \beta^2}{2} (V'(h_{i+1} - h_i) - V'(h_i - h_{i-1})) + \beta \frac{dB_i}{dt}. \end{aligned} \tag{1.25}$$

where $\beta, \gamma > 0$ and B_i's are independent Brownian motions. If we interpret h_i as the height at site i, and write $r_i = h_i - h_{i-1}$, for the height difference, then

$$\frac{dr_i}{dt} = \left(\sigma + \frac{\gamma\beta^2}{2}\right) V'(r_{i+1}) + \left(\sigma - \frac{\gamma\beta^2}{2}\right) V'(r_i) - \left(\sigma + \frac{\gamma\beta^2}{2}\right) V'(r_i)$$
$$- \left(\sigma - \frac{\gamma\beta^2}{2}\right) V'(r_{i-1}) + \beta\left(\frac{dB_i}{dt} - \frac{dB_{i-1}}{dt}\right).$$

Writing D_i and $D_{i,i+1}$ for $\frac{\partial f}{\partial r_i} - \frac{\partial f}{\partial r_{i+1}}$ and $\frac{p}{\partial r_i}$ respectively, the generator of $\mathbf{h}$ can be written as $\mathcal{L} = \sigma\mathcal{A} + \beta^2\mathcal{S}$, where

$$\mathcal{A} = \sum_i (V'(r_{i+1}) + V'(r_i))D_{i,i+1} = \sum_i (V'(r_{i+1}) - V'(r_{i-1}))D_i,$$

$$\mathcal{S} = \frac{1}{2}\sum_i D_{i,i+1}^2 - \gamma(V'(r_i) - V'(r_{i+1}))D_{i,i+1}. \tag{1.26}$$

We now argue that $\mathcal{A}$ is invariant with respect to

$$\nu_\alpha(d\mathbf{r}) = \prod_i \frac{1}{Z(\alpha)} e^{\alpha r_i - \gamma V(r_i)},$$

with $Z(\alpha) = \int e^{\alpha r - V(r)} dr$.

$$\int g\ \mathcal{S}f\ d\nu_\alpha = \int f\ \mathcal{S}g\ d\nu_\alpha = -\frac{1}{2}\int (D_{i,i+1}f)(D_{i,i+1}g)d\nu_\alpha,$$

and $\int \mathcal{A}f\ d\nu_\alpha = 0$, for every nice (local) function f and g. To see this, observe

$$\int \sum_i (V'(r_{i+1}) - V'(r_{i-1}))\frac{\partial f}{\partial r_i} d\nu_\alpha = \int \sum_i (V'(r_{i+1}) - V'(r_{i-1}))(-\alpha + \gamma V'(r_i))d\nu_\alpha = 0.$$

We may rescale the height as

$$h^\varepsilon(x, t) = \varepsilon h_{[x/\varepsilon]}(t/\varepsilon),$$

where ε represents the ratio of the macroscopic and microscopic scales. We wish to study the behavior of h^ε as $\varepsilon \to 0$. To ease the notation, let us assume that $\beta = \gamma = 1$. We also write $\hat{h}_i$ and $\hat{r}_i$ for $\varepsilon h_i(t/\varepsilon)$ and $r_i(t/\varepsilon)$ respectively. Observe

$$\frac{d\hat{h}_i}{dt} = \sigma(V'(\hat{r}_{i+1}) + V'(\hat{r}_{i-1})) + \frac{1}{2}(V'(\hat{r}_{i+1}) - V'(\hat{r}_i)) + \varepsilon^{3/2}\frac{d\hat{B}_i}{dt}. \tag{1.27}$$

From this, we expect that $\hat{h} \to u$ in low ε limit, with u satisfying a Hamilton-Jacobi PDE of the form

$$u_t = H(u_x), \tag{1.28}$$

where $H(\rho) = 2\sigma\lambda(\rho)$ and λ is given explicitly by

$$\int r\nu_\lambda(dr) = \rho.$$

One can derive (1.28) using the method of [R] provided that V is a convex function.

Regarding (1.28) as a law of large numbers, we may attempt to establish a central limit theorem (CLT) to capture the nature of the fluctuations of $\mathbf{h}$ about the macroscopic profile u. For this, let us write $\hat{u}(x,t)$ and $\hat{\rho}(x,t)$ for $\hat{h}_{[x/e]}(t)$ and $\hat{r}_{[x/\varepsilon]}(t)$. Inspired by (1.27), we make an ansatz that $\hat{h}$ roughly satisfies

$$\hat{u}_t \approx H(\hat{\rho}) + 2^{-1}\varepsilon\lambda(\hat{\rho})_x + \varepsilon\xi. \tag{1.29}$$

Here we are using the fact that $\varepsilon^{1/2}\dot{B}_i(t)$ has the same law as $\xi(i/\varepsilon, t)$. We now start the system from the equilibrium measure with density $\bar{\rho}$. Initially,

$$\hat{u}(x,0) = \bar{\rho}x + \varepsilon^{1/2}B_{\bar{\rho}}(x) + o(\varepsilon^{1/2}),$$

for a Brownian motion $B_{\bar{\rho}}$. Assuming that we have a similar CLT at later times, we can write

$$\hat{u}(x,t) = \bar{\rho}x + tH(\bar{\rho}) + \varepsilon^{1/2}w(x,t) + o(\varepsilon^{1/2}), \qquad \hat{\rho}(x,t) = \bar{\rho} + \varepsilon^{1/2}w_x(x,t) + o(\varepsilon^{1/2}).$$

From (1.29), we expect

$$w_t \approx 2\sigma\lambda'(\bar{\rho})w_x + \varepsilon^{1/2}\sigma\lambda''(\bar{\rho})w_x^2 + 2^{1/2}\varepsilon\lambda'(\bar{\rho})w_{xx} + \varepsilon^{1/2}\xi.$$

We now choose $\sigma = \varepsilon^{1/2}/2$ and speed up the time. If

$$\hat{w}(x,t) = w(x,t/\varepsilon),$$

then $\hat{w}$ is expected to satisfy

$$\hat{w}_t \approx \varepsilon^{-1/2}\lambda'(\bar{\rho})\hat{w}_x + 2^{-1}\lambda''(\bar{\rho})\hat{w}_x^2 + 2^{1/2}\lambda'(\bar{\rho})w_{xx} + \xi.$$

To get ride of the first term on the right-hand side, we set

$$k^\varepsilon(x,t) = k(x,t) = \hat{w}\left(x - \varepsilon^{-1/2}\lambda'(\bar{\rho})t, t\right),$$

so that

$$k_t \approx +2^{-1}\lambda''(\bar\rho)k_x^2 + 2^{1/2}\lambda'(\bar\rho)k_{xx} + \xi.$$

From this we conjecture that $k_\varepsilon \to h$ in low ε limit, with h satisfying KPZ equation

$$h_t = 2^{-1}\lambda''(\bar\rho)h_x^2 + 2^{1/2}\lambda'(\bar\rho)h_{xx} + \xi. \tag{1.30}$$

This has been rigorously verified by Goncalves and Jara [GJ] for a discrete variant of the above model know as Simple Exclusion Process. Later Gubinelli and Perkowski use the same approach to establish (1.30) for GZ model. Chapter 6 is devoted to the derivation of (1.30).

2 Rough Path Integration

In this section we focus on differential equations that are driven by rough paths. More precisely, let $x : [0,T] \to \mathbb{R}^\ell$ be a Hölder continuous function of Hölder exponent α, and consider

$$\dot{y}(t) = \sigma(y(t))\ \dot{x}(t), \tag{2.1}$$

where $\sigma : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d\times\ell}$ is a C^2 function that takes value in the space of $d \times \ell$ matrices. As we mentioned in the Introduction, we may attempt to make sense of (2.1) by rewriting it as

$$y(t) - y(s) = \int_s^t \sigma(y(\theta))\ dx(\theta), \tag{2.2}$$

and try to make sense the integral that appears on the right-hand side. Since $x \in \mathcal{C}^\alpha$, we expect $y \in \mathcal{C}^\alpha$. As a natural strategy, we may write

$$I := \int_s^t \sigma(y(\theta))\ dx(\theta) = \sum_{i=0}^n \int_{t_i}^{t_{i+1}} \sigma(y(\theta))\ dx(\theta),$$

for a mesh $\pi = \{s = t_0 < t_1 < \cdots < t_n < t_{n+1} = t\}$, with

$$|\pi| = \max_{i=0}^n (t_{i+1} - t_i),$$

and approximate

$$I_i := \int_{t_i}^{t_{i+1}} \sigma(y(\theta))\ dx(\theta) = \sigma(y(t_i))(x(t_{i+1}) - x(t_i)) + O(|t_{i+1} - t_i|^{2\alpha}). \tag{2.3}$$

It is not hard to show that when $\alpha > 1/2$, the limit

$$I = \lim_{|\pi|\to 0} \sum_{i=0}^n \sigma(y(t_i)(x(t_{i+1}) - x(t_i)),$$

exists, which offers a natural candidate for the right-hand side of (2.2), because

$$\lim_{|\pi|\to 0} \sum_{i=0}^n O(|t_{i+1} - t_i|^{2\alpha}) = 0. \tag{2.4}$$

The convergence (2.4) does not hold when $\alpha \le 1/2$, which calls for a better approximation than what we used in (2.3). Perhaps we should try

$$\begin{aligned}I_i =&\sigma(y(t_i))(x(t_{i+1})-x(t_i))+\sum_{r=1}^{d}\sigma_{y_r}(y(t_i))\int_{t_i}^{t_{i+1}}(y_r(\theta)-y_r(t_i))\ dx(\theta)+O(|t_{i+1}-t_i|^{3\alpha})\\ =&\sigma(y(t_i))(x(t_{i+1})-x(t_i))+\sum_{r=1}^{d}\sum_{j=1}^{\ell}\sigma_{y_r}(y(t_i))\sigma^{rj}(y(t_i))\int_{t_i}^{t_{i+1}}(x_j(\theta)-x_j(t_i))\ dx(\theta)\\ &+O(|t_{i+1}-t_i|^{3\alpha}).\end{aligned}$$

Note that when $\alpha \in (1/3, 1/2)$, then

$$\lim_{|\pi|\to 0}\sum_{i=0}^{n} O(|t_{i+1}-t_i|^{3\alpha}) = 0.$$

As a result, we may use our approximation for I_i to find a candidate for I as a limit, provided that we have a candidate for integrals of the form

$$\mathbb{X}(s,t)=\int_s^t\int_s^\theta x(d\theta')\otimes x(d\theta)=\int_s^t x(s,\theta)\otimes x(d\theta)=\int_s^t x(\theta)\otimes x(d\theta)-x(s)\otimes x(s,t),$$

where

$$x(s,t) := x(t)-x(s).$$

Observe that if $z(t)$ denotes our candidate for $\int_0^t x(\theta)\otimes x(d\theta)$, then we expect

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

In other words, $\int_s^t x(\theta)\otimes x(d\theta)$ is an increment of a function z. This expressed in terms of $\mathbb{X}$ takes the form

$$\mathbb{X}(s,t)=\mathbb{X}(s,u)+\mathbb{X}(u,t)+x(s,u)\otimes x(u,t), \tag{2.5}$$

which is known as *Chen's relation*. Note that for an error of order $O(|t_{i+1}-t_i|^{3\alpha})$, we need $|\mathbb{X}(s,t)| = O(|t_{i+1}-t_i|^{2\alpha})$. These considerations suggest the following definition.

Definition 2.1(i) We write $\mathcal{R}^\alpha = \mathcal{R}^\alpha([0,T];\mathbb{R}^\ell)$ for the set of pairs $\mathbf{x} = (x, \mathbb{X})$ such that $x : [0,T] \to \mathbb{R}^\ell,\ \mathbb{X} : [0,T]^2 \to \mathbb{R}^{\ell\times\ell}$, (2.5) holds, and

$$\|\mathbf{x}\|_{\alpha,2\alpha} = |x(0)| + [(x,\mathbb{X})]_{\alpha,2\alpha} = |x(0)| + [x]_\alpha + [\mathbb{X}]_{2\alpha} < \infty,$$

where

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

(ii) We say $(x,\mathbb{X}) \in \mathcal{R}^\alpha$ is *weakly geometric* if

$$\mathbb{X}(s,t) + \mathbb{X}(s,t)^* = x(s,t) \otimes x(s,t), \tag{2.6}$$

where $\mathbb{X}^*$ denotes the transpose of the matrix $\mathbb{X}$. The set of weakly geometric $\mathbf{x}$ is denoted by $\mathcal{R}^\alpha_g$.

(iii) If $x : [0,T] \to \mathbb{R}$ is smooth, then we write

$$\mathbb{X}^x_0(s,t) = \int_s^t x(s,t) \otimes x(d\theta),$$

where the write-hand side is the standard integral a la Riemann. We write $\mathcal{R}^\alpha_{sg}$ for the closure of the set

$$\left\{(x, \mathbb{X}^x_0) : \ x : [0,T] \to \mathbb{R} \ \text{ is smooth}\right\},$$

with respect to $\|\cdot\|_{\alpha,2\alpha}$-norm. Note that $\mathcal{R}^\alpha_{sg} \subset \mathcal{R}^\alpha_g$. We refer to $\mathbf{x} \in \mathcal{R}^\alpha_{sg}$ as *(strongly) geometric rough paths.*

(iv) Given $x \in \mathcal{C}^\alpha$, we write $\mathcal{G}^\alpha(x) = \mathcal{G}^\alpha(x;[0,T];\mathbb{R}^{d\times\ell})$ for the set of pairs $\mathbf{y} = (y,\hat{y})$ such that $y : [0,T] \to \mathbb{R}^{d\times\ell}$, $\hat{y} : [0,T] \to \mathbb{R}^{d\times\ell\times\ell}$, such that

$$[(y,\hat{y})]_{x;\alpha,2\alpha} = [\hat{y}]_\alpha + [y,\hat{y}]_{x;2\alpha}, \qquad [y,\hat{y}]_{x;2\alpha} := \sup_{s\neq t} \frac{|y(s,t) - \hat{y}(s)x(s,t)|}{|s-t|^{2\alpha}} < \infty.$$

Here $\hat{y} = [\hat{y}^{ijk}]$, $\hat{y}(s)x(s,t) = [(\hat{y}(s)x(s,t))^{ij}]$, with

$$(\hat{y}(s)x(s,t))^{ij} = \sum_{k=1}^{\ell} \hat{y}(s)^{ijk} x(s,t)_k.$$

We refer to $\hat{y}$ as a *Gubinelli derivative* of y and regard it as a candidate for dy/dx.

(v) Given $\mathbf{x} = (x,\mathbb{X})$, $\mathbf{x}' = (x',\mathbb{X}') \in \mathcal{R}^\alpha$, and $\mathbf{y} = (y,\hat{y}) \in \mathcal{G}^\alpha(\mathbf{x})$, $\mathbf{y}' = (y',\hat{y}') \in \mathcal{G}^\alpha(\mathbf{x}')$, we define $[\mathbf{y};\mathbf{y}']_{x,x';\alpha,2\alpha} = [\hat{y} - \hat{y}']_\alpha + [\mathbf{y};\mathbf{y}']_{x,x';2\alpha}$, where

$$[\mathbf{y};\mathbf{y}']_{x,x';2\alpha} := \sup_{s\neq t} \frac{|y(s,t) - \hat{y}(s)x(s,t) - y'(s,t) + \hat{y}'(s)x'(s,t)|}{|s-t|^{2\alpha}}.$$

□

Remark 2.1(i) As we stated in the Introduction, according to a result of Lyons and Victoir, for every $x(\cdot) \in \mathcal{C}^\alpha$, there exists $z \in \mathcal{C}^\alpha$ and a constant c such that $z(0) = 0$, and

$$|z(t) - z(s) - x(s) \otimes x(s,t)| \le c|t-s|^{2\alpha}, \tag{2.7}$$

for every $s, t \in [0,T]$. As a consequence, if

$$\mathcal{R}^\alpha(x) = \{\mathbb{X} : \ (x, \mathbb{X}) \in \mathcal{R}^\alpha\},$$

is nonempty. Note that if $\mathbb{X}, \mathbb{X}' \in \mathcal{R}^\alpha(x)$, then there exist function $w : [0,T] \to \mathbb{R}^{\ell \times \ell}$ such that $w \in \mathcal{C}^{2\alpha}$, and

$$\mathbb{X}'(s,t) = \mathbb{X}(s,t) + w(t) - w(s). \tag{2.8}$$

The converse is also true: If $\mathbb{X} \in \mathcal{R}^\alpha(x)$, and $\mathbb{X}'$ is given by (2.8), then $\mathbb{X}' \in \mathcal{R}^\alpha(x)$. It is worth mentioning that Lyons-Victoir result when $\ell = 1$ is trivial because the function $z(t) = x^2(t)/2$ does satisfy (2.7). The Reconstruction Theorem of Hairer is the generalization of Lyons-Victoir's result to higher dimensions. This generalization will be presented in Chapter 4.

(ii) Recall that a candidate of the integral $\int_0^t x \otimes dx$ yields a candidate for the distribution $x \otimes \dot{x}$. Our interest in geometric rough paths stems from the fact that the product rule of differentiation in calculus is valid for such paths. Indeed the condition (2.6) means

$$\int_s^t (x_i dx_j + x_j dx_i) = x_i(t)x_j(t) - x_i(s)x_j(s), \tag{2.9}$$

or our candidate for the $x_i\dot{x}_j + x_i\dot{x}_i$ coincides with the distribution derivative of x_ix_j. If we write $\mathcal{R}_g^\alpha(x)$ for the set of $\mathbb{X}$ such that $(x, \mathbb{X}) \in \mathcal{R}^\alpha$ is geometric, then $\mathbb{X} + \mathbb{X}^*$ is uniquely determined. Hence when $\ell = 1$, there is only one such $\mathbb{X}$. More generally, if $\mathbb{X}, \mathbb{X}' \in \mathcal{R}_g^\alpha$, then there exists an antisymmetric $w \in \mathcal{C}^\alpha$ such that (2.8) holds.

(iii) We assert that we can use (2.7) to show that if $x \in \mathcal{C}^\alpha$, then the set $\mathcal{R}_g^\alpha(x) \neq \emptyset$. Let us demonstrate this when $\ell = 2$. Without loss of generality, we may assume that $x(0) = 0$. If $x = (x_1, x_2)$, then by (2.7), there exists $z(t) = z_{12}(t)$ such that

$$|z(t) - z(s) - x_1(s)x_2(s,t)| \le c|t-s|^{2\alpha}.$$

We claim that if we set $z_{ii} = x_i^2/2$, and $z_{21} = x_1x_2 - z_{12}$, then

$$\mathbb{X} = z(t) - z(s) - x(s) \otimes x(s,t), \qquad z = [z_{ij}]_{i,j=1}^2,$$

is in $\mathcal{R}_g^\alpha(x)$. To show this, observe

$$\begin{aligned}
|z_{21}(t) - z_{21}(s) - x_2(s)x_1(s,t)| &= |(x_1x_2)(t) - (x_1x_2)(s) - z(t) + z(s) - x_2(s)x_1(s,t)| \\
&= |x_1(t)x_2(s,t) - z(t) + z(s)| \\
&\le |x_1(s)x_2(s,t) - z(t) + z(s)| + |x_1(s,t)x_2(s,t)| \\
&\le (c + [x_1]_\alpha [x_2]_\alpha)|t-s|^{2\alpha},
\end{aligned}$$

which shows that $\mathbb{X} \in \mathcal{R}^\alpha$. By construction, $(x, \mathbb{X}) \in \mathcal{R}_g^\alpha$.

(iv) Given $x \in \mathcal{C}^\alpha$, define

$$\mathcal{R}_g^\alpha(x) = \left\{\mathbb{X} : \ (x, \mathbb{X}) \in \mathcal{R}_g^\alpha\right\},$$

is nonempty. Note that if $\mathbb{X}, \mathbb{X}' \in \mathcal{R}_g^\alpha(x)$, then there exist function $w : [0, T] \to \mathbb{R}^{\ell\times\ell}$ such that $w \in \mathcal{C}^{2\alpha}$, w is antisymmetric, and (2.8) holds

$$\mathbb{X}'(s,t) = \mathbb{X}(s,t) + w(t) - w(s). \tag{2.10}$$

We can fix $\mathbb{X} \in \mathcal{R}_g^\alpha(x)$ and vary w as above to produce all members of $\mathcal{R}_g^\alpha(x)$. □

Example 2.1 Given a C^2 function $F : \mathbb{R}^\ell \to \mathbb{R}^\ell$, the pair $(y, \hat{y}) \in \mathcal{G}^\alpha$, for $y(t) = F(x(t))$, and $\hat{y}(t) = DF(x(t))$. □

We now prove a theorem of Gubinelli that is due to Lyons when $\mathbf{y}$ is as in Example 2.1.

Theorem 2.1 **(i)** *Assume that $\alpha \in (3^{-1}, 1)$, and let $\mathbf{x} = (x, \mathbb{X}) \in \mathcal{R}^\alpha$, $\mathbf{y} = (y, \hat{y}) \in \mathcal{G}^\alpha(x)$. Let $\pi_n = \{0 = t_0^n < t_1^n < \cdots < t_n^n < t_{n+1}^n = t\} : \ n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ be a family of partitions of $[0, t]$ such that $|\pi_n| \to 0$ in large n limit, and $\pi_n \subset \pi_{n+1}$ for every $n \in \mathbb{N}_0$. Then the limit*

$$\mathcal{I}(\mathbf{x}, \mathbf{y})(t) := \int_0^t \mathbf{y}\, d\mathbf{x} := \lim_{n\to\infty} \sum_{i=0}^{n} \left[y(t_i^n)x(t_i^n, t_{i+1}^n) + \hat{y}(t_i^n) : \mathbb{X}(t_i^n, t_{i+1}^n)\right], \tag{2.11}$$

exists. (Here by $\hat{y} : \mathbb{X}$ we mean a vector with the i-th component given by $\sum_{jk} \hat{y}_{ijk}\mathbb{X}_{ij}$.) Moreover, for every $s, t \in [0, T]$,

$$|\mathcal{I}(\mathbf{x}, \mathbf{y})(t) - \mathcal{I}(\mathbf{x}, \mathbf{y})(s) - y(s)x(s,t) - \hat{y}(s)\mathbb{X}(s,t)| \le c_0(\alpha)[\mathbf{x}]_{\alpha,2\alpha}[\mathbf{y}]_{x;\alpha,2\alpha}\ |t - s|^{3\alpha}, \tag{2.12}$$

where

$$c_0(\alpha) = 2^{3\alpha} \sum_{n=1}^{\infty} n^{-3\alpha}.$$

(ii) *Given $\mathbf{x}$ as above, the map $\mathcal{I}_\mathbf{x}(\mathbf{y}) = (\mathcal{I}(\mathbf{x}, \mathbf{y}), y)$ defines a linear operator*

$$\mathcal{I}_\mathbf{x} : \mathcal{G}^\alpha(x; [0, T]; \mathbb{R}^{d\times\ell}) \to \mathcal{G}^\alpha(x; [0, T]; \mathbb{R}^d),$$

that is continuous;

$$\begin{aligned} &[\mathcal{I}_\mathbf{x}(\mathbf{y})]_{x;\alpha,2\alpha} \le (c_0(\alpha)[\mathbf{x}]_{\alpha,2\alpha} + T^\alpha[\mathbb{X}]_{2\alpha})\ [\mathbf{y}]_{x;\alpha,2\alpha}, \\ &[\mathcal{I}_\mathbf{x}(\mathbf{y}); \mathcal{I}_{\mathbf{x}'}(\mathbf{y}')]_{x,x';\alpha,2\alpha} \le c_1\left([\mathbf{x} - \mathbf{x}']_{\alpha,2\alpha} + |\hat{y}'(0) - \hat{y}(0)| + T^\alpha[\mathbf{y}; \mathbf{y}']_{x,x';\alpha,2\alpha}\right). \end{aligned} \tag{2.13}$$

Proof(i) Let $\pi = \{s = t_0 < t_1 < \cdots < t_n < t_{n+1} = t\}$ is a partition of the interval $[s,t]$. Let us write

$$I(\pi) := \sum_{i=0}^{n} \left[y(t_i)x(t_i,t_{i+1}) + \hat{y}(t_i) : \mathbb{X}(t_i,t_{i+1}) \right].$$

Pick $i \in \{1,2,\dots,n\}$ so that $|t_{i+1} - t_{i-1}| \le 2n^{-1}|t-s|$. Observe

$$\begin{aligned} I(\pi) - I\left(\pi \setminus \{y_i\}\right) =& y(t_{i-1})x(t_{i-1},t_i) + \hat{y}(t_{i-1}) : \mathbb{X}(t_{i-1},t_i) + y(t_i)x(t_i,t_{i+1}) + \hat{y}(t_i) : \mathbb{X}(t_i,t_{i+1}) \\ & - y(t_{i-1})x(t_{i-1},t_{i+1}) - \hat{y}(t_{i-1}) : \mathbb{X}(t_{i-1},t_{i+1}) \\ =& y(t_{i-1},t_i)x(t_i,t_{i+1}) + \hat{y}(t_{i-1},t_i) : \mathbb{X}(t_i,t_{i+1}) - \hat{y}(t_{i-1}) : \left(x(t_{i-1},t_i) \otimes x(t_i,t_{i+1})\right) \\ =& \left[y(t_{i-1},t_i) - \hat{y}(t_{i-1})x(t_{i-1},t_i)\right] \cdot x(t_i,t_{i+1}) + \hat{y}(t_i,t_{i-1}) : \mathbb{X}(t_{i-1},t_i). \end{aligned}$$

As a result

$$\begin{aligned} |I(\pi) - I\left(\pi \setminus \{t_i\}\right)| &\le \left[[y,\hat{y}]_{x;2\alpha}[x]_\alpha + [\hat{y}]_\alpha[\mathbb{X}]_{2\alpha}\right] |t_{i+1} - t_i|^{3\alpha} \\ &\le 2^{3\alpha} \left[[y,\hat{y}]_{x,2\alpha}[x]_\alpha + [\hat{y}]_\alpha[\mathbb{X}]_{2\alpha}\right] n^{-3\alpha}|t-s|^{3\alpha} =: Cn^{-3\alpha}|t-s|^{3\alpha}. \end{aligned}$$

From this, it is not hard to deduce the convergence in (2.10). Moreover, by induction,

$$|I(\pi) - I(\pi_0)| \le c_0(\alpha)C|t-s|^{3\alpha},$$

where $\pi_0 = \{t_0 = s < t_1 = t\}$. This is exactly (2.12).

(ii) The first inequality in (2.13) is an immediate consequence of (2.12). For the second inequality..... □

Remark 2.2(i) We note that when $\alpha > 1/2$, then $\mathcal{R}^\alpha(x)$ is a singleton, and we may simply write

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

and the term $\hat{y}(t_i^n) : \mathbb{X}(t_i^n, t_{i+1}^n)$ does not contribute to the integral and can be dropped. More generally with a verbatim argument we can show that the integral $\int_0^t y\ dx$ is well defined when $y \in \mathcal{C}^\beta,\ x \in \mathcal{C}^\alpha$, with $\alpha + \beta > 1$. We refer to the corresponding integral as Young integral.

(ii) Note that if $\mathbb{X}, \mathbb{X}' \in \mathcal{R}^\alpha$ are related by (2.8), then

$$\int_0^t \mathbf{y}\ d(x,\mathbb{X}') = \int_0^t \mathbf{y}\ d(x,\mathbb{X}) + \int_0^t \hat{y}\ dw,$$

where the second integral on the right-hand side is a Young integral. In Exercise **(iv)**, an example of w is given when $w(t) - w(s) = (t-s)C$, for an antisymmetric matrix C. □

When $\alpha > 1/2$, the existence of (unique) z satisfying (2.7) is due to Young as we stated in (1.6). In fact the proof of (1.6) can be carried out by an approximation scheme that is similar to (2.11). What (1.6) requires is that the function h near s can be approximated by $h(s) + A(s,t)$, where $A(s,t) = f(s)(g(t) - g(s))$. The condition $\alpha + \beta > 1$ yields a regularity of $A(s,\cdot)$ as we vary the base point s. More generally, we may formulate the following condition:

Definition 2.2 Let $A : [0,T]^2 \to \mathbb{R}$ be a continuous function. Given $\gamma > 0$, we say that A is γ-*coherent* if there exists a constant c_0 such that

$$|A(s,t) - A(s,u) - A(u,t)| \le c_0 |t-s|^{1+\gamma}, \tag{2.14}$$

for every $0 \le s \le u \le t \le T$. □

Example 2.2 Let $f, g : [0,T] \to \mathbb{R}$, with $f \in \mathcal{C}^\alpha,\ g \in \mathcal{G}^\beta$ with $\alpha + \beta > 1$. Then $A(s,t) = f(s)(g(t) - g(s))$ is γ coherent for $\gamma = \alpha + \beta - 1$. □

The following *Sewing Lemma* of Gubinelli guarantees the existence of a Hölder continuous function h that can be approximated by a coherent A.

Theorem 2.2 *Assume that A is γ-coherent for some $\gamma > 0$. Let $\{\pi_n : \ n \in \mathbb{N}_0\}$ be as in Theorem 1.1. Then*

$$h(t) = \lim_{n\to\infty} \sum_{i=0}^{n} A(t_i^n, t_{i+1}^n), \tag{2.15}$$

exists, and satisfies

$$|h(t) - h(s) - A(s,t)| \le c_0 c_1(\gamma) |t-s|^{1+\gamma},$$

where

$$c_1(\gamma) = 2^{1+\gamma} \sum_{n=1}^{\infty} n^{-1-\gamma}.$$

Proof Let π and t_i be as in Theorem 2.1, and set

$$I(\pi) = \sum_{i=0}^{n} A(t_i, t_{i+1}).$$

Then

$$\begin{aligned} |I(\pi) - I\left(\pi \setminus \{t_i\}\right)| =& |A(t_{i-1}, t_{i+1}) - A(t_{i-1}, t_i) - A(t_i, t_{i+1})| \\ \le& c_0 |t_{i+1} - t_{i-1}|^{1-\gamma} \le c_0 |t-s|^{1+\gamma} (2n)^{-1-\gamma}, \end{aligned}$$

and we can argue as in Theorem 2.1. □

Remark 2.3 When $A(s,t) = f(s)(g(t) - g(s))$, and $F_s = f(s)g'$, then F_s is a distribution that is assigned to the point s. If we set $\varphi = 1\!\!1_{[0,1]}$, and $\varphi_s^\delta(\theta) = \delta^{-1}\varphi(\delta^{-1}(\theta - s))$, then

$$\begin{aligned} A(s,t) - A(s,u) - A(u,t) &= (f(s)g')(1\!\!1_{[s,t]}) - (f(s)g')(1\!\!1_{[s,u]}) - (f(u)g')(1\!\!1_{[u,t]}) \\ &= (f(s)g')(1\!\!1_{[u,t]}) - (f(u)g')(1\!\!1_{[u,t]}) = (f(s) - f(u))\ g'(1\!\!1_{[u,t]}) \\ &= \delta(f(s) - f(u))\ g'(\varphi_u^\delta), \end{aligned}$$

where $\delta = t - u$. Hence we can rewrite (2.14) as

$$|(F_s - F_u)(\varphi_u^\delta)| \leq c_0\delta^{-1}\ (|s-u| + \delta)^{\gamma+1}. \tag{2.16}$$

This requires a regularity of F_s with respect to its base point (when $|s - u|$ is small), and a control on the order of singularity of the distribution F_u (when δ is small). It is the formulation (2.15) that can be generalized to higher dimension as we will see in Chapter 3. □

In search for a better understanding the space of rough paths associated with a path, let us observe that the pair $(x(s,t), \mathbb{X}(s,t))$ can be interpreted as some kind of increments if we use the correct algebraic interpretation. Note $x(s,t)$ is a vector (1-tensor), while $\mathbb{X}(s,t)$ is a matrix (2-tensor). Our construction does not go beyond 2-tensors because $\alpha > 1/3$. (As we may guess, we need tensors of orders up to k if $\alpha > 1/k$.) We may consider *truncated tensor algebra* of the form $V = \mathbb{R} \oplus \mathbb{R}^\ell \oplus \mathbb{R}^{\ell\times\ell}$, which is an algebra with the multiplication rule,

$$(a, v, A) \otimes (a', v', A') = (aa', av' + a'v, aA' + a'A + v \otimes v'). \tag{2.17}$$

Alternatively, we may write $a + v + A$ for (a, v, A), and derive (2.17) by multiplying out $(a + v + A) \otimes (a' + v' + A')$, and truncate (replace with 0) all tensors of order higher than 2. This suggests interpreting $\mathbf{x}(s,t) = (x(s,t), \mathbb{X}(s,t))$ as

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

which takes value in the set

$$G := \left\{1 + v + A :\ v \in \mathbb{R}^\ell,\ A \in \mathbb{R}^{\ell\times\ell}\right\},$$

which is a group $G \subset V$. We note that 1 plays the role of the unit, and

$$(1 + v + A)^{-1} = 1 - (v + A) + (v + A) \otimes (v + A) = 1 - (v + A) + v \otimes v,$$

is indeed the inverse of $1 + v + A$. Hence G is a Lie group. More importantly,

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

by Chen's relation. This gives an elegant (and compact when α is low) reformulation of Chen's relation. In particular $\mathbf{x}(0,s) \otimes \mathbf{x}(s,t) = \mathbf{x}(0,t)$ would lead to

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

Hence $\mathbf{x}(s,t)$ is an increment of a path $\mathbf{x}(t) := \mathbf{x}(0,t)$ with respect to the group structure. In summary, $\mathcal{R}^\alpha$ can be interpreted as the space of $\mathbf{x} : [0,T] \to G$ that is Hölder continuous of exponent α provide that we equip G with the right metric. First, we define

$$N(1 + v + A) := \max\{|v|, \sqrt{2|A|}\}, \quad \|\mathbf{x}\| := \frac{1}{2}(N(\mathbf{x}) + N(\mathbf{x}^{-1})), \quad d(\mathbf{x}, \mathbf{x}') := \|\mathbf{x}^{-1} \otimes \mathbf{x}'\|,$$

This yields a left-invariant metric on G. Now the Hölder norm

$$[\mathbf{x}(\cdot)]_\alpha = \sup_{s \neq t} \frac{d(x(s), x(t))}{|s-t|^\alpha},$$

is equivalent to

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

As for the weakly geometric rough path, observe that our condition (2.6) is equivalent to saying

$$\mathbb{X}(s,t) = \frac{1}{2} x(s,t) \otimes x(s,t) + \mathbb{Y}(s,t),$$

where $\mathbb{Y}$ is anti-symmetric. This means that if

$$\hat{G} = \left\{ 1 + v + \left(C + \frac{1}{2} v \otimes v \right) : \ v \in \mathbb{R}^\ell, \ C \in \mathbb{R}^{\ell \times \ell}, \ C^* + C = 0 \right\},$$

then the α-Hölder path $\mathbf{x}$ is weakly geometric iff it takes value in $\hat{G}$. Observe that $\hat{G}$ is a subgroup of G because the product

$$\left[1 + v + \left(C + \frac{1}{2} v \otimes v \right) \right] \otimes \left[1 + v' + \left(C' + \frac{1}{2} v' \otimes v' \right) \right],$$

equals

$$
\begin{aligned}
1+&(v+v') + \left(C + C' + v\otimes v' + \frac{1}{2}(v\otimes v + v'\otimes v')\right)\\
&= 1 + (v+v') + \left(C + C' + \frac{1}{2}\left(v\otimes v' - v'\otimes v\right) + \frac{1}{2}(v+v')\otimes(v+v')\right),
\end{aligned}
$$

and

$$\left[1 + v + \left(C + \frac{1}{2}v\otimes v\right)\right]^{-1} = \left[1 - v + \left(\frac{1}{2}v\otimes v - C\right)\right] \in \hat{G}.$$

2.1 Rough ODE

Given a rough path $\mathbf{x} = (x, \mathbb{X}) \in \mathcal{R}^\alpha$, we wish to solve (2.1) provided that σ is sufficiently smooth. This solution is unique and can be expressed as $y = \mathcal{S}(y(0), \mathbf{x})$ where $\mathcal{S}$ is a continuous function. The map $\mathcal{S}$ was constructed by Lyons as a deterministic analog of Itô's solution to SDE, and is known as Itô-Lyons map. Here is the precise statement of our main result of this subsection:

Theorem 2.3 *Assume that $\sigma \in C^3$, and that $D\sigma, D^2\sigma$ and $D^3\sigma$ are bounded. Given a rough path $\mathbf{x} = (x, \mathbb{X}) \in \mathcal{R}^\alpha$, with $\alpha \in (3^{-1}, 2^{-1}]$, $y^0 \in \mathbb{R}^d$, there exists a unique $\mathbf{y} = (y, \hat{y}) \in \mathcal{G}^\alpha(x)$ such that $\hat{y} = \sigma(y)$, and*

$$y(t) = y^0 + \int_0^t (\sigma(y), D\sigma(y)\sigma(y))\ d\mathbf{x}. \tag{2.18}$$

Moreover if $\mathcal{S}(y^0, \mathbf{x})$ is this solution, then there exists a constant $C_1 = C_1(T)$ such that

$$[\mathcal{S}(y^0, \mathbf{x}) - \mathcal{S}(y^1, \mathbf{x}')]_\alpha \le C_1 \left[|y^0 - y^1| + [\mathbf{x} - \mathbf{x}']_{\alpha, 2\alpha}\right]. \tag{2.19}$$

Proof *(Step 1)* We first construct the solution for short times, and a bootstrap would allow us to prove the existence in the interval $[0, T]$. This solution is constructed as a fixed point of an operator $\mathcal{F} = \mathcal{F}^1 \circ \mathcal{F}^0$, where $\mathcal{F}^0_x$, and $\mathcal{F}^1_{y^0, \mathbf{x}}$ are two continuous operators:

$$
\begin{aligned}
&\mathcal{F}^1_{y^0,\mathbf{x}} : \mathcal{G}^\alpha(x; [0, t_0]; \mathbb{R}^{d\times\ell}) \to \mathcal{G}^\alpha(x; [0, t_0]; \mathbb{R}^d),\\
&\mathcal{F}^0_x : \mathcal{G}^\alpha(x; [0, t_0]; \mathbb{R}^d) \to \mathcal{G}^\alpha(x; [0, t_0]; \mathbb{R}^{d\times\ell}).
\end{aligned}
$$

For $\mathcal{F}^1_{y^0,\mathbf{x}}$, we simply have $\mathcal{F}^1_{y^0,\mathbf{x}}(z, \hat{z}) = (w, z)$, where

$$w(t) = y^0 + \int_0^t \mathbf{z}\ d\mathbf{x}.$$

By (2.13), we already know that this linear operator is bounded with

$$[\mathcal{F}^1(\mathbf{z}) - \mathcal{F}^1(\mathbf{z}')]_{x;\alpha,2\alpha} \le (c_0(\alpha)[\mathbf{x}]_{\alpha,2\alpha} + T^\alpha[\mathbb{X}]_{2\alpha})\ [\mathbf{z}-\mathbf{z}']_{x;\alpha,2\alpha}. \tag{2.20}$$

The operator $\mathcal{F}^0_x$ is defined by $\mathcal{F}^0_x(y,\hat{y}) = (v,\hat{v}) := (\sigma(y), D\sigma(y)\hat{y})$ where $\hat{v} = [\hat{v}^{ijk}]$, with

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

Note

$$\begin{aligned}|\sigma(y(t)) - \sigma(y(s)) - \hat{v}(s)x(s,t)| =& |\sigma(y(t)) - \sigma(y(s)) - D\sigma(y(s)) : (\hat{y}(s)x(s,t))| \\ \le& |D\sigma(y(s)) : (y(s,t) - \hat{y}(s)x(s,t))| + \|D^2\sigma\|_{L^\infty}|y(s,t)|^2 \\ \le& \left[\|D\sigma\|_{L^\infty} + \|D^2\sigma\|_{L^\infty}\right]\left([\mathbf{y}]_{x;2\alpha} + [y]_\alpha^2\right)|s-t|^{2\alpha},\end{aligned}$$

which implies

$$[\mathbf{v}]_{x;\alpha,2\alpha} \le \left[2\|D\sigma\|_{L^\infty} + \|D^2\sigma\|_{L^\infty}\right]\left([\mathbf{z}]_{x;\alpha,2\alpha} + [\mathbf{z}]^2_{x;\alpha,2\alpha}\right). \tag{2.21}$$

(Step 2) Observe that if $\mathbf{a} = (a,\hat{a})$ with $a(t) = y^0 + \sigma(y^0)x(0,t),\ \hat{a}(t) = \sigma(y^0)$, then $\mathbf{a} \in \mathcal{G}^\alpha(x)$ with $[\mathbf{a}]_{x,\alpha,2\alpha} = 0$. Let us write $\widehat{\mathcal{G}}_{t_0}(x)$ for the set of $\mathbf{y} \in \mathcal{G}^\alpha(x;\mathbb{R}^d)$ such that $y(0) = y^0,\ \hat{y}(0) = \sigma(y^0)$. Note that $\mathbf{y} \in \widehat{\mathcal{G}}_{t_0}$, and $[\mathbf{y}]_{x;\alpha,2\alpha} = 0$ is equivalent to $\mathbf{y} = \mathbf{a}$. Clearly $\widehat{\mathcal{G}}_{t_0}$ is invariant under $\mathcal{F}$, and the semi-norm $[\cdot]_{x,\alpha,2\alpha}$ induces a metric D_α on $\widehat{\mathcal{G}}_{t_0}$, defined by

$$D_\alpha(\mathbf{y},\mathbf{y}') = [\mathbf{y}-\mathbf{y}']_{x,\alpha,2\alpha}.$$

We set

$$\mathcal{B}_\alpha(r) = \left\{\mathbf{y} \in \widehat{\mathcal{G}}_{t_0} :\ D_\alpha(\mathbf{y},\mathbf{a}) \le r\right\}.$$

Note

$$D_\alpha(\mathbf{y},\mathbf{a}) = [\hat{y}]_\alpha + [\mathbf{y}]_{x,2\alpha} = [\mathbf{y}]_{x,\alpha,2\alpha}.$$

Hence $[\mathbf{y}]_{x,\alpha,2\alpha} \le r$ for $\mathbf{y} \in \mathcal{B}_\alpha(r)$. From this, (2.18) and (2.20) we learn

$$D_\alpha(\mathcal{F}(\mathbf{y}),\mathbf{a}) \le (r + r^2)\left[2\|D\sigma\|_{L^\infty} + \|D^2\sigma\|_{L^\infty}\right](c_0(\alpha)[\mathbf{x}]_{\alpha,2\alpha} + t_0^\alpha[\mathbb{X}]_{2\alpha}).$$

If the right-hand side were at most r, we could have claimed that the set $\mathcal{B}_\alpha(r)$ is invariant under the map $\mathcal{F}$. Observe that since $\mathcal{C}^\alpha \subset \mathcal{C}^\beta$, for any $\beta < \alpha$, and

$$[\mathbf{x}]_{\beta,2\beta} \le [\mathbf{x}]_{\alpha,2\alpha}\ \max\{t_0^{\alpha-\beta}, t_0^{2\alpha-2\beta}\}, \tag{2.22}$$

we pick some $\beta \in (3^{-1},\alpha)$, so that the set $\mathcal{B}_\beta(1)$ is invariant under $\mathcal{F}$ for sufficiently small t_0. Note that the choice of t_0 depends on $\mathbf{x}$ and σ only.

(Step 3) We claim that if t_0 is sufficiently small, then $\mathcal{F}$ is a contraction on $\mathcal{B} = \mathcal{B}_\beta(1)$. For this, we first show that $\mathcal{F}^0$ is a Lipschitz map. Take $\mathbf{y} \in \widehat{\mathcal{G}}_{t_0}(x)$, $\mathbf{y}' \in \widehat{\mathcal{G}}_{t_0}(x')$, and put

$$\begin{aligned}
&z = y - y', \quad \hat{z} = \hat{y} - \hat{y}', \quad v = \sigma(y) - \sigma(y'), \quad k = D\sigma(y) - D\sigma(y'),\\
&h = D\sigma(y)\hat{y} - D\sigma(y')\hat{y}', \quad w = \int_0^1 D\sigma(\theta y + (1-\theta)y')\, d\theta\\
&\gamma(s,t) := y(s,t) - \hat{y}(s)x(s,t), \quad \gamma'(s,t) := y'(s,t) - \hat{y}'(s)x'(s,t).
\end{aligned}$$

Note that $v = wz$, and

$$\begin{aligned}
|v(s,t)| =& |(wz)(t) - (wz)(s)| \le |w(s,t)||z(t)| + |w(s)||z(s,t)|\\
&\le \left(\|D^2\sigma\|_{L^\infty} \max\{[y]_\beta, [y']_\beta\}\|y - y'\|_{L^\infty} + \|D\sigma\|_{L^\infty}[y - y']_\beta\right)|t - s|^\beta\\
&\le \left(\|D^2\sigma\|_{L^\infty} \max\{[y]_\beta, [y']_\beta\} t_0^\beta + \|D\sigma\|_{L^\infty}\right)[y - y']_\beta |t - s|^\beta,
\end{aligned} \tag{2.23}$$

which leads to the bound

$$[\sigma(y) - \sigma(y')]_\beta \le \left(\|D^2\sigma\|_{L^\infty} \max\{[y]_\beta, [y']_\beta\} t_0^\beta + \|D\sigma\|_{L^\infty}\right)[y - y']_\beta \le c_1[\sigma]_2[y - y']_\beta,$$

for any $y, y' \in \mathcal{B}$, where $[\sigma]_2 = \|D\sigma\|_{L^\infty} + \|D^2\sigma\|_{L^\infty}$. In the same fashion, we can show that there exists a constant c_2 such that

$$[k]_\beta = [D\sigma(y) - D\sigma(y')]_\beta \le c_2[\sigma]_3\ [y - y']_\beta,$$

for any $y, y' \in \mathcal{B}_\alpha$. Now write

$$h = k\hat{y} + D\sigma(y)\hat{z},$$

and we argue as in (2.23) to assert

$$\begin{aligned}
|(k\hat{y})(s,t)| &\le |k(s,t)||\hat{y}(t)| + |k(s)||\hat{y}(s,t)| \le 2[k]_\beta[\hat{y}]_\beta\ t_0^\beta\ |t - s|^\beta\\
&\le 2c_2[\sigma]_3[\hat{y}]_\beta\ t_0^\beta\ [y - y']_\beta\ |t - s|^\beta,\\
|(D\sigma(y)\hat{z})(s,t)| &\le |(D\sigma(y))(s,t)||\hat{z}(t)| + |D\sigma(y)(s)||\hat{z}(s,t)|\\
&\le \|D\sigma\|_\infty \left[[y]_\beta\ t_0^\beta + 1\right][\hat{z}]_\beta\ |t - s|^\beta.
\end{aligned}$$

From this we learn

$$[D\sigma(y)\hat{y} - D\sigma(y')\hat{y}']_\beta \le c_3[\sigma]_3\left([y - y']_\beta + [\hat{y} - \hat{y}']_\beta\right), \tag{2.24}$$

for a constant c_3.

(Step 4) Let us define

$$\eta(\theta) = \sigma(\theta y(t) + (1-\theta)y(s)) - \sigma(\theta y'(t) + (1-\theta)y'(s)),$$

so that

$$\eta(1)-\eta(0)-\eta'(0)=v(s,t)-D\sigma(y(s))y(s,t)+D\sigma(y'(s))y'(s,t)=:v(s,t)-\zeta(s,t),$$
$$\eta''(\theta)=D^2\sigma(\theta y(t)+(1-\theta)y(s))\ y(s,t)y(s,t)-D^2\sigma'(\theta y'(t)+(1-\theta)y'(s))\ y'(s,t)y'(s,t).$$

as in (2.23) we can readily show

$$|\eta''(\theta)|\le\Big[\|D^3\sigma\|_{L^\infty}\ t_0^\beta\ [y]_\beta^2+\|D^2\sigma\|_{L^\infty}([y]_\beta+[y']_\beta)\ \Big]\ [y-y']_\beta|t-s|^{2\beta}.$$

As a result

$$|v(s,t)-\zeta(s,t)|\le[\sigma]_3\Big[t_0^\beta\ [y]_\beta^2+[y]_\beta+[y']_\beta\Big]\ [y-y']_\beta \tag{2.25}$$

Moreover

$$\begin{aligned}&|\zeta(s,t)-D\sigma(y(s))\hat{y}(s)\ x(s,t)+D\sigma(y'(s))\hat{y}'(s)\ x'(s,t)|\\ &\quad=|D\sigma(y(s))y(s,t)-D\sigma(y'(s))y'(s,t)-D\sigma(y(s))\hat{y}(s)\ x(s,t)+D\sigma(y'(s))\hat{y}'(s)\ x'(s,t)|\\ &\quad=|D\sigma(y(s))\gamma(s,t)-D\sigma(y'(s))\gamma'(s,t)|\\ &\quad\le|(D\sigma(y(s))-D\sigma(y'(s))\gamma(s,t)|+|D\sigma(y'(s))(\gamma(s,t)-\gamma'(s,t))|\\ &\quad\le[\sigma]_2\left(\|y-y'\|_{L^\infty}[\mathbf{y}]_{x,2\beta}+[\mathbf{y}-\mathbf{y}']_{x,x',2\beta}\right)|t-s|^{2\beta}.\end{aligned}$$

In particular when for $x=x'$,

$$|\zeta(s,t)-h(s)x(s,t)|\le[\sigma]_2\left(\|y-y'\|_{L^\infty}[\mathbf{y}]_{x,2\beta}+[\mathbf{y}-\mathbf{y}']_{x,2\beta}\right)|t-s|^{2\beta}.$$

From this and (2.25) we learn,

$$[(v,h)]_{x,2\beta}\le[\sigma]_3\left(\left(t_0^\beta\ [\mathbf{y}]_{x,2\beta}+t_0^\beta\ [y]_\beta^2+[y]_\beta^2+[y']_\beta^2\right)[y-y']_\beta+[\mathbf{y}-\mathbf{y}']_{x,2\beta}\right),$$

From this, and (2.24) we deduce there exists a constant c_4 such that

$$[\mathcal{F}^0(\mathbf{y})-\mathcal{F}^0(\mathbf{y}')]_{x,\beta,2\beta}\le c_4[\sigma]_3[\mathbf{y}-\mathbf{y}']_{x,2\beta}, \tag{2.26}$$

for $\mathbf{y},\mathbf{y}'\in\mathcal{B}_\beta$.

(Step 5) From (2.26), (2.22), (2.20) we deduce that $\mathcal{F}$ is a contraction on $\mathcal{B}_\beta$ for t_0 sufficiently small. From this, we deduce the existence of a solution in $[0,t_0]$ in $\mathcal{G}^\beta(x)$. Since t_0 depends on σ and $\mathbf{x}$ only, we can apply or result to $[t_0,2t_0],\dots,$ to assert the existence of a global solution. The solution we have constructed is in $\mathcal{G}^\beta(x)$. We now use (2.18) to conclude that the solution is in $\mathcal{G}^\alpha(x)$ provided that $2\beta\ge\alpha$. This is an immediate consequence of (2.12).

(Step 6) We now turn to the proof of (2.19). Since the solution is the fixed point of $\mathcal{F}$, we need to study the stability of $\mathcal{F}$. We already know that $\mathcal{F}$ is locally Lipschitz with respect

to $\mathbf{y}$. Let us examine the Lipschitz regularity of $\mathcal{F}$ with respect to $\mathbf{x}$. We already know that $\mathcal{F}^1$ is Lipschitz in $(\mathbf{x}, \mathbf{y})$ by (2.13). The local Lipschitzness of $\mathcal{F}^0$ with respect to $(\mathbf{x}, \mathbf{y})$ we carried out in Step 4. □

Remark 2.1.1 Note that if $\sigma \in C^3$, we can still prove the existence of a solution for sufficiently small T. Note that for $\mathbf{y} \in \mathcal{B}_\beta$, we have $|y(s) - y^0| \le t_0^\beta$ for $s \in [0, t_0]$, we calls for a uniform bound of $\Delta\sigma,\ D^2\sigma,\ D^3\sigma$ on the set $\{y:\ |y - y^0| \le t_0^\beta$. □

2.2 Rough Paths of Low Regularity

As we saw before, when $\alpha \in (1/3, 12]$ and $x \in \mathcal{C}^\alpha$, then its lift $\mathbf{x}$ can be regarded as a path in the truncated vector algebra $T^{(2)}(\mathbb{R}^\ell) = \mathbb{R} \oplus \mathbb{R}^\ell \oplus \mathbb{R}^{\ell\times\ell}$. In this section, we show that when $\alpha \in (1/(n+1), 1/n)$, the lift of a geometric path takes value $T^{(n)}(\mathbb{R}^\ell)$. As we vary n, we will be dealing with the full tensor algebra $H = T(\mathbb{R}^d)$. In the case of non-geometric paths, we need to go beyond H as we will discuss later. We first argue that in order to represent the geometric property of a rough path, we need to equip the tensor algebra H with the *shuffle product* $\sqcup\!\sqcup$. We refer to Example C.3**(ii)** for a detailed discussion of shuffle algebra that turns H to a Hopf algebra. More precisely, if $I = \{1, \dots, \ell\}$, and $\{e_i, i \in I\}$ denotes the standard basis for $\mathbb{R}^\ell$, then $H = \oplus_{n\ge0} H_n$, where $H_n = T_n(V)$ is spanned by $\{e_a:\ a \in I^n\}$, with $I^0 = \{\emptyset\}$,

$$e_\emptyset = 1, \qquad e_{(i_1,\dots,i_n)} = e_{i_1} \otimes \dots \otimes e_{i_n}.$$

We think of I as the set of alphabet, and $a \in I^n$ as a word of length n. When there is no danger of confusion we write a for e_a. Also, we write $a_j = (i_1, \dots, i_j)$, $\hat{a}_j = (i_{j+1}, \dots, i_n)$, when $a = (i_1, \dots, i_n)$, and write $a\ell$ for $(i_1, \dots, i_n, \ell)$. The Hopf algebra $(H; \sqcup\!\sqcup, \mathbb{1}; \Delta, \mathbb{1}'; S)$ is equipped with the *shuffle product* $\sqcup\!\sqcup$, that is defined inductively by

$$a \sqcup\!\sqcup \emptyset = \emptyset \sqcup\!\sqcup a = a, \qquad (ak) \sqcup\!\sqcup (b\ell) = (a \sqcup\!\sqcup (b\ell))\, k + ((ak) \sqcup\!\sqcup b)\, \ell, \tag{2.27}$$

a coproduct $\Delta : H \to H \boxtimes H$, that is defined by

$$\Delta(a) = a \boxtimes \mathbb{1} + \mathbb{1} \boxtimes a + \sum_{k=1}^{n-1} a_k \boxtimes \hat{a}_k, \qquad a_k = (i_1, \dots, i_k), \quad \hat{a}_k = (i_{k+1}, \dots, i_n), \tag{2.28}$$

for $a = (i_1, \dots, i_n)$, the counit $\mathbb{1}'(e_a) = \delta_{a,\emptyset}$, and the antipode

$$S(v_1 \otimes \dots \otimes v_n) = (-1)^n S(v_n \otimes \dots \otimes v_1).$$

The Hopf algebra $(H; \sqcup\!\sqcup, \mathbb{1}; \Delta, \mathbb{1}'; S)$ has a dual $(H^*; \bullet, \mathbb{1}'^*; \sqcup\!\sqcup^*, \mathbb{1}^*; S^*)$, where $\bullet$ is Δ^*. This dual is also a Hopf algebra. We write

$$\langle f, h\rangle = f(h), \qquad f \in H^*, \quad h \in H,$$

for the pairing between H and H^*.

When $x : [0, t_0] \to \mathbb{R}^\ell$ is a smooth path, we lift it to a geometric $\mathbf{x} : [0, t_0]^2 \to H^*$ by

$$
\begin{aligned}
(2.29) \qquad & \langle \mathbf{x}(s,t), \emptyset \rangle = 1, \qquad \langle \mathbf{x}(s,t), i) \rangle = \int_s^t dx^i(\theta), \\
& \langle \mathbf{x}(s,t), (i_1, \dots, i_n) \rangle = \int_s^t \int_s^{\theta_n} \cdots \int_s^{\theta_2} dx^{i_1}(\theta_1) \dots dx^{i_n}(\theta_n),
\end{aligned}
$$

for $n \geq 2$. The space of geometric path is the completion smooth geometric paths as above with respect to the locally uniform convergence. The following result justifies the relevance of shuffle product.

Proposition 2.1 *Given a smooth path $x : [0, t_0] \to \mathbb{R}^\ell$, define $\mathbf{x}$ by (2.29). Then*

$$
\begin{aligned}
(2.30) \qquad & \langle \mathbf{x}(s,t), a \sqcup b \rangle = \langle \mathbf{x}(s,t), a \rangle \langle \mathbf{x}(s,t), b \rangle, \\
(2.31) \qquad & \mathbf{x}(s,t) = \mathbf{x}(s,u) \bullet \mathbf{x}(u,t),
\end{aligned}
$$

for every $u \in (s,t)$, and words a and b.

Proof Given $a = (i_1, \dots, i_m)$, and $b = (i_{m+1}, \dots, i_{m+n})$, the right-hand side of (2.30) can be written as

$$
\mathcal{I} := \int_C dx^{i_1}(\theta_1) \dots dx^{i_{m+n}}(\theta_{m+n}),
$$

where

$$
C = \{\theta = (\theta_1, \dots, \theta_{m+n}) : \; s < \theta_1 < \cdots < \theta_s < t, \quad s < \theta_{m+1} < \cdots < \theta_{m+n} < t\}.
$$

Let us write $Sh_{m,n}$ for the set of permutations σ of $\{1, \dots, m+n\}$, such that $\sigma^{-1}(1) < \cdots < \sigma^{-1}(m)$, and $\sigma^{-1}(m+1) < \cdots < \sigma^{-1}(m+n)$. Given $\theta \in C$, we can find a $\sigma \in Sh_{m,n}$ such that

$$
(2.32) \qquad \theta_{\sigma(1)} < \cdots < \theta_{\sigma(m+n)}.
$$

This suggests writing C_σ for the set of $\theta = (\theta_1, \dots, \theta_{m+n})$ such that (2.32) holds. Evidently,

$$
\begin{aligned}
\mathcal{I} &= \sum_{\sigma \in Sh_{m,n}} \int_{C_\sigma} dx^{i_1}(\theta_1) \dots dx^{i_{m+n}}(\theta_{m+n}) \\
&= \sum_{\sigma \in Sh_{m,n}} \langle \mathbf{x}(s,t), \big(i_{\sigma(1)}, \dots, i_{\sigma(m+n)}\big) \rangle = \sum_{c \in Sh(a,b)} \langle \mathbf{x}(s,t), c \rangle,
\end{aligned}
$$

completing the proof of (2.30).

To verify the proof of (2.32), pick $a = (i_1, \dots, i_n) \in I^n$, so that

$$\begin{aligned}
\langle \mathbf{x}(s,t), a \rangle &= \int_s^t \int_s^{\theta_n} \cdots \int_s^{\theta_2} dx^{i_1}(\theta_1) \dots dx^{i_n}(\theta_n) \\
&= \sum_{k=0}^{n} \int_s^{\theta_{n+1}} \cdots \int_s^{\theta_2} \mathbb{1}\left(\theta_k < u < \theta_{k+1}\right) \; dx^{i_1}(\theta_1) \dots dx^{i_n}(\theta_n) \\
&= \sum_{k=0}^{n} \int_u^{\theta_{n+1}} \cdots \int_u^{\theta_{k+2}} \int_s^{u} \int_s^{\theta_k} \cdots \int_s^{\theta_2} dx^{i_1}(\theta_1) \dots dx^{i_n}(\theta_n) \\
&= \sum_{k=0}^{n} \int_s^{u} \int_s^{\theta_k} \cdots \int_s^{\theta_2} dx^{i_1}(\theta_1) \dots dx^{i_k}(\theta_k) \int_u^{\theta_{n+1}} \cdots \int_u^{\theta_{k+2}} dx^{i_{k+1}}(\theta_{k+1}) \dots dx^{i_n}(\theta_n) \\
&= \sum_{k=0}^{n} \langle \mathbf{x}(s,u), a^k \rangle \langle \mathbf{x}(u,t), \hat{a}^k \rangle .
\end{aligned}$$

with the convention $\theta_0 = s, \; \theta_{n+1} = t$. As a result,

$$\langle \mathbf{x}(s,t), a \rangle = \langle \mathbf{x}(s,u) \boxtimes \mathbf{x}(u,t), \Delta a \rangle'$$

which implies (2.32). Moreover, $\bullet$ coincides with the tensor product.

We may identify H^* with $T(\mathbb{R}^{d*})$ with a basis consisting of e_a^*, where e_a^* is dual to e^a for each word a.WE now claim $e_a \bullet e_b = e_a \otimes e_b = e_{ab}$ for every pair of words a and b. For this, observe

$$\begin{aligned}
\langle e_a^* \bullet e_b^*, e_c \rangle = \langle e_a^* \boxtimes e_b^*, \Delta e_c \rangle = \sum_{i=0}^{|c|} \langle e_a^* \boxtimes e_b^*, e_{c_i} \boxtimes e_{\hat{c}_i} \rangle = \sum_{i=0}^{|c|} \langle e_a^*, e_{c_i} \rangle \langle e_b^*, e_{\hat{c}_i} \rangle \\
= \mathbb{1}\left(a = c_i, b = \hat{c}_i \quad \text{for some } i\right) = \mathbb{1}(ab = c) = \langle e_{ab}^*, e_c \rangle,
\end{aligned}$$

as desired. □

From Proposition 2.1, we learn that geometric rough paths should take value in the space of characters $G(H) \subset H^*$, which is a group by Proposition C.1 of Appendix C. Moreover, since

$$\mathbf{x}(s,t) = \mathbf{x}(0,s)^{-1} \bullet \mathbf{x}(0,t),$$

we only need to study Hölder continuous $\mathbf{x} : [0, t_0] \to G(H)$. We write $\mathcal{C}^\alpha([0,t_0]; G(H))$ for the set of such $\mathbf{x}$ such that

$$[\langle \mathbf{x}, a \rangle] = \sup_{s \neq t} \frac{|\langle \mathbf{x}(s,t), a \rangle|}{|t - s|^{\alpha |a|}} < \infty,$$

for every word a.

For non-geometric paths, Gubinelli [G2] discovered that the tensor algebra $T(\mathbb{R}^\ell)$ must be replaced with the Hopf algebra $\mathcal{H}$ of Connes-Kreimer [CK] that was used in the context of renormaliztion theory. More precisely, $\mathcal{H}$ is the space of polynomials with the set of labeled rooted trees T serving as the commuting indeterminants, and the product $\cdot$ simply given by the polynomial product.

2.3 Stochastic Sewing Lemma

Definition 2.3.1 Let $(\Omega, \widehat{\mathcal{F}}, \mathbb{P})$ be a probability measure, and assume that $\widehat{\mathcal{F}} = (\mathcal{F}_t : \ t \in [0,T])$ is a filtration, i.e., each $\mathcal{F}_t$ is a complete σ-algebra, $\mathcal{F}_s \subseteq \mathcal{F}_t$, for $s \le t$, and $\mathcal{F}_0$ is the trivial σ-algebra. Let $A : \Omega \times [0,T]^2 \to \mathbb{R}$ be a measurable function, such that $(s,t) \mapsto A(\omega,s,t)$ is a continuous function for every ω. Given $\gamma_1, \gamma_2, C_1, C_2 > 0$, and $r \ge 1$, we say that A is $(\gamma_1,\gamma_2;r;C_1,C_2)$-*coherent* if the following conditions are true:

- When $s \le t$, the function $A(\cdot,s,t)$ is $\mathcal{F}_t$-measurable.
- If $0 \le s \le u \le t \le T$, and $A(\omega,s,u,t) = A(\omega,s,t) - A(\omega,s,u) - A(\omega,u,t)$, then

  $$\mathbb{E}|A(s,u,t)|^r \le C_1^r|t-s|^{r\gamma_1}, \quad \mathbb{E}\left|\mathbb{E}^s A(s,u,t)\right|^r \le C_2^r|t-s|^{r\gamma_2}. \tag{2.33}$$

  where
  $$\mathbb{E}^s A := \mathbb{E}(A \mid \mathcal{F}_s).$$

□

The following *Stochastic Sewing Lemma* of Le guarantees the existence of a Hölder continuous function h that can be approximated by a coherent A.

Theorem 2.4 *Assume that A is $(\gamma_1,\gamma_2;r;C_1,C_2)$-coherent for some $\gamma_1 > 1/2$ and $\gamma_2 > 1$. Let $\{\pi_n : \ n \in \mathbb{N}_0\}$ be the partition*

$$\pi_n = \left\{t_i^n = i2^{-n}t : \ i = 1,\dots,2^n\right\}.$$

Then

$$h(t) = \lim_{n\to\infty} \sum_{i=0}^{n} A(t_i^n, t_{i+1}^n). \tag{2.34}$$

exists in $L^r(\mathbb{P})$. Moreover, there satisfies a constant C_3 such that

$$|h(t) - h(s)| \le C_3 C_1 |t-s|^{\gamma_1} + C_3' C_2 |t-s|^{\gamma_1}, \tag{2.35}$$

where

$$C_3(\gamma_1,r) = c_0(r)(1 - 2^{-(\gamma_1 - 1/2)}, \quad C_3'(\gamma_2,r) = c_0(r)(1 - 2^{-\gamma_2}.$$

Proof Let π and t_i be as in Theorem 2.1, and set

$$I_n = \sum_{i=0}^{2^n-1} A(t_i, t_{i+1}).$$

Then

$$\begin{aligned}
\|I_{n+1} - I_n\|_{L^r} =& \left\| \sum_{i=0}^{2^n-1} A(t_{i-1}^n, t_{2i+1}, t_{i+1}^n) \right\|_{L^r} \\
\leq& \left\| \sum_{i=0}^{2^n-1} \left(A(t_{i-1}^n, t_{2i+1}, t_{i+1}^n) - \mathbb{E}^{t_i^n} A(t_{i-1}^n, t_{2i+1}^{n+1}, t_{i+1}^n) \right) \right\|_{L^r} \\
&+ \left\| \sum_{i=0}^{2^n-1} \mathbb{E}^{t_i^n} A(t_{i-1}^n, t_{2i+1}^{n+1}, t_{i+1}^n) \right\|_{L^r} \\
\leq& c_0(r) \left(\sum_{i=0}^{2^n-1} \left\| \left(A(t_{i-1}^n, t_{2i+1}^{n+1}, t_{i+1}^n) - \mathbb{E}^{t_i^n} A(t_{i-1}^n, t_{2i+1}^{n+1}, t_{i+1}^n) \right) \right\|_{L^r}^2 \right)^{1/2} \\
&+ \sum_{i=0}^{2^n-1} \left\| \mathbb{E}^{t_i^n} A(t_{i-1}^n, t_{2i+1}^{n+1}, t_{i+1}^n) \right\|_{L^r} \\
\leq& c_0(r) \left[C_1 2^{n/2} t^{\gamma_1} 2^{-n\gamma_1} + C_2 2^n t^{\gamma_2} 2^{-n\gamma_2} \right],
\end{aligned}$$

where we used BDG inequality for the second inequality. As a result,

$$\sum_{n=0}^{\infty} \|I_n\|_{L^r} < \infty,$$

which implies the existence of the limit in (2.34). We then use

$$h(t) = \sum_{n=0}^{\infty} (I_{n+1} - I_n),$$

to prove (2.35) when $s = 0$. The proof of general s is similar. $\square$

Example 2.3.1 Let B be a d-dimensional standard Brownian motion, let $X : \mathbb{R}^d \to \mathbb{R}^d$ be a vector field in C^γ. Consider

$$F(s,t) := \int_s^t \left(X(B(\theta) + x) - X(B(\theta) + y) \right) \, d\theta, \qquad A(s,t) := \mathbb{E}^s F(s,t),$$

for $s \le t$. In other words, if β is a standard Brownian motion, and

$$E(z,\tau) = \mathbb{E}\int_0^\tau \big(X(z+x+\beta(\theta)) - X(z+y+\beta(\theta))\big)\, d\theta,$$

then $A(s,t) = E(B(s), t-s)$. Writing P_θ for the heat semigroup, we have,

$$\begin{aligned}|E(z,\tau)| =& \left|\int_0^\tau \big[(P_\theta X)(z+x) - (P_\theta X)(z+y)\big]\, d\theta\right| \\ \le& |x-y| \int_0^\tau \|\nabla(P_\theta X)\|_{L^\infty}\, d\theta \le c|x-y|\tau^{(\gamma+1)/2}\|X\|_{C^\gamma},\end{aligned}$$

by parabolic Schauder estimate. Hence

$$|A(s,t)| \le c|x-y|(t-s)^{(\gamma+1)/2}\|X\|_{C^\gamma}. \tag{2.36}$$

Note

$$A(s,u,t) = \big(\mathbb{E}^s - \mathbb{E}^u\big)\int_u^t \big(X(B(\theta)+x) - X(B(\theta)+y)\big)\, d\theta.$$

Clearly

$$\mathbb{E}^s A(s,u,t) = 0.$$

Moreover, by (2.36),

$$|A(s,u,t)| \le 3c|x-y|(t-s)^{(\gamma+1)/2}\|X\|_{C^\gamma}.$$

Hence, A is $\big((\gamma+1)/2, 0, r\big)$-coherent for any $r \ge 1$. On the other hand, if $\Pi = \{s = t_0 \le t_1 \le \dots \tau_n = t\}$ is any mesh of $[s,t]$, then

$$\begin{aligned}\left\|F(s,t) - \sum_{i=0}^{n-1} A(t_i,t_{i+1})\right\|_{L^2(\mathbb{P})}^2 =& \left\|\sum_{i=0}^{n-1}\big(F(t_i,t_{i+1}) - \mathbb{E}^{t_i}F(t_i,t_{i+1})\big)\right\|_{L^2(\mathbb{P})}^2 \\ =& \sum_{i=0}^{n-1}\big\|F(t_i,t_{i+1}) - \mathbb{E}^{t_i}F(t_i,t_{i+1})\big\|_{L^2(\mathbb{P})}^2 \\ \le& \|X\|_{L^\infty}^2 \sum_{i=0}^{n-1}(t_{i+1}-t_i)^2 \le \|X\|_{L^\infty}^2 \max_i(t_{i+1}-t_i)\ (t-s).\end{aligned}$$

Hence

$$F(s,t) = \lim_{\|\Pi\|\to 0}\sum_{i=0}^{n-1} A(t_i,t_{i+1}),$$

in $L^2(\mathbb{P})$ sense. We may use Theorem 2.4 to assert that for a C^γ vector field X, the process

$$f(x,t) = \int_0^t X(x + B(\theta))\, d\theta,$$

satisfies

$$\big[\mathbb{E}|f(x,t) - f(y,s)|^r\big]^{1/r} \leq c(r)\|X\|_{C^\gamma}|x-y|\ |t-s|^{(\gamma+1)/2}. \tag{2.37}$$

By Kolmogorov's theorem (see Theorem 3.1 in Chapter 3), (2.37) implies that f is in C^{1-} with respect to x, and f is in $C^{(\gamma+1)/2-}$ with respect to t. □

Exercise(i) Assume that $0 < \alpha < \beta \leq 1$. Consider the Banach space $(\mathcal{C}^\alpha, \|\cdot\|_\alpha)$. Show that the closure of $\mathcal{C}^\beta$ in this Banach space is exactly the set $\widehat{\mathcal{C}}^\alpha$, which consists of $x : [0,T] \to \mathbb{R}^d$ such that

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

(ii) Show that if $\mathbf{x}_n \in \mathcal{C}^\alpha$ converges uniformly to $\mathbf{x}$, and $\sup_n \|\mathbf{x}_n\|_{\alpha,2\alpha} < \infty$, then $\mathbf{x}_n$ converges in $\mathcal{C}^\beta$ for any $\beta < \alpha$.

(iii) Pick $\alpha \in (0,1]$. Let $f : \mathbb{R} \to \mathbb{R}^d$ be a 1-periodic Lipschitz function and define

$$x_n(t) = n^{-\alpha} f(nt),$$

for $t \in [0,1]$, and $n \in \mathbb{N}$. Show

$$\sup_{n\in\mathbb{N}} \|x_n\|_\alpha < \infty.$$

(iv) Let x_n be as in the previous problem and set

$$\mathbb{X}_n(s,t) = \int_s^t (x_n(\theta) - x_n(s)) \otimes x_n(d\theta),$$

where the integral is the standard Riemann integral. Show that the sequence $\mathbf{x}_n := (x_n, \mathbb{X}_n)$ converges when $\alpha \geq 1/2$. Determine the uniform limit

$$(x, \mathbb{X}) = \lim_{n\to\infty} (x_n, \mathbb{X}_n).$$

Show

$$\sup_n \|\mathbf{x}_n\|_{\frac{1}{2},1} < \infty,$$

when $\alpha = 1/2$. From this deduce that $\mathbf{x}_n$ converges in $\|\cdot\|_{\beta,2\beta}$ for every $\beta < 1/2$.

(v) Assume $y : [0, T] \to \mathbb{R}^d$ is C^1, and define $\hat{x}_n = x_n + y$, and

$$\widehat{\mathbb{X}}_n(s,t) = \int_s^t (\hat{x}_n(\theta) - \hat{x}_n(s)) \otimes \hat{x}_n(d\theta),$$

with x_n as in the previous problem. Determine $\hat{\mathbf{x}}$, the large n limit of $(\hat{x}_n, \widehat{\mathbb{X}}_n)$ in $\mathcal{C}^\alpha$ with $\alpha \in (1/3, 1/2)$. Evaluate

$$\int F(\hat{\mathbf{x}}) \cdot d\hat{\mathbf{x}}.$$

(vi) Let $f : \mathbb{R}^2 \to \mathbb{R}^d$ be a C^1 function with $f(s,t)$ 1-periodic in the second variable, and define $x_n(t) = n^{-1/2} f(t, nt)$ for $t \in [0,1]$. Define $\mathbb{X}_n$ as in part **(iv)**. Show

$$C(s,t) = [C^{ij}]_{i,j=1}^d := \lim_{n\to\infty} \mathbb{X}_n(s,t),$$

exists with

$$C^{ij}(s,t) = \int_s^t \left[\int_0^1 f^i(\theta_1, \theta_2) f^j_{\theta_2}(\theta_1, \theta_2)\, d\theta_2 \right] d\theta_1.$$

□

3 SDE

To solve (2.1) for a randomly selected x such as Brownian motion, we need to come up with an approxiamtion scheme to produce a candidate for $\mathbb{X}$. For such random x, we often have some information about its marginals and we need to learn how to use such information to control $\mathbf{x} = (x, \mathbb{X})$ in a suitable rough path space $\mathcal{R}^\alpha$. For example if

$$A(x) = \int_0^T \int_0^T \Psi\left(\frac{|x(t) - x(s)|}{p(|t-s|)}\right) dsdt < \infty,$$

for increasing functions $\Psi, p : [0, \infty) \to [0, \infty)$, with $\Psi(0) = p(0) = 0$ and $\Psi(\infty) = \infty$, then by the celebrated *Garsia–Rodemich–Rumsey inequality*,

$$|x(t) - x(s)| \leq 8 \int_0^{|t-s|} \Psi^{-1}\left(\frac{4A(x)}{\theta^2}\right) p(d\theta). \tag{3.1}$$

Note that if $\mathbb{E}$ denotes the expected value with respect to the randomness of x, then $\mathbb{E}A(x) < \infty$ guarantees (3.1) for almost all realizations of x. On the other hand, the validity of $\mathbb{E}A(x) < \infty$ can be checked if we have some control on the 2-dimensional marginals of x.

In particular the choices of $\Psi(a) = a^q$, and $p(a) = a^{\alpha + \frac{1}{q}}$ lead to

$$|x(t) - x(s)| \leq c_0(q, \alpha) B(x)^{\frac{1}{q}} \; |t - s|^{\alpha - \frac{1}{q}},$$

where

$$B(x) = \int_0^T \int_0^T \frac{|x(t) - x(s)|^q}{|t-s|^{q\alpha+1}} \; dsdt.$$

Note that $\mathbb{E}B(x) < \infty$ if there exists $\varepsilon > 0$ and c such that

$$[\mathbb{E}|x(t) - x(s)|^q]^{\frac{1}{q}} \leq c|t-s|^{\alpha+\varepsilon}$$

for all $s, t \in [0, T]$. This Hölder continuity of x is the celebrated Kolmogorov's continuity theorem. We now formulated and prove a generalization of this result for rough paths.

Theorem 3.1 *Let $q \geq 2$ and $\beta > q^{-1}$. Assume $\mathbb{P}$ is a probability measure on the set of measurable maps $x : [0, T] \to \mathbb{R}^\ell$, $\mathbb{X} : [0, T]^2 \to \mathbb{R}^{\ell \times \ell}$ such that (2.5) holds. If*

$$\mathcal{N}_{\beta,q}(\mathbf{x}) := \sup_{s \neq t} \mathbb{E}\left[\frac{|x(s,t)| + \sqrt{|\mathbb{X}(s,t)|}}{|t-s|^\beta}\right]^q < \infty$$

then there exists a continuos version of $\mathbf{x} = (x, \mathbb{X})$ such that

$$\mathbb{E}[\mathbf{x}]_{\gamma,2\gamma}^q < \infty,$$

for every $\gamma \in (0, \beta - q^{-1})$.

Proof Withous loss of generality, we assume that $T = 1$. Put

$$D_n = \{i2^{-n} : \; i = 0, \dots, 2^n\}, \qquad D = \cup_{n=0}^{\infty} D_n.$$

Define

$$C_n(x) = \sup_{t \in D_n} \left| x(t + 2^{-n}) - x(t) \right|, \qquad \mathbb{C}_n(\mathbb{X}) = \sup_{t \in D_n} \left| \mathbb{X}(t, t + 2^{-n}) \right|.$$

Observe

$$\begin{aligned} \mathbb{E}\left[|C_n(x)| + \sqrt{|\mathbb{C}_n(\mathbb{X})|} \right]^q \leq & \sum_{t \in D_n} \mathbb{E}\left[|x(t, t+2^{-n})| + |\mathbb{X}(t, t+2^{-n})|^{1/2} \right]^q \\ \leq & 2^{n(1-\beta q)} \mathcal{N}_{\beta,q}(\mathbf{x}) \end{aligned} \tag{3.2}$$

We wish to bound $[\mathbf{x}]_{\gamma,2\gamma}$ in terms of $C_n(x)$ and $\mathbb{C}_n(\mathbb{X})$. For this we pick $s, t \in D$ with $s < t$, and choose $m \in \mathbb{N}$ so that $2^{-(m+1)} < t - s \leq 2^{-m}$. For sure, there exists a unique $\theta \in (s, t) \cap D_{m+1}$. We use the binary expansions of $t - \theta$ and $\theta - s$ to write

$$t - \theta = 2^{-m_1} + \cdots + 2^{-m_r}, \qquad \theta - s = 2^{-m'_1} + \cdots + 2^{-m'_{r'}},$$

such that $m + 1 \leq m_1 < m_2 < \cdots < m_r$, $m + 1 \leq m'_1 < m'_2 < \cdots < m_{r'}$. If we define two finite sequences

$$s_0 = t_0 = \theta, \qquad s_{i+1} = s_i - 2^{-m_i}, \qquad t_{i+1} = t_i + 2^{-m'_i},$$

inductively, then

$$\begin{aligned} |x(t) - x(s)| \leq & |x(t_{r+1} - x(t_r)| + \cdots + |x(t_1) - x(t_0)| \\ & + |x(s_0) - x(s_1)| + \cdots + |x(s_{r'}) - x(s_{r'+1})| \\ \leq & \sum_{i=1}^{r} C_{m_i} + \sum_{i=1}^{r'} C_{m'_i} \leq 2 \sum_{n=m+1}^{\infty} C_n. \end{aligned}$$

As a result,

$$\frac{|x(t) - x(s)|}{|t - s|^\gamma} \leq 2 \; 2^{(m+1)\gamma} \sum_{n=m+1}^{\infty} C_n \leq 2 \sum_{n=m+1}^{\infty} 2^{n\gamma} C_n \leq 2 \sum_{n=1}^{\infty} 2^{n\gamma} C_n.$$

This yields

$$\sup_{s,t \in D, s \neq t} \frac{|x(t) - x(s)|}{|t - s|^\gamma} \leq 2 \sum_{n=1}^{\infty} 2^{n\gamma} C_n(x).$$

In the same fashion,

$$\begin{aligned}|\mathbb{X}(s,t)| \leq & \sum_{i=0}^{r} |\mathbb{X}(t_i,t_{i+1})| + \left[\sum_{i=0}^{r} |x(t_i,t_{i+1})|\right]^2 \\ & + \sum_{i=0}^{r'} |\mathbb{X}(s_i,s_{i+1})| + \left[\sum_{i=0}^{r'} |x(s_i,s_{i+1})|\right]^2 \\ \leq & 2\sum_{n=m+1}^{\infty} \mathbb{C}_n + 2\left[\sum_{n=m+1}^{\infty} C_n\right]^2,\end{aligned}$$

which yields

$$\begin{aligned}\sup_{s,t\in D, s\neq t} \frac{\sqrt{|\mathbb{X}(s,t)|}}{|t-s|^\gamma} \leq & \left[2\sum_{n=1}^{\infty} 2^{2n\gamma}\mathbb{C}_n(\mathbb{X})\right]^{1/2} + \sqrt{2}\sum_{n=1}^{\infty} 2^{n\gamma} C_n(x) \\ \leq & 2\sqrt{2}\sum_{n=1}^{\infty} 2^{n\gamma}\left[C_n(x) + \sqrt{\mathbb{C}_n(\mathbb{X})}\right].\end{aligned}$$

From this and (3.2) we deduce

$$\begin{aligned}\left\| \sup_{s,t\in D, s\neq t} \frac{|x(s,t)| + \sqrt{|\mathbb{X}(s,t)|}}{|t-s|^\gamma} \right\|_{L^q(\mathbb{P})} \leq & 2(\sqrt{2}+1)\left\| \sum_{n=1}^{\infty} 2^{n\gamma}\left[C_n(x) + \sqrt{\mathbb{C}_n(\mathbb{X})}\right]\right\|_{L^q(\mathbb{P})} \\ \leq & 2(\sqrt{2}+1)\sum_{n=1}^{\infty} 2^{n\gamma}\left\| C_n(x) + \sqrt{\mathbb{C}_n(\mathbb{X})}\right\|_{L^q(\mathbb{P})} \\ \leq & 2(\sqrt{2}+1)\sum_{n=1}^{\infty} 2^{-n(\beta-q^{-1}-\gamma)}\, \mathcal{N}_q(\mathbf{x})^{1/q},\end{aligned}$$

which is finite if $\gamma < \beta - q^{-1}$. This yields the desired Hölder regularity, if we replace $[0,1]$ with D. We can then extend $\mathbf{x}$ to $[0,1]$ by

$$x'(t) = \lim_{t_n \to t} x(t_n), \qquad \mathbb{X}'(s,t) = \lim_{(s_n,t_n)\to(s,t)} \mathbb{X}(s_n,t_n),$$

where $s_n, t_n \in D$. Note that now $\mathbf{x}' = (x', \mathbb{X}') \in \mathcal{R}^\gamma$, and

$$\mathbb{E}|x(t)-x'(t)|^q = \mathbb{E}\liminf_{n\to\infty} |x(t)-x(t_n)|^q \leq \liminf_{n\to\infty} \mathbb{E}|x(t)-x(t_n)|^q \leq \liminf_{n\to\infty} |t-t_n|^{\beta q}\mathcal{N}_{\beta,q}(\mathbf{x}) = 0.$$

Hence $x = x'$ almost everywhere. The same reasoning yields $\mathbb{X} = \mathbb{X}'$ a.e. □

Given a stochastic x, we may design a scheme to approximate $\mathbb{X}$ and use Theorem 3.1 to verify the convergence of our approximation in a suitable $\mathcal{R}^\gamma$. Depending on the type of randomness we have we may appeal to different techniques to bound $\mathcal{N}_{\beta,q}(\mathbf{x})$. For example if x is a centered Gaussian process, a mild regularity of $D_x(s,t) := (\mathbb{E}|x(s,t)|^2)^{1/2}$ would guarantees the existence of nice rough paths $\mathbf{x} = (x, \mathbb{X}) \in \mathcal{R}^\alpha$. We remark that it takes a very mild regularity of D_x to guarantee the continuity of x (see Exercise **(i)** below). Though for a Hölder regularity, it suffices to have

$$D_x(s,t) \le c_0|t-s|^{-\tau},$$

for some $\tau > 0$, because for every $n \in \mathbb{N}$,

$$\mathbb{E}|x(s,t)|^{2n} = \frac{(2n)!}{2^n n!}\left[\mathbb{E}|x(s,t)|^2\right]^n. \tag{3.3}$$

This and Theorem 3.1 imply that $x \in \mathcal{C}^\alpha$ for every $\alpha \in (0,\tau)$. As an example, consider the fractional Brownian motion of *Hurst index* τ that is defined as a centered Gaussian process with $x(0) = 0$, and the correlation

$$\mathbb{E}x(t)x(s) = \frac{1}{2}\left(t^{2\tau} + s^{2\tau} - |t-s|^{2\tau}\right).$$

We now study two classical examples of rough paths associated with a Brownian motion. When $\tau = 1/2$, our Gaussian process is a standard Brownian motion. Let us write B for the standard ℓ-dimensional Brownian motion. We also write $t_i^n = i2^{-n}$ for the dyadic points. The following approximation schemes lead to Itô and Stratonovich integrals respectively:

$$\begin{aligned} \mathbb{B}(s,t) &= \lim_{n\to\infty} \sum_{i:t_i^n\in[s,t)} B(t_i^n) \otimes B(t_i^n, t_{i+1}^n) - B(s) \otimes B(s,t), \\ \widehat{\mathbb{B}}(s,t) &= \lim_{n\to\infty} \sum_{i:t_i^n\in[s,t)} \frac{B(t_i^n) + B(t_{i+1}^n)}{2} \otimes B(t_i^n, t_{i+1}^n) - B(s) \otimes B(s,t), \end{aligned} \tag{3.4}$$

where $B(s,t) = B(t) - B(s)$.

Theorem 3.2 *The limits in* (3.4) *exist in* $L^2(\mathbb{P})$, *where* $\mathbb{P}$ *denotes the Wiener measure, and*

$$\widehat{\mathbb{B}}(s,t) = \mathbb{B}(s,t) + (t-s)I/2. \tag{3.5}$$

Moreover

$$\sup_{s\ne t} \frac{|B(s,t)| + \sqrt{|\mathbb{B}(s,t)|}}{|t-s|^\alpha} \in L^q(\mathbb{P}), \tag{3.6}$$

for every $q \ge 1$, *and every* $\alpha \in (0,1/2)$.

Proof *(Step 1)* Set

$$\mathbb{B}_n(s,t) = \sum_{i:t_i^n\in[s,t)} B(t_i^n)\otimes B(t_i^n,t_{i+1}^n),$$
$$\widehat{\mathbb{B}}_n(s,t) = \sum_{i:t_i^n\in[s,t)} \frac{B(t_i^n)+B(t_{i+1}^n)}{2}\otimes B(t_i^n,t_{i+1}^n).$$

Then

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

From this and $|a\otimes b|^2=|a|^2|b|^2$ we learn,

$$\begin{aligned}\mathbb{E}|\mathbb{B}_{n+1}(s,t)-\mathbb{B}_n(s,t)|^2 &= \sum_{i:t_i^n\in[s,t)} \mathbb{E}|B(t_i^n,t_{2i+1}^{n+1})|^2|B(t_{2i+1}^{n+1},t_{i+1}^n)|^2\\ &= \sum_{i:t_i^n\in[s,t)} (t_{2i+1}^{n+1}-t_i^n)(t_{i+1}^n-t_{2i+1}^{n+1})\approx (t-s)2^{-n},\end{aligned}$$

which implies the $L^2(\mathbb{P})$-convergence of $\mathbb{B}_n(s,t)$. We may regard $\{\mathbb{B}_n\|_{n\in\mathbb{N}}$ as Cauchy in $L^2(m\times\mathbb{P})$, where $m(ds,dt)=ds\ dt$ is the Lebesgue measure on $[0,T]^2$. As a result $\mathbb{B}$ and $\widehat{\mathbb{B}}$ are well-defined as functions measurable functions of (s,t,B) in $L^2(m\times\mathbb{P})$. We can easily check that the Chen's relation holds for both $\mathbb{B}$ and $\widehat{\mathbb{B}}$).

As for (3.5), observe

$$\mathbb{E}\left[\widehat{\mathbb{B}}_n(s,t)-\mathbb{B}_n(s,t)\right] = \frac{1}{2}\sum_{i:t_i^n\in[s,t)}(t_{i+1}^n-t_i^n)I\approx\frac{1}{2}(t-s)I.$$

On the other hand, we can readily show

$$\lim_{n\to\infty}\mathbb{E}\left|\sum_{i:t_i^n\in[s,t)}\big(B(t_i^n,t_{i+1}^n)\otimes B(t_i^n,t_{i+1}^n)-(t_{i+1}^n-t_i^n)I\big)\right|^2=0, \tag{3.7}$$

which implies (3.5). The proof of (3.6) is left to the reader (see Exercise **(iv)**).

(Step 2) We now turn our attention to the question of the regularity of $\mathbf{B}=(B,\mathbb{B})$. Fix $\beta\in(0,1/2)$. From $\mathbb{E}|B(s,t)|^2=\ell|t-s|$, (3.3), and Theorem 3.1, we learn that the random variable

$$C(B)=\sup_{s\neq t}\frac{|B(s,t)|}{|t-s|^\beta},$$

is in $L^q(\mathbb{P})$ for every $q \in [1,\infty)$. On the other hand if $t_i^n \in (s,t)$ iff $i = m, m+1, \dots, r$, and we write $X_i = B_j(s, t_i^n)$, then we may use the independence of X_i and $B_k(t_n^i, t_n^{i+1})$ to argue

$$\mathbb{E}\ e^{\sum_{i=m}^{r-1}\left(\lambda X_i B_k(t_n^i,t_n^{i+1}) - \frac{\lambda^2}{2}X_i^2(t_n^{i+1}-t_n^i)\right)} = \mathbb{E}\ e^{\sum_{i=m}^{r-2}\left(\lambda X_i B_k(t_n^i,t_n^{i+1}) - \frac{\lambda^2}{2}X_i^2(t_n^{i+1}-t_n^i)\right)} = \cdots = 1,$$

for every $\lambda \in \mathbb{R}$ such that $\lambda^2 2^n < 1$ (we need this condition to make sure that the expected value is finite). Write

$$Y_n = \sum_{i=m}^{r-1} X_i B_k(t_n^i, t_n^{i+1}), \qquad Z_n = \sum_{i=m}^{r-1} X_i^2 (t_n^{i+1} - t_n^i),$$

so that

$$1 = \mathbb{E}\ e^{\lambda Y_n - \lambda^2 Z_n/2} =: \sum_{m=0}^{\infty} \frac{t^m}{m!} K_m(Y_n, Z_n),$$

where $K_m(Y,Z)$ is a polynomial of degree m in (Y,Z). We may express these polynomials in terms of *Hermite* polynomials. To see this, write

$$e^{ta - t^2/2} =: \sum_{m=0}^{\infty} \frac{t^m}{m!} He_m(a),$$

and recall $He_{m+1}(a) = aHe_m(a) - mHe_{m-1}(a)$, which in particular implies that $He_m(0) = 0$ when m is odd, and He_m is even (odd) when m is even (odd). Hence,

$$K_m(Y,Z) = Ha_m\left(\frac{Y}{Z^{1/2}}\right) Z^{m/2}.$$

In particular

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

After an application of a weighted Schwartz inequality

$$Y^{2(m-j)} Z^j \le \left(1 - \frac{j}{m}\right)(\varepsilon Y^{2(m-j)})^{\frac{m}{m-j}} + \frac{j}{m}(Z^j/\varepsilon)^{\frac{m}{j}} = \left(1 - \frac{j}{m}\right)\varepsilon^{\frac{m}{m-j}} Y^{2m} + \frac{j}{m\varepsilon^{\frac{m}{j}}} Z^m,$$

we can write

$$\mathbb{E} Y_n^{2m} \le c_m \mathbb{E} Z_n^m,$$

for a constant c_m. From this and

$$|Z_n| \le C(B)^2 |t-s|^{2\beta+1},$$

we deduce

$$\mathbb{E} Y_n^{2m} \le c'_m |t-s|^{2m\beta+m}, \tag{3.8}$$

for a constant c'_m. We may send $n \to \infty$ to obtain analogous bound for $\mathcal{B}$. This and Theorem 3.1 imply that there is version of $\mathbb{B}$ that is continuous, and that (3.6) holds for $\alpha \in (0, 2^{-1}\beta + 4^{-1} - (4m)^{-1})$. This complete the proof because we can choose β close to $1/2$, and m large. □

Remark 3.1 Our calculation in the second step can be used to assert that the process

$$N(t) = e^{\int_0^t f(B(s))\ d\mathbb{B}(s) - \frac{1}{2}\int_0^t |f(B(s))|^2\ ds}, \tag{3.9}$$

is a martingale with respect to the σ-field $\mathcal{F}_t$ that is generated from $(B(s):\ s \in [0,t])$. Also, if $M(t) = \int_0^t f(B(s))\ d\mathbb{B}(s)$, then $M(t)$ is a martingale with the quadratic variation

$$\langle M\rangle(t) = \int_0^t |f(B(s))|^2\ ds.$$

From the exponential martingale (3.9) we learn that $K_m(M, \langle M\rangle)$ is also a martingale for each m. We may use the celebrated the Burkholder-Davis-Gundy inequality

$$\mathbb{E} \sup_{s\in[0,t]} |M(s)|^q \le c_q \mathbb{E}\langle M\rangle(t)^{q/2},$$

to establish (3.8). □

From Theorem 3.1, we now have two candidates $\mathbf{B}$ and $\widehat{\mathbf{B}}$ in $\mathcal{R}^\alpha(B)$. From the definition, it is clear

$$\widehat{\mathbb{B}}(s,t) + \widehat{\mathbb{B}}(s,t)^* = \lim_{n\to\infty} \sum_{t_n^i \in [s,t)} \left(B(t_{i+1}^n) \otimes B(t_{i+1}^n) - B(t_i^n) \otimes B(t_i^n)\right) = B(s,t) \otimes B(s,t),$$

which means that $\widehat{\mathbf{B}}$ is geometric. Because of this, we expect that if we replace B with a reasonable smooth approximation $B^{(n)}$, and use Riemann integration for its lift, then the corresponding $\mathbf{B}^{(n)}$ converges to $\widehat{\mathbf{B}}$. We first consider a linear interpolation;

$$B^{(n)}(t) = \sum_i \left[\frac{t - t_i^n}{t_{i+1}^n - t_n^n} B(t_{i+1}^n) + \frac{t_{i+1}^n - t}{t_{i+1}^n - t_i^n} B(t_i^n)\right] 1\!\!1 \left(t \in (t_i^n, t_{i+1}^n)\right).$$

We then define $\mathbf{B}^{(n)} = (B^{(n)}, \mathbb{B}^{(n)})$, where

$$\mathbb{B}^{(n)}(s,t) = \int_s^t B^{(n)}(\theta) \otimes \dot{B}^{(n)}(\theta)\ d\theta.$$

The following is a rather straightforward consequence of Theorem 3.2.

Theorem 3.3 *For every $\alpha \in (0,1/2)$,*

$$\lim_{n\to\infty} d_\alpha(\mathbf{B}^{(n)}, \widehat{\mathbf{B}}) = 0, \tag{3.10}$$

almost surely, where d_α is the distance associated with $[\cdot]_{\alpha,2\alpha}$.

Proof Observe,

$$\dot{B}^{(n)}(t) = \sum_i \frac{B(t_i^n, t_{i+1}^n)}{t_{i+1}^n - t_i^n} \; 1\!\!1\left(t \in (t_n^i, t_n^{i+1})\right),$$

which implies

$$\int_{t_i^n}^{t_{i+1}^n} B^{(n)} \dot{B}^{(n)} \; dt = (t_{i+1}^n - t_i^n) \frac{B(t_i^n) + B(t_{i+1}^n)}{2} \otimes \frac{B(t_i^n, t_{i+1}^n)}{t_{i+1}^n - t_i^n}.$$

Hence,

$$\int_s^t B^{(n)} \dot{B}^{(n)} \; dt = \widehat{\mathbb{B}}_n(t_i^n, t_{i+1}^n).$$

(This gives an alternative proof for the geometric property of $\widehat{\mathbf{B}}$.) By Theorem 3.2 we already know that $\mathbf{B}^{(n)}(s,t) \to \widehat{\mathbf{B}}(s,t)$ in $L^2(\mathbb{P})$ as $n \to \infty$ for every $(s,t) \in D$. Also the proof of Theorem 3.2 guarantees

$$\sup_n \left\| \left[\mathbf{B}^{(n)}\right]_{\alpha,2\alpha} \right\|_{L^q(\mathbb{P})} < \infty$$

for every $q \geq 1$ and $\alpha \in (0,1/2)$. As a result, $\mathbf{B}^{(n)} \to \widehat{\mathbf{B}}$ uniformly, almost surely. We then use Exercise **(iii)** of Chapter 2 to deduce (3.10). □

Remark 3.2 If we write $\mathcal{F}^n$ for the σ-algebra generated by $(B(t_i^n : \; t_i^n \in [0,T])$, then we can show that indeed

$$\mathbf{B}^{(n)} = \mathbb{E}\left(\mathbf{B} \mid \mathcal{F}^n\right). \tag{3.11}$$

To see this, observe that if $0 < s < t$, then $X := \mathbb{E}(B(s) \mid B(t))$ must be a constant multiple of $B(t)$ because it is Gaussian by Exercise **(v)**, and is centered because $\mathbb{E}B(s) = 0$. Since $\mathbb{E}XB(t) = \mathbb{E}(B(s)B(t)) = s$, we must have $X = st^{-1}B(t)$. As a result, for $s \in (t_1, t_2)$,

$$\begin{aligned} \mathbb{E}(B(s) \mid B(t_1), B(t_2)) &= \mathbb{E}(B(s) \mid B(t_1), B(t_1,t_2)) = B(t_1) + \mathbb{E}(B(t_1,s) \mid B(t_1), B(t_1,t_2)) \\ &= B(t_1) + \mathbb{E}(B(t_1,s) \mid B(t_1,t_2)) = B(t_1) + \frac{s-t_1}{t_2-t_1} B(t_1,t_2) \\ &= \frac{t_2-s}{t_2-t_1} B(t_1) + \frac{s-t_1}{t_2-t_1} B(t_2). \end{aligned}$$

From this, we can readily deduce

$$B^{(n)} = \mathbb{E}\left(B \mid \mathcal{F}^n\right).$$

On the other hand if $i \neq j$, then B_i and B_j are independent, and

$$\mathbb{E}(B_i(s)B_j(s,t) \mid \mathcal{F}^n) = \mathbb{E}(B_i(s) \mid \mathcal{F}^n)\mathbb{E}(B_j(s,t) \mid \mathcal{F}^n) = B_i^{(n)}(s)B_j^{(n)}(s,t).$$

This implies (3.11) because we already have an exact formula for the diagonal entries of $\mathbb{B}^{(n)}$ which is compatible with (**??**). Note that (3.11) implies the convergence $\mathbf{B}^{(n)} \to \widehat{\mathbf{B}}$ by Doob's martingale convergence theorem. □

Exercises(i) Recall that if X is a centered Gaussian random variable, then

$$\mathbb{E}e^{tX} = e^{t^2\mathbb{E}X^2/2}.$$

Use this to show (3.3). Also use this to deduce that if $x : [0,T] \to \mathbb{R}$ is a centered Gaussian process with

$$\left(\mathbb{E}[x(t)-x(s)]^2\right)^{1/2} \leq c_0 \left(\log \frac{1}{|t-s|}\right)^{-\alpha},$$

for some $\alpha > 1/2$, then x has a continuous sample path. (*Hint:* Use $\Psi(a) = e^{ca} - 1$ and $p(a) = |\log a|^{-\alpha-1/2}$ in (3.1).)

(ii) Show that almost surely,

$$\sup_{s,t\in[0,1],0<|s-t|<1/2} \frac{|B(s,t)|}{\sqrt{|t-s||\log|t-s||}} < \infty.$$

(*Hint:* Use $\Psi(a) = e^{a^2/4} - 1$ and $p(a) = a^{1/2}$ in (3.1).)

(iii) Use (3.1) to show that there exist constants c_1 and c_2 such that if $\omega(\delta) = \sqrt{\delta}|\log\delta|$ and

$$S(x) = \sup_{s,t\in[0,1],0<|s-t|<1/2} \frac{|x(t)-x(s)|}{\omega(|t-s|)}, \qquad A(x) = \int_0^1\int_0^1 \exp\left\{\frac{|x(t)-x(s)|}{\sqrt{|t-s|}}\right\}\, dsdt,$$

then

$$S(x) \leq c_1 + c_2 \log^+ A(x).$$

Use this to show that $\mathbb{E}e^{\lambda S(B)} < \infty$ for small $\lambda > 0$.

(iv) Verify (3.7).

(v) Let $(X,Y) \in \mathbb{R}^d \times \mathbb{R}^{d'}$ be a Gaussian random variable with density

$$(2\pi)^{-(d+d')/2}(\det A)^{1/2}e^{-\frac{1}{2}A_{11}x\cdot x - A_{12}x\cdot y - \frac{1}{2}y\cdot y}\, dxdy,$$

where the matrix

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix},$$

positive definite. Show that $\mathbb{E}(X|Y)$ is again Gaussian associated with that matrix

$$B = A_{11} - A_{12}A_{22}^{-1}A_{12}^*.$$

4 Reconstruction Theorem and Regularity Structure

Recall that if $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$, for some $\alpha, \beta \in (0,1)$, then according to Lyons and Victoir [LV], there exists a function h such that (1.6) holds. We may express this in terms of the distribution $T = h'$,

$$\big|(T - F_t)(\varphi_t^\delta)\big| \preceq \delta^\gamma, \tag{4.1}$$

where F_t is the distribution $f(t)\dot{g}$, and $\gamma = \alpha + \beta - 1$. The bound (4.1) is uniform over t in a bounded set, and the test functions $\varphi \in \mathcal{D}$ such that $|\varphi| \le 1$ and the support of φ is contained in a fixed interval, say $(-1,1)$. We proved (4.1) when $\gamma > 0$ in Chapter 2. We also observed that the family $F = (F_t : \ t \in [0,T])$ enjoys a regularity in the form of (2.16). In this Chapter, we will learn how to prove the existence of T satisfying (4.1) even when $\gamma \le 0$, and extend it to higher dimensions. More precisely, given $f \in \mathcal{C}^\alpha(\mathbb{R}^d), \ g \in \mathcal{C}^\beta(\mathbb{R}^d)$, we wish to come up with a candidate T for the distribution $f\nabla g$ satisfying

$$\big|(T - F_x)(\varphi_x^\delta)\big| \preceq \delta^\gamma, \tag{4.2}$$

where $F = (F_x : \ x \in \mathbb{R}^d)$, with $F_x = f(x)\nabla g$, and the inequality is uniform for x in a bounded set, $\delta \in (0,1]$, and $\varphi \in \mathcal{D}_0$ (See Definition A.1**(iii)** in the Appendix). But first we need to discuss the analog of (2.16) in higher dimensions. As a warm-up, let us work out an example, where our candidate for F_x is a polynomial.

Example 4.1 Assume that $u \in \mathcal{C}^\gamma_{loc}$ for some $\gamma > 0$, and put

$$P_a(x) = \sum_{|k| < \gamma} \frac{\partial^k u(a)}{k!}(x-a)^k. \tag{4.3}$$

We certainly have $u(x) = P_a(x) + R_0(x,a)$, where

$$|R_0(x,a)| \preceq |x-a|^\gamma,$$

holds locally uniformly. This certainly implies

$$\big|\langle (u - P_x), \varphi_x^\delta \rangle\big| \preceq \delta^\gamma. \tag{4.4}$$

On the other hand, since

$$\partial^k u(b) = \sum_{|m| < \gamma - |k|} \frac{\partial^{k+m} u(a)}{m!}(b-a)^m + R_k(a,b),$$

with R satisfying

$$|R(a,b)| \preceq |a-b|^{\gamma - |k|},$$

we have

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

In particular

$$\left|\langle P_a - P_b, \varphi_b^\delta\rangle\right| \preceq \sum_{|k|<\gamma} |a-b|^{\gamma-|k|}\delta^k \preceq (|a-b|+\delta)^\gamma. \tag{4.5}$$

□

Given a $\mathcal{D}'$-valued function F, we wish to find a distribution T that is well-approximated locally by F as in (4.2). Naturally, we may wonder what regularity/consistency condition on F would guarantee the existence of T. We now formulate such a condition that generalizes (2.16) and (4.5).

Definition 4.1(i) By a *germ*, we mean a measurable map $F : \mathbb{R}^d \to \mathcal{D}'$. We also write F_x for $F(x)$.

(ii) Let $\gamma \in \mathbb{R}$, and $\tau = \big(\tau_K :\ K \text{ compact subset of } \mathbb{R}^d\big)$, be a collection of non-negative numbers such that $\gamma + \tau_K \geq 0$, for every K. We say that a germ F is $(-\tau, \gamma)$-*coherent* if there exists $\varphi \in \mathcal{D}$ such that $\int \varphi\, dx \neq 0$, and for very compact set K,

$$\left|(F_x - F_y)(\varphi_y^\delta)\right| \preceq \delta^{-\tau_K}(|x-y|+\delta)^{\gamma+\tau_K}, \tag{4.6}$$

uniformly for $x, y \in K$, and $\delta \in (0,1]$. We say F is γ-coherent, if F is $(-\tau,\gamma)$-coherent for some τ. The set of such germs is denoted by $\mathcal{CG}_\gamma(\mathbb{R}^d)$. We also define

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

□

Example 4.2(i) Let $f \in \mathcal{C}^\alpha(\mathbb{R}^d)$, $g \in \mathcal{C}^\beta(\mathbb{R}^d)$, with $\alpha, \beta \in (0,1)$, and define $F_x = f(x)\nabla g$. Note

$$\begin{aligned}\left|\langle \nabla g, \varphi_y^\delta\rangle\right| &= \left|\langle \nabla(g-g(y)), \varphi_y^\delta\rangle\right| = \left|\langle g - g(y), div\ \varphi_y^\delta\rangle\right|\\ &=\delta^{-1}\left|\langle g-g(y), (div\ \varphi)_y^\delta\rangle\right| \leq [g]_\beta \|\varphi\|_{C^1}\ \delta^{-1+\beta},\end{aligned}$$

which in turn implies

$$\left|\langle F_x - F_y, \varphi_y^\delta\rangle\right| = \leq [f]_\alpha [g]_\beta \|\varphi\|_{C^1}\ \delta^{-1+\beta}|x-y|^\alpha \leq [f]_\alpha [g]_\beta \|\varphi\|_{C^1}\ \delta^{-1}(|x-y|+\delta)^{\gamma+1},$$

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

(ii) For $u \in \mathcal{C}^\alpha_{loc}$, $\alpha > 0$, the germ P with P_a as in (4.3) is $(0,\gamma)$-coherent by (4.5). □

We now argue that if (4.2) holds locally uniformly in x, and uniformly in

$$\mathcal{D}_r = \left\{\varphi \in \mathcal{D} :\ \varphi(x) = 0 \ \text{ for } \ x \notin B_1,\ \|\varphi\|_{C^r} \leq 1\right\},$$

then F must be γ-coherent. The following result is due to Caravenna and Zambotti [CZ].

Proposition 4.1 *Let F be a germ, and suppose that there exist $\gamma \in \mathbb{R}$, $T \in \mathcal{D}'$ and a constant C such that*

$$\left|(T - F_x)(\varphi_x^\delta)\right| \leq C\delta^\gamma, \tag{4.7}$$

for every $x \in K$, $\delta \in (0,1]$, and $\varphi \in \mathcal{D}_r$. Then

$$\left|(F_x - F_y)(\varphi_y^\delta)\right| \leq 2C\delta^{-\tau}(|x-y|+\delta)^{\gamma+\tau}, \tag{4.8}$$

for every $\delta \in (0,1/2]$, every $\varphi \in \mathcal{D}_r$, and every $x, y \in K$ with $|x-y| \leq 1/2$, where $\tau = d + r$.

Proof Assume that (4.7) holds. Pick any $x, y \in K$, $\delta \in (0,1/2]$, and $\varphi \in \mathcal{D}_r$. We certainly have

$$\begin{aligned}\left|(F_x - F_y)(\varphi_y^\delta)\right| &\leq \left|(F_x - T)(\varphi_y^\delta)\right| + \left|(T - F_y)(\varphi_y^\delta)\right| \leq \left|(F_x - T)(\varphi_y^\delta)\right| + C\delta^\gamma \\ &\leq \left|(F_x - T)(\varphi_y^\delta)\right| + C\delta^{-\tau}(|x-y|+\delta)^{\gamma+\tau}.\end{aligned}$$

It remains to show

$$\left|(F_x - T)(\varphi_y^\delta)\right| \leq C\delta^{-\tau}(|x-y|+\delta)^{\gamma+\tau}. \tag{4.9}$$

To use (4.7), we find $\psi \in \mathcal{D}$ and $\varepsilon \in (0,1]$ so that $\varphi_y^\delta = \psi_x^\varepsilon$. Indeed

$$\begin{aligned}\varphi_y^\delta(z) =& \delta^{-d}\ \varphi\left(\frac{z-y}{\delta}\right) = \delta^{-d}\ \varphi\left(\frac{(z-x)-(y-x)}{\delta}\right) = \varepsilon^{-d}(\varepsilon')^{-d}\ \varphi\left(\frac{(z-x)-a\varepsilon}{\varepsilon\varepsilon'}\right) \\ =& \varepsilon^{-d}(\varepsilon')^{-d}\ \varphi\left(\frac{\varepsilon^{-1}(z-x)-a}{\varepsilon'}\right) = \varepsilon^{-d}\ \varphi_a^{\varepsilon'}\left(\frac{z-x}{\varepsilon}\right) = \psi_x^\varepsilon(x),\end{aligned}$$

where

$$\varepsilon = |y-x|+\delta, \qquad \varepsilon' = \frac{\delta}{|y-x|+\delta}, \qquad a = \frac{y-x}{\varepsilon}, \qquad \psi = \varphi_a^{\varepsilon'}.$$

Observe the support of ψ is contained in $B_a(\varepsilon') \subset B_{|a|+\varepsilon'}(0)$, and $|a|+\varepsilon' = 1$, which allows us to use (4.7) to assert

$$\left|(F_x - T)(\varphi_y^\delta)\right| = |(F_x - T)(\psi_x^\varepsilon)| \le C\varepsilon^\gamma \|\psi\|_{C^r}.$$

This and the elementary bound

$$\|\psi\|_{C^r} \le (\varepsilon')^{-d-r}\|\varphi\|_{C^r},$$

yields (4.9). □

If a distribution T is of order $r = r_K$ in a compact set K, then we have the bound

$$|T(\varphi_x^\delta)\| \preceq \|\varphi_x^\delta\|_{C^r} \le \delta^{-d-r}\|\varphi\|_{C^r}, \tag{4.10}$$

whenever the support of φ_x^δ is contained in K. We now argue that for a coherent distribution, we have a similar bound for F_x locally uniformly in x.

Proposition 4.2 *Suppose that* (4.6) *holds, and set* $\beta = \beta_K = \max\{\tau_K, r_{K'}+d\}$, *where*

$$K' = \{x : \ |x-y| \le 1 \ \ \textit{for some } y \in K\}.$$

Then

$$|F_x(\varphi_x^\delta)| \preceq \delta^{-\beta}, \tag{4.11}$$

uniformly for $x \in K$, *and* $\delta \in (0,1]$.

Proof Fix $a \in K$ and let $r = r_{a,K}$ be the order of F_a in K'. Then by (4.10),

$$|F_a(\varphi_x^\delta)| \preceq \delta^{-r-d},$$

uniformly for $x \in K$, and $\delta \in (0,1]$. We then have

$$|F_x(\varphi_x^\delta)| \le |(F_x - F_a)(\varphi_x^\delta)| + |F_a(\varphi_x^\delta)| \preceq \delta^{-\tau_K}(|x-a|+\delta)^{\gamma+\tau} + \delta^{-r-d} \preceq \delta^{-\tau_K} + \delta^{-r-d},$$

uniformly for $x \in K$, and $\delta \in (0,1]$. This completes the proof of (4.11). □

We now state Reconstruction Theorem of Hairer.

Theorem 4.1 *For each $\gamma \in \mathbb{R}$, there exists a linear continuous operator $\mathcal{T}_\gamma : \mathcal{CG}_\gamma \to \mathcal{D}'$ such that*

$$\big|(\mathcal{T}_\gamma(F) - F_x)(\psi_x^\delta)\big| \preceq [F]_{K,\varphi,\tau,\gamma} \begin{cases} \delta^\gamma & \gamma \neq 0, \\ |\log \delta| & \gamma = 0, \end{cases} \tag{4.12}$$

uniformly for $x \in K$, $\delta \in (0,1]$, and $\psi \in \mathcal{D}_r$. The operator $\mathcal{T}_\gamma$ is unique when $\gamma > 0$.

Proof *(Step 1)* We first give a recipe for the operator $\mathcal{T}_\gamma$. For a compact notation, we write $\hat\rho_x^n$ for $\rho_x^{2^{-n}}$, and set $\tilde\rho^n(x) := \hat\rho^n(-x)$. Note that if $\rho \in \mathcal{D}$ with $\int \rho = 1$, then we always have $T * \hat\rho^n \to T$ as $n \to \infty$ (see (A.6) of Appendix A). In the support of $\hat\rho_x^n$, we should be able to replace T with F_x. Motivated by this, we define

$$\mathcal{T}_\gamma^{(n)}(F)(x) := T_n(x) := (F_x * \tilde\rho^n)(x) = F_x(\hat\rho_x^n),$$

which is a measurable function for each $n \in \mathbb{N}$. Here, the function ρ is a suitable test function that is related to φ of (4.6). When $\gamma > 0$, we define

$$\mathcal{T}_\gamma(F)(\psi) := T(\psi) := \lim_{n\to\infty} \langle \mathcal{T}_\gamma^{(n)}(F), \psi\rangle = \lim_{n\to\infty} \langle T_n, \psi\rangle. \tag{4.13}$$

Observe (see (A.6) below)

$$F_x(\psi) = \lim_{n\to\infty} (F_x * \tilde\rho^n)(\psi) =: \lim_{n\to\infty} \langle G_x^{(n)}, \psi\rangle = \lim_{n\to\infty} F_x(\hat\rho^n * \psi),$$

where $G_x^{(n)} = F_x * \tilde\rho^n$. Conveniently, we may write

$$T_n = T_1 + \sum_{k=1}^{n-1} (T_{k+1} - T_k), \qquad G_x^{(n)} = G_x^{(1)} + \sum_{k=1}^{n-1} (G_x^{(k+1)} - G_x^{(k)}), \tag{4.14}$$

which is convergent if we find a exponential decay bound on $T_{k+1} - T_k$. Our choice of ρ is written as $\rho = \eta * \varphi$ so that

$$(T_{k+1} - T_k)(y) = F_y(\hat\rho_y^{k+1} - \hat\rho_y^k) = F_y(\hat m_y^k), \qquad (G_x^{(k+1)} - G_x^{(k)})(y) = F_x(\hat m_y^k), \tag{4.15}$$

where

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

Observe that if $\eta = \varphi^2$, then

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

Also observe

$$
\begin{aligned}
(T_{k+1}-T_k)(y) =& F_y(\hat\varphi^k * \hat\zeta^k_y) = \int F_y(\hat\varphi^k_z)\ \hat\zeta^k_y(z)\ dz \\
=& \int F_z(\hat\varphi^k_z)\ \hat\zeta^k_y(z)\ dz + \int (F_y - F_z)(\hat\varphi^k_z)\ \hat\zeta^k_y(z)\ dz \\
=& : A_k(y) + B_k(y), \\
(G^{(k+1)}_x - G^{(k)}_x)(y) =& \int F_x(\hat\varphi^k_z)\ \hat\zeta^k_y(z)\ dz =: C_k(y).
\end{aligned}
$$

This allows us to write $T_n = S_n + U_n$, where

$$
S_n = T_1 + \sum_{k=1}^{n-1} A_k, \qquad U_n = \sum_{k=1}^{n-1} B_k. \tag{4.16}
$$

We define

$$
\mathcal{S}_\gamma(\psi) := \langle S, \psi\rangle = \lim_{n\to\infty}\langle S_n, \psi\rangle = T_1(\psi) + \sum_{k=1}^{\infty}\langle A_k, \psi\rangle. \tag{4.17}
$$

When $\gamma \le 0$, our $\mathcal{T}_\gamma$ is simply defined as $\mathcal{T}_\gamma = S$.

(Step 2) In this step, we focus on the A_k sequence. To show the existence of the limit in (4.17), we need to assume that for some $r \in \mathbb{N}$,

$$
\int \varphi(x) x^k\ dx = 0, \qquad \text{for } 0 < |k| < r. \tag{4.18}
$$

Since $\zeta = \varphi^{1/2} - \varphi^2$, we also have $\int \zeta = 0$, which in turn implies

$$
\int \zeta(x) P(x)\ dx = 0, \tag{4.19}
$$

for every polynomial P of degree at most $r-1$. As a consequence,

$$
|(\hat\zeta^k * \psi)(x)| = \left| \int \hat\zeta^k(z-z')\left(\psi(z') - P^{(r)}_z(z')\right)\ dz'\right| \preceq 2^{-rk}, \tag{4.20}
$$

where $P^{(r)}_z$ is the Taylor polynomial of degree $r-1$ of ψ at z. From this and (4.11) we learn

$$
\begin{aligned}
|\langle A_n, \psi\rangle| &= \left|\iint F_z(\hat\varphi^k_z)\ \hat\zeta^k_y(z)\psi(y)\ dzdy\right| = \left|\int F_z(\hat\varphi^k_z)\ (\hat\zeta^k * \psi)(z)\ dz\right| \\
&\preceq \sup_{z\in K} |F_z(\hat\varphi^k_z)|\ 2^{-rk} \preceq 2^{(\beta_K - r)k}.
\end{aligned}
$$

This yields the existence of the limit in (4.17) provided that $r > \beta_K$.

(Step 2) So far we know that S exists. For (4.12), we need to study $(S - F_x)(\psi_x^\delta)$, which can be expressed as a sum of terms of the form $A_k - C_k$. Observe

$$\begin{aligned}\left|\langle A_k - C_k, \psi_x^\delta\rangle\right| &= \left|\int (F_z - F_x)(\hat{\varphi}_z^k)\ \hat{\zeta}_y^k(z)\psi_x^\delta(y)\ dzdy\right| \\ &= \left|\int (F_z - F_x)(\hat{\varphi}_z^k)\ (\hat{\zeta}^k * \psi_x^\delta)(y)\ dy\right| \\ &\le \sup_{z\in B_{2^{-n}+\delta}(x)} |(F_z - F_x)(\hat{\varphi}_z^k)| \int \left|(\hat{\zeta}^k * \psi_x^\delta)(y)\right|\ dy \\ &\preceq 2^{\tau k}(\delta + 2^{-k+1})^{\gamma+\tau} \int |\hat{\zeta}^k * \psi_x^\delta)(y)|\ dy.\end{aligned}$$

Clearly, $\|\hat{\zeta}^k * \psi_x^\delta\|_{L^1} \preceq 1$, which yields

$$\left|\langle A_k - C_k, \psi_x^\delta\rangle\right| \preceq 2^{\tau k}(\delta + 2^{-k+1})^{\gamma+\tau} \preceq 2^{-k\gamma}, \tag{4.21}$$

whenever $\delta \le 2^{-k}$. The bound $\|\hat{\zeta}^k * \psi_x^\delta\|_{L^1} \preceq 1$ can be improved when $2^{-k} \le \delta$; as in (4.20),

$$\begin{aligned}\int |\hat{\zeta}^k * \psi_x^\delta)(y)|\ dy = &\int_{B_x(\delta+2^{-n})} |\hat{\zeta}^k * \psi_x^\delta)(y)|\ dy = \int_{B_{\delta+2^{-n}}(x)} \left|\int \hat{\zeta}^k(y - z')(\psi_x^\delta - \hat{P}_y^{(r)})(z'))\right|\ dy \\ &\preceq \|\psi_x^\delta\|_{C^r}\ 2^{-kr}|B_x(\delta + 2^{-k})| \preceq \delta^{-r-d}\ 2^{-kr}\ (\delta + 2^{-k})^d \preceq \delta^{-r}2^{-kr},\end{aligned}$$

whenever $2^{-k} \le \delta$. As a result,

$$\left|\langle A_k - C_k, \psi_x^\delta\rangle\right| \preceq 2^{\tau k}(\delta + 2^{-k+1})^{\gamma+\tau}\delta^{-r}2^{-kr} \preceq \delta^{\gamma+\tau-r}2^{(\tau-r)k},$$

whenever $2^{-k} \le \delta$. Hence, for $r > \tau$,

$$\sum_{2^{-k}\le\delta} \left|\langle A_k - C_k, \psi_x^\delta\rangle\right| \preceq \delta^\gamma. \tag{4.22}$$

On the other hand, when $\gamma \le 0$, we may use (4.21) to assert

$$\sum_{2^{-k}\ge\delta} \left|\langle A_k - C_k, \psi_x^\delta\rangle\right| \preceq \sum_{2^{-k}\ge\delta} 2^{-k\gamma} \preceq \begin{cases}\delta^\gamma & \gamma < 0, \\ |\log\delta| & \gamma = 0.\end{cases} \tag{4.23}$$

Note

$$\begin{aligned}|\langle T_1 - G_x^{(1)}, \psi_x^\delta\rangle| &= \left|\int (F_y - F_x)(\hat{\rho}_y^1)\psi_x^\delta(y)\ dy\right| = \left|\iint (F_y - F_x)(\hat{\varphi}_z^1)\hat{\eta}_y^1(z)\psi_x^\delta(y)\ dydz\right| \\ &\preceq \left[(|x - z| + 2^{-1})^{\gamma+\tau} + (|z - y| + 2^{-1})^{\gamma+\tau}\right] \left|\iint |\hat{\eta}_y^1(z)\psi_x^\delta(y)|\ dydz\right| \preceq 1.\end{aligned}$$

From this, (4.23), (4.22), (4.17), and (4.15) we deduce (4.12) when $\gamma \le 0$.

(Step 3) We now assume that $\gamma > 0$. We certainly have,

$$\begin{aligned}\left|\langle B_k, \psi_x^\delta\rangle\right| =& \left|\iint (F_y - F_z)(\hat\varphi_z^k)\ \hat\zeta_y^k(z)\psi_x^\delta(y)\ dzdy\right| \\ \preceq& \iint 2^{k\tau}(|y-z| + 2^{-n})^{\gamma+\tau}\ \left|\hat\zeta_y^k(z)\psi_x^\delta(y)\right|\ dzdy \\ \preceq& 2^{-k\gamma}\|\zeta\|_{L^1}\|\psi\|_{L^1}.\end{aligned}$$

From this (by choosing $x = 0$ and $\delta = 1$) it is not hard to deduce that $\sum_k B_k$ is convergent in $\mathcal{D}'$. Moreover,

$$\sum_{2^{-k}\le\delta} \left|\langle B_k, \psi_x^\delta\rangle\right| \preceq \sum_{2^{-k}\le\delta} 2^{-k\gamma} \preceq \delta^\gamma.$$

From this and (4.22) we learn

$$\left|[(T - F_x) - (T_n - G_x^{(n)})](\psi_x^\delta)\right| \preceq \delta^\gamma, \tag{4.24}$$

provide $2^{-n} \le \delta \le 2^{-n+1}$. It remains to verify

$$\left|\langle T_n - G_x^{(n)}, \psi_x^\delta\rangle\right| \preceq \delta^\gamma, \tag{4.25}$$

for such n. Indeed

$$\begin{aligned}\left|\langle T_n - G_x^{(n)}, \psi_x^\delta\rangle\right| =& \left|\int (F_y - F_x)(\hat\rho_y^n)\ \psi_x^\delta(y)\ dy\right| = \left|\iint (F_y - F_x)(\hat\varphi_z^n)\ \hat\eta_y^n(z)\ \psi_x^\delta(y)\ dydz\right| \\ =& \left|\iint [(F_y - F_z) + (F_z - F_x)](\hat\varphi_z^n)\ \hat\eta_y^n(z)\ \psi_x^\delta(y)\ dydz\right| \\ \preceq& \left|\iint 2^{n\tau}\left[(|z-y| + 2^{-n})^{\gamma+\tau} + (|x-y| + 2^{-n})^{\gamma+\tau}\right]\left|\hat\eta_y^n(z)\ \psi_x^\delta(y)\right|\ dydz\right| \\ \preceq& \ 2^{-n\gamma}\|\eta^n * \psi_x^\delta\|_{L^1} \preceq \delta^\gamma,\end{aligned}$$

because 2^{-n} and δ are comparable. This completes the proof of (4.25). From this and (4.24) we deduce (4.11) when $\gamma > 0$.

(Step 4) So far we have established (4.11) provided that φ satisfies (4.18). If we start from $\varphi \in \mathcal{D}$ with in support in B_1 and $\int \varphi = 1$, then we can build a suitable $\bar\varphi$ that is built from φ and satisfies (4.18). This construction is done in such a way that if φ satisfies (4.6), then $\bar\varphi$ also satisfies (4.6) provided that τ_K is modified. Indeed given distinct positive numbers $\lambda_0, \dots, \lambda_{r-1}$, we select constants $c_0, \dots, c_{r-1}$, so that a desired $\bar\varphi$ can be expressed as

$$\bar\varphi = \sum_{i=0}^{r-1} c_i \varphi^{\lambda_i}.$$

Indeed from the calculation

$$\int x^k \bar{\varphi}(x)\ dx = \sum_{i=0}^{r-1} c_i \int x^k \varphi^{\lambda_i}(x)\ dx = \sum_{i=0}^{r-1} c_i \lambda_i^{|k|},$$

we learn that $c_0, \dots, c_{r-1}$ must satisfy,

$$\sum_{i=0}^{r-1} c_i \lambda_i^s = 0, \qquad \sum_{i=0}^{r-1} c_i = 1, \tag{4.26}$$

for $s \in \{1, \dots, r_1\}$. The matrix of the coefficients of this linear system is a Vandermonde matrix of determinant $\prod_{i<j}(\lambda_i - \lambda_j)$ which is non-zero. Hence there exists a unique solution $(c_0, \dots, c_{r-1})$ that satisfies (4.26). On the other hand, so long as $\lambda_i \in (0,1)$, it is straightforward to see that if (4.6) holds also when φ is replaced with φ^{λ_i}.

(Step 5) We now discuss the uniqueness of T when $\gamma > 0$. Indeed if T and T' both satisfy (4.12), then $S = T - T'$ satisfies

$$|S(\psi_x^\delta)| \preceq \delta^\gamma.$$

From this we learn that for every $\phi \in \mathcal{D}$ with $\int \phi = 1$,

$$S(\psi^\delta * \phi) = \int S(\psi_x^\delta)\phi(x)\ dx \preceq \delta^\gamma.$$

Sending $\delta \to 0$ yields $S(\phi) = 0$. Thus $S = 0$. □

Remark 4.1(i) If we $\alpha, \beta \in (0,1)$, and F as in Example 4.2**(i)**, then the restriction of $\mathcal{T}$ to such F yields a bilinear continuous operator $\mathcal{A} : \mathcal{C}^\alpha \times \mathcal{C}^\beta \to \mathcal{C}^{\beta-1}$. When $\gamma = \alpha + \beta - 1 > 0$, we simply have

$$\mathcal{A}(f,g) = \lim_{n\to\infty} f(\nabla g * \tilde{\rho}^n) = \lim_{n\to\infty} f\nabla(g * \tilde{\rho}^n).$$

In other words, from

$$\delta_0 = \lim_{n\to\infty} \tilde{\rho}^n = \tilde{\rho}^1 + \sum_{k=1}^{\infty} \tilde{m}^k,$$

we deduce a convergent representation

$$\mathcal{A}(f,g) = f(\nabla g * \tilde{\rho}^1) + \sum_{k=1}^{\infty} f(\nabla g * \tilde{m}^k).$$

When $\gamma = \alpha + \beta - 1 > 0$, and $g \in C^1$, then $\mathcal{A}(f,g) = f\nabla g$ by uniqueness because the continuous function $T = f\nabla g$ satisfies (4.12). Hence the operator $\mathcal{A}(f,g)$ is an extension of the classical $\mathcal{A}_0 : C^1 \times C^1 \to C$ that is defined by $\mathcal{A}_0(f,g) = f\nabla g$. When $\gamma \le 0$ however,

$\mathcal{A}(f,g) \neq f\nabla g$ for $f,g \in C^1$. In our construction we simply dropped the terms B_k to have convergence. Indeed,

$$\mathcal{A}(f,g) = f(\nabla g * \rho) + \sum_{n=1}^{\infty} (f(\nabla g * \hat{\varphi}^n)) * \hat{\zeta}^n =: \mathcal{A}_0(f,g) + \widehat{\mathcal{A}}(f,g). \tag{4.27}$$

(ii) Note that when $\gamma < 0$, and $G \in \mathcal{C}^\gamma$, then $\mathcal{T}' = \mathcal{T}_\gamma + G$ also satisfies (4.12). In our definition of coherence seminorm $[F]$ and the formulation of (4.12) we have been using Hölder norm and Hölder spaces. However Theorem 4. 1 has been extended to Besov spaces $\mathcal{B}^\gamma_{p,q}$ in Hairer and Labbé [HL]. □

4.1 Regularity structure

In Chapter 2 we learned how to solve the ODE (2.1) with rough x. A solution was constructed as a fixed point of an operator that acted on Gubinelli pairs $(y,\hat{y})$ with $\hat{y}$ playing the role of dy/dx. The derivative $\hat{y}$ corresponds to a Taylor-like approximation for y, namely $y(s,t) = \hat{y}(s)x(s,t) + O(|t-s|^2)$. When we study SPDE of the type we discussed in the introduction, we encounter various terms of different degrees of singularities. To manage such SPDES, we first attach a Taylor-like expansion to our potential solution to each spatial point. For our purposed we need a generalization of polynomials where monomials are certain distributions of various degrees of singularities (orders). To manage this in an orderly and systematic fashion, Hairer formulated *regularity structures* to be able to perform algebraic manipulations with a (often finite) set of relevant distributions. We now give a detailed presentation of such structures.

Definition 4.2(i) By a regularity structure we mean a triplet (A,T,Γ) where $A \subset \mathbb{R}$ is a discrete set that is bounded below and $0 \in \mathbb{R}$;

$$T = \oplus_{\alpha \in A} T_\alpha,$$

with each T_α a Banach space with norm $\|\cdot\|_\alpha$, and G is a group of linear continuous transformation $\Gamma : T \to T$ such that if $\tau \in T_\alpha$, then

$$\Gamma\tau - \tau \in \oplus_{<\alpha} T_\beta := \oplus_{\beta<\alpha} T_\beta.$$

We always assume that $\dim T_0 = 1$. We also use the notation $T_\alpha = \langle \tau^1, \dots, \tau^k \rangle$, when $\dim T_\alpha = k$ and $\{\tau^1, \dots, \tau^k\}$, is a basis for T_α. We write $\{\mathbb{1}\}$ for the basis of T_0. In other words, $T_0 = \langle \mathbb{1} \rangle$. Also, when $\tau \in T$, we write $\|\tau\|_\alpha$ for the norm of its α component. We write $Pr_\alpha \tau$ for the α-component of τ so that $\|\tau\|_\alpha = \|Pr_\alpha \tau\|$.

(ii) We write $\mathcal{L}(T) = \mathcal{L}(T, \mathcal{D}'(\mathbb{R}^d))$ for the set of linear continuous maps $L : T \to \mathcal{D}'(\mathbb{R}^d)$. By a *model* for (A,T,G) we mean a pair of measurable maps $M = (\Pi, \Gamma)$, with

$$\Pi : \mathbb{R}^d \to \mathcal{L}(T), \quad \Gamma : \mathbb{R}^d \times \mathbb{R}^d \to G,$$

such that $\Pi_x = \Pi(x)$ and $\Gamma_{xy} = \Gamma(x, y)$ satisfy

$$\Pi_x \Gamma_{xy} = \Pi_y, \qquad \Gamma_{xy}\Gamma_{yz} = \Gamma_{xz}, \tag{4.28}$$

for every $x, y \in \mathbb{R}^d$. Moreover, if r is the smallest integer with $r > |\min A|$, $\alpha, \beta \in A$ with $\beta < \alpha$, and $K \subset \mathbb{R}^d$ is compact, then

$$\sup_{x\in K}\sup_{\delta\in(0,1]}\sup_{\varphi\in\mathcal{D}_r}\sup_{\tau\in A_\alpha} \frac{(\Pi_x\tau)(\varphi_x^\delta)}{\delta^\alpha\|\tau\|_\alpha} < \infty, \qquad \sup_{x,y\in K}\sup_{\tau\in A} \frac{\|\Gamma_{xy}\tau\|_\beta}{|x-y|^{\alpha-\beta}\|\tau\|_\alpha} < \infty. \tag{4.29}$$

(iii) Given a regularity structure (A, T, G), its model M, and $\gamma \in \mathbb{R}$, we define $\mathcal{C}_M^\gamma$ to be the set of maps $f : \mathbb{R}^d \to T_{<\gamma}$ such that for every compact set $K \subset \mathbb{R}^d$, and $\alpha \in A$ with $\alpha < \gamma$, we have

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

This turns $\mathcal{C}_M^\gamma$ to a Frechet Space (locally convex complete metric vector space). We also write $\mathcal{C}_{\beta,M}^\gamma = \mathcal{C}_\beta^\gamma$, for the set of those $f \in \mathcal{C}_M^\gamma$ such that $f(x) \in \oplus_{\beta\le\alpha<\gamma} T_\alpha$. □

The following is a corollary to Theorem 4.1.

Theorem 4.2 *For each $\gamma \in \mathbb{R}$, there exists a linear continuous operator $\mathcal{T}_M^\gamma : \mathcal{C}_M^\gamma \to \mathcal{D}'$ such that*

$$\left|(\mathcal{T}_M^\gamma(f) - \Pi_x f(x))(\psi_x^\delta)\right| \preceq \begin{cases} \delta^\gamma & \gamma \neq 0, \\ |\log\delta| & \gamma = 0, \end{cases} \tag{4.30}$$

uniformly over $\delta \in (0, 1]$, x in a compact set, and $\psi \in \mathcal{D}_r$.

Proof Set $F_x = \Pi_x f(x)$. (4.30) would follow from (4.12) if we can show that $F \in \mathcal{CG}^\gamma$. Indeed

$$\begin{aligned}
\left|(F_x - F_y)(\varphi_y^\delta)\right| &= \left|(\Pi_x f_x - \Pi_y f_y)(\varphi_y^\delta)\right| = \left|(\Pi_x(f_x - \Gamma_{xy}f_y))(\varphi_y^\delta)\right| \\
&\le \sum_{\alpha<\gamma} \left|(\Pi_x Pr_\alpha (f_x - \Gamma_{xy}f_y))(\varphi_y^\delta)\right| \\
&\preceq \sum_{\alpha<\gamma} \delta^\alpha \|f_x - \Gamma_{xy}f_y\|_\alpha \preceq \|f\|_{\alpha,K,M} \sum_{\alpha<\gamma} \delta^\alpha |x-y|^{\gamma-\alpha} \\
&= \|f\|_{\alpha,K,M}\ \delta^{-r} \sum_{\alpha<\gamma} \delta^{\alpha+r}|x-y|^{\gamma-\alpha+r} \\
&\preceq \|f\|_{\alpha,K,M}\ \delta^{-r}(\delta + |x-y|)^{\gamma+r},
\end{aligned}$$

as desired. ∎

Example 4.3(i) Assume that $A = \mathbb{N}_0$, and $T = \mathbb{R}[X_1, \dots, X_d]$ is the space of real polynomials of the d-commuting variables $X_1, \dots, X_d$. For each $r \in \mathbb{N}_0$, the space T_r is the subspace of homogeneous polynomials of degree r, and $T_{\le r}$ is the space of polynomial of degree at most r. The collection $\{X^k : \ |k| = r\}$ is a basis for T_r, hence $\dim T_r = \binom{d+r-1}{r-1}$. Using this basis, we equip T_r with the standard Euclidean norm. The group G consists of operators $\Gamma_h, \ h = (h_1, \dots, h_d) \in \mathbb{R}^d$ that is formally defined by

$$\Gamma_h X^k = \prod_{i=1}^{d} (X_i + h_i \mathbb{1}),$$

with the convention that $X_i \mathbb{1} = \mathbb{1} X_i = X_i$. We define a model $P = (\Pi, \Gamma)$ by

$$(\Pi_a X^k)(x) = \prod_{i=1}^{d} (x_i - a_i), \qquad \Gamma_{ab} = \Gamma_{a-b}.$$

The properties (4.28) follow from $\Gamma_h \Gamma_{h'} = \Gamma_{h+h'}$. Evidently,

$$\left|\langle \Pi_a X^k, \varphi_a^\delta \rangle\right| = \delta^{|k|} \left| \int x^k \varphi(x) \ dx \right| \preceq \delta^{|k|}, \qquad \|\Gamma_h X^k\|_s \preceq |h|^{|k|-s},$$

for every $s < |k|$. We next choose $\gamma = n + \gamma_0$ with $n \in \mathbb{N}_0$ and $\gamma_0 \in (0, 1)$, and study corresponding $\mathcal{M}_M^\gamma$. Clearly if $f \in \mathcal{C}_M^\gamma$, then

$$f(x) = \sum_{|k| \le n} a_k(x) X^k,$$

is a polynomial of degree n with the following property:

$$\left\| \sum_{|k| \le n} \left(a_k(y + h) X^k - a_k(y)(X + h\mathbb{1})^k \right) \right\|_r \preceq |h|^{\gamma - r}, \tag{4.31}$$

for every integer $r \le n$, and locally uniformly in $y, h \in \mathbb{R}^d$. The choice of $r = n$ yields

$$|a_k(y + h) - a_k(y)| \preceq |h|^{\gamma_0},$$

whenever $|k| = n$, which means that $a_k \in \mathcal{C}_{loc}^{\gamma_0}(\mathbb{R}^d)$. More generally

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

for $r \leq n$ (by $k \geq \ell$, we mean $k_i \geq \ell_i$ for $i = 1, \dots, d$). For $r \leq n-1$, we can write

$$\sum_{|\ell|=r} \left| a_\ell(y+h) - a_\ell(y) - \sum_{i=1}^{d} (\ell_i + 1) a_{\ell+\delta_i}(y) h_i \right| \preceq |h|^2 + |h|^{\gamma - r} \preceq |h|^{1+\gamma_0},$$

which implies that $\partial_{y_i} a_\ell(y) = (\ell_i + 1) a_{\ell+\delta_i}(y)$. Inductively, we deduce

$$a_\ell = \frac{\partial^\ell a_0}{\ell!}.$$

Hence $f(y)$ is the Taylor polynomial of the function a_0 at y, and $a_0 \in \mathcal{C}^\gamma(\mathbb{R}^d)$. In other words, $\mathcal{C}_M^\gamma$ is isomorphic to the Hölder space $\mathcal{C}^\gamma$. For such a function f, we simply have $\mathcal{T}_M^\gamma f = a_0$.

(ii) Given $\alpha \in (1/3, 1/2)$, let $A = \{\alpha - 1, 2\alpha - 1, 0, \alpha\}$, with $r = 1$, and define

$$T_0 = \langle \mathbb{1} \rangle, \quad T_\alpha = \langle X^1, \dots, X^\ell \rangle, \quad T_{\alpha-1} = \langle \dot{X}^1, \dots, \dot{X}^\ell \rangle, \quad T_{2\alpha-1} = \langle \dot{\mathbb{X}}^{ij} : \ 1 \leq i, j \leq \ell \rangle.$$

What we have in mind is that $X = (X_1, \dots, X_\ell)$ represents a Hölder continuous path x and $\dot{X}$ represents its derivative. However we wish to have a candidate for $x \otimes \dot{x}$. Abstractly we use the symbol $\mathbb{X} = X \otimes \dot{X}$ to represents such a product. Also, $G = \{\Gamma_h : \ h \in \mathbb{R}^\ell\}$, with

$$\Gamma_h \mathbb{1} = \mathbb{1}, \quad \Gamma_h X^i = X^i + h_i \mathbb{1}, \quad \Gamma_h \dot{X}^i = \dot{X}^i, \quad \Gamma_h \dot{\mathbb{X}}^{ij} = \dot{\mathbb{X}}^{ij} + h_i \dot{X}^j.$$

The last definition is motivated by the formal manipulation

$$\Gamma_h \mathbb{X} = \Gamma_h (X \otimes \dot{X}) = (X + h\mathbb{1}) \otimes \dot{X} = \mathbb{X} + h\dot{X}.$$

Given $\mathbf{x} = (x, \mathbb{X}) \in \mathcal{R}^\alpha$, we define a model $M = (\Pi, \Gamma)$ by $\Gamma_{st} := \Gamma_{x(t,s)}$, and

$$\begin{aligned}
&(\Pi_s \mathbb{1})(t) = 1, \qquad (\Pi_s X^i)(t) = x_i(s,t), \\
&(\Pi_s \dot{X}^i)(\psi) = \dot{x}_i(\psi) = -\int \dot{\psi}(t) x_i(t) \, dt, \\
&(\Pi_s \dot{\mathbb{X}}^{ij})(\psi) = \int \psi(t) \, d\mathbb{X}^{ij}(s,t) = -\int \dot{\psi}(t) \mathbb{X}^{ij}(s,t) \, dt.
\end{aligned}$$

As for the verification first equation in (4.28) in the case of $\dot{\mathbb{X}}$, observe

$$\begin{aligned}
(\Pi_s \Gamma_{ss'} \dot{\mathbb{X}})(\psi) &= \left(\Pi_s (\dot{\mathbb{X}} + x(s', s) \otimes \dot{X}) \right) (\psi) \\
&= -\int \dot{\psi}(t) \left(\mathbb{X}(s,t) + x(s', s) \otimes x(t) \right) \, dt \\
&= -\int \dot{\psi}(t) \left(\mathbb{X}(s', s) + \mathbb{X}(s,t) + x(s', s) \otimes x(s,t) \right) \, dt \\
&= -\int \dot{\psi}(t) \mathbb{X}(s', t) \, dt = (\Pi_{s'} \dot{\mathbb{X}})(\psi),
\end{aligned}$$

by the Chen's relation. As for the first property in (4.29), observe that for each s, $x_i(s,\cdot) \in \mathcal{C}^\alpha$, $\dot{x}_i \in \mathcal{C}^{\alpha-1}$, and

$$\left|(\Pi_s \dot{\mathbb{X}}^{ij})(\varphi_s^\delta)\right| = \left|\int \dot{\varphi}_s^\delta(t)\mathbb{X}^{ij}(s,t)\ dt\right| = \delta^{-1}\left|\int \dot{\varphi}(\theta)\mathbb{X}^{ij}(s,s+\delta\theta)\ d\theta\right| \le [\mathbb{X}]_{2\alpha}\|\varphi\|_{C^1}\delta^{2\alpha-1}.$$

(iii) When now study $\mathcal{C}_{0,}M^{2\alpha}$ when M is as in the previous example. Let $Y(t) = y(t)\mathbb{1} + \hat{y}(t)X \in \mathcal{C}_M^{2\alpha}$. This is equivalent to saying

$$\|Y(s) - \Gamma_{st}Y(t)\|_r = \|(y(s) - y(t) - \hat{y}(t)x(t,s))\mathbb{1} + (\hat{y}(s) - \hat{y}(t))X\|_r \preceq |s-t|^{2\alpha-r},$$

for $r = 0$ and α. Equivalently,

$$|y(s) - y(t) - \hat{y}(t)x(t,s)| \preceq \delta^{2\alpha}, \quad |\hat{y}(s) - \hat{y}(t)| \preceq \delta^\alpha.$$

Hence such $Y \in \mathcal{C}_M^{2\alpha}$ iff $\mathbf{y} = (y,\hat{y}) \in \mathcal{G}^\alpha(x)$. Moreover, it is not hard to see that in fact $\mathcal{T}^{2\alpha}Y = y$.

Now imagine that we wish to define $y\dot{x}$. First we perform this multiplication formally/abstractly, namely

$$Y\dot{X} = (y\mathbb{1} + \hat{y}X)\dot{X} = y\dot{X} + \hat{y}(X \otimes \dot{X}).$$

In the setting of our regularity structure, we may use Theorem 4.2 to turn $Y\dot{X}$ into a distribution that is indeed $\dot{z} = \mathbf{y}\cdot\mathbf{x}$ of Chapter 2. To see this, first observe then $Y\dot{X} \in \mathcal{C}_M^{3\alpha-1}$, because

$$\begin{aligned}(Y\dot{X})(s) - \Gamma_{st}(Y\dot{X})(t) &= y(s)\dot{X} + \hat{y}(s)\dot{\mathbb{X}} - \Gamma_{x(t,s)}\ (y(t)\dot{X} + \hat{y}(t)\dot{\mathbb{X}}) \\ &= y(s)\dot{X} + \hat{y}(s)\dot{\mathbb{X}} - y(t)\dot{X} - \hat{y}(t)\dot{\mathbb{X}} - \hat{y}(t)\ x(t,s) \otimes \dot{X} \\ &= (y(s) - y(t) - \hat{y}(t)x(t,s))\dot{X} + (\hat{y}(s) - \hat{y}(t))\dot{\mathbb{X}},\end{aligned}$$

with

$$\begin{aligned}&|y(s) - y(t) - \hat{y}(t)x(t,s)| \preceq |t-s|^{2\alpha} = |t-s|^{3\alpha-1-(\alpha-1)}, \\ &|\hat{y}(s) - \hat{y}(t)| \preceq |t-s|^\alpha = |t-s|^{3\alpha-1-(2\alpha-1)}.\end{aligned}$$

Since $3\alpha - 1 > 0$, by Theorem 4.2, there exists a unique operator $\mathcal{T}_M^{3\alpha-1}$ such that

$$\left|\left(\left(\mathcal{T}_M^{3\alpha-1}(Y\dot{X}) - (y(s)\dot{x} + \hat{y}(s)\mathbb{X}_t(s,\cdot)\right)(\psi_s^\delta)\right| \preceq \delta^{3\alpha-1}. \tag{4.32}$$

By approximation, one can show that if $T = \mathcal{T}_M^{3\alpha-1}(Y\dot{X})$, then $T \in \mathcal{C}^{\alpha-1}(\mathbb{R})$, and that we can choose $\psi = \mathbb{1}_{[0,1]}$. Hence if $z(t) = T([0,t])$, then $\dot{z} = T$ and z satisfies (2.12). We may also define an operator on the set of such expressions as

$$\mathcal{I}(Y\dot{X})(t) = z(t)\mathbb{1} + y(t)X,$$

where $z(t) = \mathcal{T}_M^{3\alpha-1}(Y\dot{X})(\mathbb{1}_[0,t])$. If we write $\mathcal{C}_{<0}^{3\alpha-1}$ for the set of such $Y\dot{X}$, then $\mathcal{I} : \mathcal{C}_{<0}^{3\alpha-1} \to \mathcal{C}_0^{2\alpha}$. More generally, a function $f \in \mathcal{C}_M^{3\alpha-1}$ (with $\alpha \in (1/3, 1/2]$) would look like

$$f(t) = y(t)\dot{X} + \hat{y}(t)\dot{\mathbb{X}} + h(t)\mathbb{1} = (Y\dot{X})(t) + h(t)\mathbb{1},$$

with $Y\dot{X}$ as before, and h a Hölder continuous function of Hölder exponent $3\alpha - 1 \in (0, 1]$. In this case, we simply have

$$\mathcal{T}_M^{3\alpha-1}(f) = \mathcal{T}_M^{3\alpha-1}(Y\dot{X}) + h.$$

(iv) Given a C^3 function $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times\ell}$, We may define an operator $\Sigma : \mathcal{C}_0^{2\alpha} \to \mathcal{C}_0^{2\alpha}$, by

$$\Sigma(y\mathbb{1} + \hat{y}X) = \sigma(y)\ \mathbb{1} + (D\sigma)(y)\hat{y}\ X.$$

We then set

$$\mathcal{F}(Y) = \hat{y}^0\mathbb{1} + \mathcal{I}(\Sigma(Y)\dot{X}).$$

A fixed point $Y = y\mathbb{1} + \hat{y}X$ of the operator $\mathcal{F} : \mathcal{C}_0^{2\alpha} \to \mathcal{C}_0^{2\alpha}$ yields a solution to (2.1), as discussed in Theorem 2.3.

(v) We now discuss a regularity structure associated with the iterated integrals of Section 2.3. Given $\alpha > 0$, let $A = \{n\alpha : n \in \mathbb{N}_0\}$ be the set of indices, and $T = H = T(\mathbb{R}^\delta)$ be the Tensor algebra associated with $\mathbb{R}^\ell$. We regards H as a Hopf algebra with the product $\sqcup\!\sqcup$. Its dual $H^* = T(\mathbb{R}^{d*})$ is equipped with the product $\bullet = \otimes$. Recall that the group $G(H) \subset H^*$ is the set of characters. We then define a group $G = \{\Gamma_g : g \in G(H)\}$, where the linear $\Gamma_g : H \to H$ is defined by the duality

$$\langle f, \Gamma_g h\rangle = \langle g^{-1} \bullet f, h\rangle.$$

As in Proposition C.1 of Appendix C, we know that G is a group. Also, if $h = e_a$ for a word $a = (i_1, \dots, i_n)$, then using $\Gamma_g = (g^{-1}\otimes id)\Delta$ (see Proposition C.2**(ii)** below),

$$\Gamma_g h = \sum_{m=0}^{n} g^{-1}(h_m)\hat{h}_m,$$

where $h_m = (i_1, \dots, i_m)$, and $\hat{h}_m = (i_{m+1}, \dots, h_n)$. In particular

$$\Gamma_g h - h = \sum_{m=1}^{n} g^{-1}(h_m)\hat{h}_m \in H_n,$$

because $h_m = e_{a_m}$, with $|a_m| = n - m < n$.

To define a model, we take a path $\mathbf{x} : [0, t_0] \to G(H)$, we set $\mathbf{x}(s, t) = \mathbf{x}(s)^{-1} \bullet \mathbf{x}(t)$, and define the linear function $\Pi = \Pi^{\mathbf{x}} : H \to C([0, t_0]; \mathbb{R})$ and $\Gamma_{s,t} \in G$ by

$$(\Pi_s a)(t) = \langle \mathbf{x}(s,t), a \rangle, \qquad \Gamma_{s,t} := \Gamma_{\mathbf{x}(s,t)},$$

for every word a. Note

$$\Gamma_{s,u} \circ \Gamma_{u,t} = \Gamma_{\mathbf{x}(s,u)} \circ \Gamma_{\mathbf{x}(u,t)} = \Gamma \mathbf{x}(s,u) \bullet \mathbf{x}(u,t) = \Gamma_{\mathbf{x}(s,t)} = \Gamma_{s,t}.$$

(vi) In this example we discuss a multidimensional analogue of **(ii)**. Given $\alpha \in (1/3, 1/2]$, let $A = \{\alpha - 1, 2\alpha - 1, 0, \alpha\}$, and define $T_0 = \langle 1 \rangle$, $T_\alpha = \langle F^1, \dots, F^\ell \rangle$,

$$T_{\alpha-1} = \langle F^i_j : \ i = 1, \dots, d, \ j = 1, \dots, \ell \rangle, \qquad T_{2\alpha-1} = \langle \mathbb{F}^{k,i}_j : \ i = 1, \dots, d, \ j, k = 1, \dots, \ell \rangle.$$

Similarly, for $h \in \mathbb{R}^\ell$,

$$\Gamma_h \mathbb{1} = \mathbb{1}, \qquad \Gamma_h F = F + h\mathbb{1}, \qquad \Gamma_h F^i_j = F^i_j, \qquad \Gamma_h \mathbb{F}^{k,i}_j = \mathbb{F}^{k,i}_j + h_k F^i_j.$$

What we have in mind is that $F = (F^1, \dots, F^\ell)$ represents an α-Hölder continuous function $f = (f^1, \dots, f^\ell) : \mathbb{R}^d \to \mathbb{R}^\ell$, F^i_j represents the partial derivative $\partial_j f^i := f^i_{x_j}$, and $\mathbb{F}^{k,i}_j$ represents the product $f^k \partial_j f^i$. Let $h^{k,i}_j$ be a candidate for $h^{k,i}_j$, so that

$$|\langle h^{k,i}_j - f^k(x) \partial_j f^i, \psi^\delta_x \rangle| \preceq \delta^{2\alpha - 1}.$$

We can use f and h to build a model M for our regularity structure, with $\Gamma_{xy} = \Gamma_{f(y) - f(x)}$. Given a sufficiently differentiable function $\eta : \mathbb{R}^\ell \to \mathbb{R}$, we wish to make sense of $\eta(f) \partial_j f^i$. For this, we define

$$H(x) = \eta(f(x))\mathbb{1} + \sum_{r=1}^{\ell} \partial_r \eta(f(x)) F^r.$$

We claim that $H \in \mathcal{C}^{2\alpha}_M$. For this, we need

$$\|H(x) - \Gamma_{xy} H(y)\|_\beta \preceq |x - y|^{2\alpha - \beta},$$

for $\beta = 0, \alpha$. Since $H(x) - \Gamma_{xy} H(y)$ equals

$$\left(\eta(f(x)) - \eta(f(y)) - \nabla \eta(f(y)) \cdot (f(x) - f(y))\right) \mathbb{1} + \left(\nabla \eta(f(x)) - \nabla \eta(f(y))\right) \cdot F,$$

it suffices to assume that $\eta \in C^2$. We wish to have a unique candidate for $\eta(f) \cdot f_{x_j}$. Formally,

$$H F^i_j = (\eta \circ f \ \mathbb{1} + \nabla \eta \circ f \cdot F) F^i_j = \eta \circ f \ F_j + \nabla \eta \circ f \cdot \mathbb{F}^i_j.$$

As in **(iii)**, we can readily show that $HF^i_j \in \mathcal{C}^{3\alpha - 1}_M$. This allows us to apply the reconstruction theorem to find a unique candidate for $\eta(f) \cdot f_{x_j}$. We may define a continuous operator $\mathcal{I}(f, h)$ such that for smooth f,

$$\mathcal{I}(f, f \nabla f) = \eta(f) \nabla f.$$

□

4.2 Schauder Estimate

The classical Schauder estimate asserts that if $u \in \mathcal{C}^\alpha$, then $\Delta^{-1}u \in \mathcal{C}^{\alpha+2}$. Note that when $d \geq 3$, then $\Delta^{-1}u = u * G$, where G is a constant multiple of $|x|^{2-d}$. Observe that we may write $G = K + \hat{K}$, with $\hat{K}$, smooth, and K a function that is smooth off of the origin, with support in the unit ball, and satisfying the bounds

$$|\partial^k K(x)| \preceq |x|^{2-d-|k|}. \tag{4.33}$$

For may of the PDEs we discussed in the Introduction, we are interested in the regularity gain of operator $(\partial_t - \Delta)^{-1}$, which is associated with the heat kernel

$$p(x,t) = (4\pi t)^{-d/2} e^{-\frac{|x|^2}{4t}} 1\!\!1(t > 0).$$

The *parabolic Schauder estimate* assert that there is a gain of 2 (parabolic) derivatives as we apply the operator $(\partial_t - \Delta)^{-1}$ on a function provided that we use parabolic distances (which in practice means two spatial derivatives, and one temporal derivative). Note that the only singularity of p occurs at $(0,0)$. Moreover, p satisfies a bound similar to (4.33) provided that we use the parabolic dimension $d+2$.

Lemma 4.1 *For every $r \in \mathbb{N}_0$, there exists a constant $\bar{c}_r$ such that for every $z = (x,t) \in \bar{B}_1$, with $t > 0$, and every k with $|k|_{par} = r$,*

$$|\partial^k p(z)| \leq \bar{c}_r |z|_{par}^{2-(d+2)-r} = \bar{c}_r |z|_{par}^{-d-r}. \tag{4.34}$$

Proof First observe that if $\ell_{d+1} = 0$, then

$$\partial^\ell p(x,t) = t^{-(d+|\ell|)/2} P_\ell \left(\frac{x}{\sqrt{t}} \right) e^{-\frac{|x|^2}{4t}} =: t^{-(d+|\ell|)/2} R_\ell \left(\frac{x}{\sqrt{t}} \right), \tag{4.35}$$

where P_ℓ is a polynomial of degree $|\ell| = |\ell|_{par}$. We can readily verify (4.35) by induction on $|\ell|$. Differentiating (4.35) with respect to t yields

$$\partial_t^s \partial^\ell p(x,t) = t^{-(d+|\ell|+2s)/2} P_{\ell,s} \left(\frac{x}{\sqrt{t}} \right) e^{-\frac{|x|^2}{4t}} =: t^{-(d+|\ell|+2s)/2} R_{\ell,s} \left(\frac{x}{\sqrt{t}} \right), \tag{4.36}$$

where $P_{\ell,s}$ is a polynomial of degree $|\ell| + 2s$. Again the proof of (4.36) can be carried out by induction on s. On account of (4.36), (4.34) would follow if we can find $\bar{c}_r$ such that

$$P_{\ell,s}(a) e^{-\frac{a^2}{4}} \leq \bar{c}_{|\ell|+2s} (a+1)^{-d-|\ell|-2s},$$

or equivalently,

$$P_{\ell,s}(a)(a+1)^{d+|\ell|+2s} \leq \bar{c}_{|\ell|+2s}\, e^{\frac{a^2}{4}}.$$

This is evidently true. □

We now state and prove our elliptic Schauder estimate.

Theorem 4.3 **(i)** *Fix $\beta > 0$. Let $K : \mathbb{R}^d \to \mathbb{R}$ be a function that is supported in $B_1(0)$, is smooth off of the origin, and satisfies*

$$|\partial^k K(x)| \preceq |x|^{\beta-d-|k|}, \tag{4.37}$$

for every $k \in \mathbb{N}_0^d$. Then

$$[K * u]_{\hat{\mathcal{C}}^{\beta+\alpha}} \preceq [u]_{\hat{\mathcal{C}}^{\alpha}}. \tag{4.38}$$

(ii) *Fix $\beta > 0$. Let $K : \mathbb{R}^{d+1} \to \mathbb{R}$ be a function that is supported in $\bar{B}_1(0)$, is smooth off of the origin, and satisfies*

$$|\partial^k K(z)| \preceq |z|^{\beta-d-2-|k|}, \tag{4.39}$$

for every $k \in \mathbb{N}_0^d$. Then

$$[K * u]_{\hat{\mathcal{C}}_{par}^{\beta+\alpha}} \preceq [u]_{\hat{\mathcal{C}}_{par}^{\alpha}}. \tag{4.40}$$

Proof We only present a proof for the first part because the second part can be treated with a verbatim argument.

(Step 1) We first express our kernel as a sum of smooth functions. To achieve this, we start

4.3 Approximation Using Semigroups

Recall that by the Reconstruction Theorem, we can make sense of fg, where $f \in \mathcal{C}^\alpha,\ g \in \mathcal{C}^\beta$, with $\alpha \in (0,1),\ \beta \in (-1,0)$. Let us write $P_t = e^{t\Delta}$. We write

$$P_t(fg) = fP_tg + \sum_{n=0}^{\infty} \left(P_{t(1-2^{-(n+1)})}\big(fP_{t2^{-(n+1)}}g\big) - P_{t(1-2^{-n})}\big(fP_{t2^{-n}}g\big)\right).$$

Note that each term on the right-hand side is well-defined. If we show that the series is convergent, then the right-hand side is well-defined. This gives a candidate for the left-hand side. If we write A_n for the n-th term in the sum, then

$$A_n = P_{t-t_n}\big(P_{t_{n+1}}(fP_{t_{n+1}}g) - fP_{t_n}g\big),$$

where $t_n = t2^{-n}$. Note

$$\Gamma_s(f,g)(z) := \big(P_s(fP_sg) - fP_{2s}g\big)(z) = \int K_s(z,y)(f(y) - f(z))\, dy,$$

where

$$K_s(z,y) = p_s(z-y)(P_s g)(y).$$

Observe,

$$|K_s(z,y)| \preceq [g]_\beta \; s^{\beta/2} \; p_s(z-y).$$

Hence,

$$\begin{aligned}\|\Gamma_s(f,g)\|_{L^\infty} \leq & [f]_\alpha \sup_z \int |K_s(z,y)| \; |z-y|^\alpha \; dy \\ \preceq & [f]_\alpha [g]_\beta \; s^{\beta/2} \; \sup_z \int p_s(z-y) \; |z-y|^\alpha \; dy \\ \preceq & [f]_\alpha [g]_\beta \;\; s^{(\alpha+\beta)/2}.\end{aligned}$$

As a consequence,

$$\|A_n\|_{L^\infty} \preceq [f]_\alpha [g]_\beta \;\; t_{n+1}^{(\alpha+\beta)/2},$$

because P_θ is a contraction. From this we learn

$$\begin{aligned}\|P_t(fg)\|_{L^\infty} &\leq \|fP_t g\|_{L^\infty} + \|P_t(fg) - fP_t g\|_{L^\infty} \\ &\preceq \|f\|_{L^\infty}[g]_\beta t^{-\beta/2} + [f]_\alpha [g]_\beta \;\; t^{-(\alpha+\beta)/2} \\ &\preceq \|f\|_\alpha [g]_\beta \;\; t^{\beta/2}.\end{aligned}$$

Hence,

$$\|fg\|_{\mathcal{C}^\beta} \preceq \|f\|_\alpha [g]_\beta.$$

This means that the bilinear operator $\mathcal{A}(f,g) = fg$ has a continuous extention to $\mathcal{C}^\alpha \times \mathcal{C}^\beta$ when $\alpha + \beta > 0$. By approximation, we also have

$$\big\|P_t\big(\mathcal{A}(f,g) - f \;\, P_t g\big\|_{L^\infty} \preceq \|f\|_\alpha [g]_\beta \;\; t^{(\alpha+\beta)/2}.$$

This means

$$\sup_x \big|\big(\mathcal{A}(f,g) - f(x)g\big)(\psi_x^\delta)\big| \preceq \|f\|_\alpha [g]_\beta \; \delta^{\alpha+\beta},$$

where $\delta = t^{1/2}$, and $\psi = p$ is the Gaussian density.

In fact

$$\big(P_t(fg) - fP_t g\big)(x) = \int R_t(x,y,z)(f(y) - f(z)) \; dydz,$$

where

$$R_t(x,y,z) = \sum_{n=0}^{\infty} p_{t-t_n}(x-z) K_{t_{n+1}}(z,y).$$

We have,

$$|R_t(x,y,z)| \preceq [g]_\beta \sum_{n=0}^{\infty} t_{n+1}^{\beta/2}\ p_{t-t_n}(x-z)p_{t_{n+1}}(z-y),$$

$$\int |R_t(x,y,z)|\ dz \preceq [g]_\beta \sum_{n=0}^{\infty} t_{n+1}^{\beta/2}\ p_{t-t_{n+1}}(x-y).$$

4.4 Exercises

(i) Let $H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a bounded continuous funcion. Assume that the funcion $H(x,y)$ is 1-periodic in y-varible. (Here $y \in \mathbb{R}^d$, and F is periodic in each y-coordinate.) Define $h_n(x) := H(x, nx)$. Show

$$\lim_{n\to\infty} h_n = \bar{h},$$

in weak sense, where

$$\bar{h}(x) = \int_{[0,1]^d} H(x,y)\ dy.$$

More precisely, show that for any continuous $\varphi : \mathbb{R}^d \to \mathbb{R}$ of compact support,

$$\lim_{n\to\infty} \int h_n(x)\ \varphi(x)\ dx = \int \bar{h}(x)\ \varphi(x)\ dx.$$

(ii) Let $F, G : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be two bounded C^1 functions such that $F(x,y)$ and $G(x,y)$ are 1-periodic in y. Given $\alpha, \beta \in (0,1]$, define

$$f_n(x) := n^{-\alpha} F(x,nx), \qquad g_n(x) =: n^{-\beta} G(x,nx)$$

Show

$$\sup_n \|f_n\|_{\mathcal{C}^\alpha} < \infty, \qquad \sup_n \|g_n\|_{\mathcal{C}^\beta} < \infty.$$

Show that

$$\lim_{n\to\infty} f_n \nabla g_n = 0,$$

if $\alpha + \beta > 1$.

(iii) Show that when $\alpha = \beta = 1/2$ in **(ii)**, then

$$\lim_{n\to\infty} \langle f_n \nabla g_n, \varphi\rangle = \int_{\mathbb{R}^d} \left[\int_{[0,1]^d} F(x,y) G_y(x,y)\ dy\right] \cdot \varphi(x)\ dx,$$

for every coninuous $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ of compact support.

(iv) Let $\tau_a : \Omega \to \Omega;\ a \in \mathbb{R}^d$, be family of measurable maps such that $\tau_a \circ \tau_b = \tau_{a+b},\ \tau_0 = id$. Let $\mathbb{P}$ be an ergodic invariant measure for the group action τ. Assume $\hat{F}, \hat{G} : \mathbb{R}^d \times \Omega \to \mathbb{R}$ are two measurable functions and set

$$F(x, y; \omega) = \hat{F}(x, \tau_y \omega), \qquad G(x, y; \omega) = \hat{G}(x, \tau_y \omega).$$

Assume that F and G and bounded C^1 functions $\mathbb{P}$ almost surely. Define

$$f_n(x, \omega) = n^{-\alpha} F(x, nx; \omega), \qquad g_n(x, \omega) = n^{-\alpha} G(x, nx; \omega).$$

When $\alpha = 1/2$, show

$$\lim_{n\to\infty} \langle f_n \nabla g_n, \varphi \rangle = \int_{\mathbb{R}^d} \left[\int_\Omega \hat{F}(x, \omega) \nabla \hat{G}(x, \omega)\ \mathbb{P}(d\omega) \right] \cdot \varphi(x)\ dx,$$

$\mathbb{P}$-almost surely.

(v) Show that in (4.32) we can choose $\psi = 1\!\!1_{[0,1]}$. *Hint:* Start from $\varphi \in \mathcal{D}$ with $\varphi \geq 0$, $\int \varphi = 1$, *supp* $\varphi \subset [0, 1]$, and from it build $\hat{\varphi}^n(t) = 2^n \varphi(2^n t)$, $\tilde{\varphi}^n(t) = 2^n \varphi(-2^n t)$. Choose η_n, ζ_n such that $\dot{\eta} = \varphi^n,\ \dot{\zeta}^n(t) = \tilde{\varphi}^n(1 - t)$. Use η_n and ζ_n to find ψ_n and $\tilde{\psi}_n$ such that

$$1\!\!1_{[0,1]} = \sum_{n=0}^{\infty} (\psi_n + \tilde{\psi}_n),$$

so that ψ_n and $\tilde{\psi}_n$ are supported in $[0, 2^{-n}]$ and $[1 - 2^{-n}, 1]$ respectively. Use this representation to derive (4.32) for $\psi = 1\!\!1_{[0,1]}$.

5 KPZ Equation

In this chapter we discuss the question of well-posedness for the KPZ equation:

$$h_t = h_{xx} + h_x^2 - C, \tag{5.1}$$

where $h : \mathbb{T} \times [0, \infty) \to \mathbb{R}$, C is a constant, and ξ is the space-time white noise. Note that if P denotes the kernel of the operator $(\partial_t - \partial_{xx})^{-1}$, then (5.1) can be rewritten as

$$h = P * h^0 + P * \left(h_x^2 + \xi - C\right). \tag{5.2}$$

Theorem 5.1 *Using parabolic scaling, we have $\xi \in \mathcal{C}^{\alpha}_{par}$, for every $\alpha < -(d+2)/2$.*

Proof Given $\varphi \in \mathcal{D}$, we write

$$\varphi^{\delta}_{(a,s)}(x,t) = \delta^{-d-2}\varphi\left(\frac{x-a}{\delta}, \frac{t-s}{\delta^2}\right).$$

By (3.3),

$$\begin{aligned}\left[\mathbb{E}\left[\xi(\varphi_z^{\delta})\right]^{2n}\right]^{\frac{1}{2n}} =& c_n\left[\mathbb{E}\left[\xi(\varphi^{\delta})\right]^{2}\right]^{1/2} = c_n\left(\int (\varphi^{\delta})^2 \ dxdt\right)^{1/2} \\ &= c_n\delta^{-d-2}\left(\int \ \varphi^2 \ dxdt\right)^{1/2}.\end{aligned}$$

From this, and a variant of Theorem 3.1, we deduce that $\xi \in \mathcal{C}_{par}^{-\frac{d+2}{2}-\frac{1}{2n}}$, almost surely. □

We wish to build a regularity structure in which we can reformulate the equation (5.2) in terms of various abstract objects so that we can find a solution in a suitable $\mathcal{C}^{\gamma}_M$, and apply the reconstruction theorem to find a candidate for a renormalized solution. We first write $P = K + \hat{K}$ so that $\hat{K}$ is smooth and K is of compact support that is smooth expect at the origin. First imagine that we can find a linear operator $\mathcal{K} : \mathcal{C}^{\gamma}_M \to \mathcal{C}^{\gamma+2}_M$ such that

$$\mathcal{T}\mathcal{K} = K * \mathcal{T}.$$

The operator $\mathcal{K} = \mathcal{I} + \mathcal{K}'$, where $\mathcal{I}$ satisfies

$$\Pi_x \mathcal{I}\tau = K * \Pi_x\tau - \sum_{|k|<\deg\tau+2} (\partial^k K * \Pi_x\tau)(x).$$

We now formulate an abstract variant of (5.2)

$$H = P * h^0 \mathbb{1} + \mathcal{I}((\partial H)^2 + \Theta) + (\mathcal{K}' + \hat{K})((\partial H)^2 + \Theta), \tag{5.3}$$

where Θ abstractly represents the white noise, ∂ abstractly represents the spatial derivative, and $(\partial H)^2$ is an abstract candidate for a product. The last term on the right-hand side of (5.3) would take value in $\otimes_{n\in\mathbb{N}} T_n$ and is polynomial like expression. Let us also write $\mathcal{I}'$ for $\partial\mathcal{I}$. To simplify our presentation, we will be using graphical notations. For Θ, we use a circle $\circ$, we use $|$ for the operator $\mathcal{I}'$, and $\wr$ for $\mathcal{I}$.

We assume $d = 1$ and write $\alpha_-$for a number $\alpha' < \alpha$ that is close to α. So far we have $\xi \in \mathcal{C}_{par}^{-(3/2)-}$ which suggests that if A is the set of indices in our regularity structure, then we have $-(3/2)_- \in A$.

6 Energy Solution to Stochastic Burgers Equation

In this chapter we derive KPZ for the Ginzburg-Landau model of Section 6.1. To simplify our presentation, we assume consider a periodic lattice instead of $\mathbb{Z}$. More precisely, we write $\mathbf{r}(t) = (r_i(t) : \ i \in \mathbb{Z}_N)$, with $\mathbb{Z}_N = \{0, \dots, N\|$, with $0 = N$, and $\mathbf{r}$ a diffusion with the generator $\mathcal{L} = \mathcal{L}_N = \varepsilon^{-2}\mathcal{S} + \varepsilon^{-3/2}\mathcal{A}$, where the operators $\mathcal{A}$ and $\mathcal{S}$ were defined by (1.26), and $\varepsilon = N^{-1}$. Note that now the summations in (1.26) are over $i = 1, \dots, N$, with the periodic conventions $N + 1 = N$, $0 = N$. Recall that product measures

$$\nu_\alpha(d\mathbf{r}) = \prod_{i=1}^{N} \Theta_\alpha(dr_i), \qquad \text{with} \qquad \Theta(r) = \frac{1}{Z(\alpha)} e^{\alpha r - V(r)}\ dr,$$

are invariant for every α such that

$$Z(\alpha) = \int e^{\alpha r - V(r)}\ dr < \infty.$$

We assume that $V = V_0 + V_1$ so that V_0 is uniformly convex, and V_1 have a bounded C^2 norm. As a result $Z(\alpha) < \infty$ for every $\alpha \in \mathbb{R}$. Define the empirical measure

$$m_\varepsilon(t, dx) = \varepsilon^{1/2} \sum_i (r_i(t) - \bar\rho)\delta_{\varepsilon i + \sqrt{\varepsilon}\lambda'(\bar\rho)t}(dx).$$

In other words, for a smooth test function φ,

$$\int \varphi(x)\ m_\varepsilon(t, dx) = \varepsilon^{1/2} \sum_i (r_i(t) - \bar\rho)\ \varphi(\varepsilon i + \sqrt{\varepsilon}\lambda'(\bar\rho)t). \tag{6.1}$$

Theorem 6.1 *Assume that the process $\mathbf{r}(t)$ starts from the equilibrium measure $\nu_{\lambda(\bar\rho)}$. Then the low ε limit of m_ε is an energy solution of* (1.30).

We first sketch the proof. Note that if we differentiate (6.1) with respect to time, we get an expression that involves sums of the form

$$\sum_i F(r_i(t))\psi(i, t),$$

where $F = V'$ and $\psi(i, t) = \varphi(\varepsilon i + \sqrt{\varepsilon}\lambda'(\bar\rho)t)$. Since the system is at equilibrium, we need a *local central limit theorem*, to replace F with its average with respect to ν. Note however that our Markov process is ergodic only when we fix the value the conserved quantity $\sum_i r_i$.

Definition 6.1(i) We write $\mathbf{r}^\ell = (r_1, \dots, r_\ell)$, and

$$\nu_\alpha^\ell(d\mathbf{r}^\ell) = \prod_{i=1}^{\ell} \Theta_\alpha(dr_i), \qquad \nu_{\ell,m}(d\mathbf{r}^\ell) = \nu_\alpha^\ell\left(d\mathbf{r}^\ell | m_\ell(\mathbf{r}^\ell) = m\right),$$

where $m_\ell(\mathbf{r}^\ell) = \ell^{-1}(r_1 + \dots + r_\ell)$. It is not hard to see

$$\nu_\alpha^\ell\left(d\mathbf{r}^\ell | m_\ell(\mathbf{r}^\ell) = m\right) = \frac{1}{Z_{\ell,m}} \exp\left(-\sum_{i=1}^{\ell-1} V'(r_i) - V'\left(\ell m - (r_1 + \dots + r_{\ell-1})\right)\right) \prod_{i=1}^{\ell-1} dr_i.$$

(ii) Define

$$M(\lambda) := \log Z(\lambda), \qquad \rho(\lambda) = \int r\Theta_\lambda(dr), \qquad \sigma^2(\lambda) = \int (r - \rho(\lambda))^2\, \Theta_\lambda(dr).$$

It is not show that $M'(\lambda) = \rho(\lambda)$, and $M''(\lambda) = \sigma(\lambda)^2$.

(iii) Given a function $F : \mathbb{R} \to \mathbb{R}$, we set

$$\hat{F}(\rho) = \int F(r_1)\ \Theta_{\lambda(\rho)}(dr), \qquad \tilde{F}_\ell(\rho) = \int F(r_1)\ \nu_{\ell,\rho}(d\mathbf{r}^\ell).$$

□

Theorem 6.2 *There exists a constant $C_0(F)$ such that*

$$\int \left|\tilde{F}_\ell(m_\ell(\mathbf{r}^\ell)) - \hat{F}(\bar\rho) - \hat{F}'(\bar\rho)(m_\ell(\mathbf{r}^\ell) - \bar\rho) - \frac{1}{2}\hat{F}''(\bar\rho)\left[(m_\ell(\mathbf{r}^\ell) - \bar\rho)^2 - \ell^{-1}\sigma^2(\bar\rho)\right]\right|^2\ \nu_{\bar\lambda}^\ell(d\mathbf{r}^\ell)$$

is bounded above by $C_0(F)\ell^{-3}$. Here $\bar\rho = \rho(\bar\lambda)$.

Given F as in Theorem 6.2, we may set

$$G(r) = F(r) - \hat{F}(\bar\rho) - \hat{F}'(\bar\rho)(r - \bar\rho), \tag{6.2}$$

so that $\hat{G}(\bar\rho) = \hat{G}'(\bar\rho) = 0$ and $\hat{G}(\bar\rho) = \hat{F}(\bar\rho)$. In this case we have the following consequence of Theorem 6.1.

Corollary 6.1 *Let $G : \mathbb{R} \to \mathbb{R}$ be a function with $\hat{G}(\bar\rho) = \hat{G}'(\bar\rho) = 0$. Then*

$$\int \left|\tilde{G}_\ell(m_\ell(\mathbf{r}^\ell)) - -\frac{1}{2}\hat{G}''(\bar\rho)\left[(m_\ell(\mathbf{r}^\ell) - \bar\rho)^2 - \ell^{-1}\sigma^2(\bar\rho)\right]\right|^2\ \nu_{\bar\lambda}^\ell(d\mathbf{r}^\ell) \le C_0(G)\ell^{-3}.$$

We now consider the process $\mathbf{r}(t)$ that starts from the equilibrium measure $\bar\nu := \nu_{\lambda(\bar\rho)}$.

Theorem 6.3 *There exists a constant $C_1(F)$ such that*

$$\mathbb{E}^{\bar\nu}\Big[\int_0^T \sum_i \tau_i\Big(F(r_i(s)) - \hat F(\bar\rho) - \hat F'(\bar\rho)(m_\ell(\mathbf{r}^\ell(s)) - \bar\rho) \\ - \frac{1}{2}\hat F''(\bar\rho)\left((m_\ell(\mathbf{r}^\ell(s)) - \bar\rho)^2 - \ell^{-1}\sigma^2(\bar\rho)\right)\Big)\zeta(i,s)\ ds\Big]^2 \\ \le C_1(F)(\varepsilon^2\ell + T\ell^{-2})\int_0^T \sum_i \zeta(i,s)^2\ ds,$$

$$\mathbb{E}^{\bar\nu}\Big[\int_0^T \sum_i \tau_i\Big(F(r_i(s)) - \hat F(\bar\rho) - \hat F'(\bar\rho)(m_\ell(\mathbf{r}^\ell(s)) - \bar\rho)\Big)\zeta(i,s)\ ds\Big]^2 \\ \le C_1(F)(\varepsilon^2\ell + T\ell^{-1})\int_0^T \sum_i \zeta(i,s)^2\ ds.$$

6.1 Uniqueness

We now focus on the solutions of KPZ equation, or SBE at equilibrium:

(6.3) $$h_t = h_{xx} + \lambda h_x^2 + \xi, \qquad h(x,0) = x\cdot\bar\rho + B(x),$$

(6.4) $$\rho_t = \rho_{xx} + \lambda(\rho^2)_x + \xi_x \qquad \rho(x,0) = \bar\rho + \eta,$$

where $B(x)$ is a two-sided standard Brownian motion, $\eta = B'(x)$ is a space white noise, and ξ is a space-time white noise. What we have in mind is that $\rho = h_x$.

Definition 6.1.1 We say that a function $\rho : [0,T] \to \mathcal{D}'$ with $\rho(x,0) = \eta(x)$ is a solution of (6.4) if there exists a process $A : [0,T] \to \mathcal{D}'$ of finite variation (in t) such that the following conditions are met for every test function $\varphi \in \mathcal{D}$:

(i) For every smooth function $\chi : \mathbb{R} \to \mathbb{R}$ of compact support with $\int \chi = 1$,

(6.5) $$A(t)(\varphi') = \lim_{\delta\to 0}\int_0^t \int (\rho(s) * \chi^\delta)^2\ \varphi'\ dx\ ds,$$

where $\chi^\delta(x) = \delta^{-1}\chi(x/\delta)$.

(ii) The processes

$$M(t,\varphi) := \rho(t)(\varphi) - \rho(0)(\varphi) - \int_0^t \rho(s)(\varphi'')\ ds + \lambda A(t)(\varphi'),$$
$$N(t,\varphi) := M(t,\varphi)^2 - 2\|\varphi'\|_0^2\ t,$$

are martingales with respect to the forward filtration $\mathcal{F}_t$, which is generated by $\{\xi(s) : s \in [0,t]\}$ (or $\{(\rho(s), A(s)) : s \in [0,t]\}$). Equivalently

$$d\rho(t)(\varphi) = \rho(s)(\varphi'')\, dt + \lambda\, dA(t)(\varphi') - dW(t)(\varphi'), \tag{6.6}$$

where W is a cylindrical Brownian motion: $W'(t) = \xi$.

(iii) Set $\hat{\rho}(t) = \rho(T-t)$, and $\hat{A}(t) = A(T-t) - A(T)$. The processes

$$\hat{M}(t,\varphi) := \hat{\rho}(t)(\varphi) - \hat{\rho}(0)(\varphi) + \int_0^t \hat{\rho}(s)(\varphi'')\, ds + \lambda \hat{A}(t)(\varphi'),$$
$$\hat{N}(t,\varphi) := \hat{M}(t,\varphi)^2 - 2\|\varphi'\|_0^2\, t,$$

are martingales with respect to the backward filtration $\widehat{\mathcal{F}}_t$, which is generated by $\{\xi(s) : s \in [T-t, T]\}$. □

Definition 6.1.2 We say that a function function $h : [0,T] \times \mathbb{T} \to \mathbb{R}$ with $h(x,0) = B(x)$, is a solution of (6.3) if there exists a process $A : [0,T] \times \mathbb{T} \to \mathbb{R}$ of finite variation (in t) such that the following conditions are met for every test function $\psi \in \mathcal{D}$:

(i) For every smooth function $\chi : \mathbb{R} \to \mathbb{R}$ of compact support with $\int \chi = 1$,

$$\int A(t,x)\psi(x)\, dx = \lim_{\delta \to 0} \int_0^t \int \left[(h(s) * \chi^\delta)^2 - \|(\chi^\delta)^2\|_0^2 \right]\, \psi\, dx\, ds, \tag{6.7}$$

where $\chi^\delta(x) = \delta^{-1}\chi(x/\delta)$.

(ii) The processes

$$M(t,\psi) := h(t)(\psi) - h(0)(\varphi) - \int_0^t h(s)(\psi'')\, ds + \lambda A(t)(\psi),$$
$$N(t,\psi) := M(t,\psi)^2 - 2\|\psi\|_0^2\, t,$$

are martingales with respect to the forward filtration $\mathcal{F}_t$, which is generated by $\{\xi(s) : s \in [0,t]\}$ (or $\{(h(s), A(s)) : s \in [0,t]\}$). **(iii)** Set $\hat{h}(t) = h(T-t)$, and $\hat{A}(t) = A(T-t) - A(T)$. The processes

$$\hat{M}(t,\psi) := \hat{h}(t)(\varphi) - \hat{h}(0)(\varphi) + \int_0^t \hat{h}(s)(\psi'')\, ds + \lambda \hat{A}(t)(\psi),$$
$$\hat{N}(t,\psi) := \hat{M}(t,\psi)^2 - 2\|\psi\|_0^2\, t,$$

are martingales with respect to the backward filtration $\widehat{\mathcal{F}}_t$, which is generated by $\{\xi(s) : s \in [T-t, T]\}$. □

We start from a solution of (6.4) and produce a solution for (6.3). We are tempted to try to make sense of formally defined $h(x,t)=\int_{-\infty}^{x}\ \rho(s,y)\ dy$. For our purposes, we choose a smooth function $\eta\ge 0$ of compact support such that $\int\eta=1$, and choose a function Θ^x such that if $h(x,t)=\rho(t)(\Theta^x)$, then

$$h_x=\rho,\qquad \langle h(t),\eta\rangle=0.$$

Indeed if we set $\Theta^x(y)=\Theta(x,y)=1\!\!1(y\le x)-\theta(y)$, then for sure $h_x=\rho$, and to satisfy the second requirement, we need

$$\int\Theta(x,y)\eta(x)\ dx=\int_y^\infty\eta(x)\ dx-\theta(y)=0.$$

Note that the function Θ^x is of compact support, but not smooth. Because of this, choose yet another smooth even function χ with $\int\chi=1$, and set

$$\rho^\delta(t)=\rho(t)*\chi^\delta,\qquad h^\delta(x,t)=\langle\Theta^x,\rho^\delta(t)\rangle.$$

Here, $\chi^\delta(x)=\delta^{-1}\chi(x/\delta)$, and ρ is chosen so that $\hat\chi=1$ on a neighborhood of 0. Since χ is even, we have

$$h^\delta(t)=\rho(t)(\Theta_{x,\delta}),\qquad\text{where}\qquad\Theta_{x,\delta}=\Theta^x*\chi^\delta.$$

Observe

$$\Theta'_{x,\delta}(y)=(\Theta^x)'*\chi^\delta=-\chi^\delta(y-x)+(\eta*\chi^\delta)(y),\qquad \partial_x\Theta_{x,\delta}(y)=(\partial_x\Theta^x)*\chi^\delta=\chi^\delta(y-x),$$

$$\begin{aligned}\partial_{xx}\Theta_{x,\delta}(y)&=\partial_x[\chi^\delta(y-x)]=\partial_y[-\chi^\delta(y-x)]=\partial_y[\Theta'_{x,\delta}(y)-\eta*\chi^\delta)(y)]\\&=\Theta''_{x,\delta}(y)-(\eta*\chi^\delta)'(y)=\Theta''_{x,\delta}(y)-(\eta'*\chi^\delta)(y),\end{aligned}$$

Because of this, we may write

$$dh^\delta(x,t)=\rho(t)(\Theta''_{x,\delta})\ dt+\lambda dA(t)(\Theta'_{x,\delta})-dW(t)(\Theta'_{x,\delta}).$$

We now set $Z^\delta(x,t)=e^{\lambda h^\delta(x,t)}$. Evidently,

$$Z^\delta_x(x,t)=\lambda Z^\delta(x,t)h^\delta_x(x,t),\qquad Z^\delta_{xx}(x,t)=\lambda Z^\delta(x,t)h^\delta_{xx}(x,t)+\lambda^2Z^\delta_x(x,t)(h^\delta_x(x,t))^2.$$

By Ito's formula,

$$\begin{aligned}dZ^\delta(x,t)=&\lambda Z^\delta(x,t)\ dh^\delta(x,t)+\frac12\lambda^2Z^\delta(x,t)\ d[h^\delta(x,t),h^\delta(x,t)]\\=&\lambda Z^\delta(x,t)\left(\rho(t)(\Theta''_{x,\delta})\ dt+\lambda dA(t)(\Theta'_{x,\delta})+\frac12\lambda^2\|\tau_x\xi^\delta-\xi^\delta*\eta\|_0^2\ dt-dW(t)(\Theta'_{x,\delta})\right)\\=&\lambda Z^\delta(x,t)\left((h^\delta_{xx}+\rho(t)(\eta'*\chi^\delta))\ dt+\lambda dA(t)(\Theta'_{x,\delta})+\frac12\lambda^2\|\tau_x\xi^\delta-\xi^\delta*\eta\|_0^2\ dt-dW(t)(\Theta'_{x,\delta})\right)\\=&Z^\delta_{xx}\ dt+\lambda Z^\delta(x,t)\left((-\lambda^2h^\delta_x(x,t))^2-\rho(t)(\eta'*\chi^\delta))\ dt+\lambda dA(t)(\Theta'_{x,\delta})+\frac12\lambda^2\|\tau_x\xi^\delta-\xi^\delta*\eta\|_0^2\ dt-\right.\end{aligned}$$

A Function Spaces

We start by recalling some standard notations:

Definition A.1(i) Given $\varphi : \mathbb{R}^d \to \mathbb{R}$, and a multi-index $k = (k_1, \dots, k_d) \in \mathbb{N}_0^d$, we write

$$\partial^k \varphi = \partial_{x_1}^{k_1} \dots \partial_{x_d}^{k_d} \varphi.$$

We also write $|k|$ for $k_1 + \dots + k_d$.

(ii) Given $r \in \mathbb{N}_0$, we write C^r for the space of functions $\varphi : \mathbb{R}^d \to \mathbb{R}$ such that $\partial^k \varphi$ exists and is continuous for every multi-index k with $|k| \le r$. We set

$$\|\varphi\|_\infty = \|\varphi\|_{L^\infty} = \sup_{x \in \mathbb{R}^d} |\varphi(x)|, \qquad \|\varphi\|_{C^r(K)} = \sup_{|k| \le r} \|\partial^k \varphi|_{L^| i(K)}, \qquad \|\varphi\|_{C^r} = \|\varphi\|_{C^r(\mathbb{R}^d)}.$$

When D is an open subset of $\mathbb{R}^d$, then the spaces $L^\infty(D)$, and $C^r(D)$ are defined analogously. Likewise, for $\varphi : D \to \mathbb{R}$, the norms $\|\varphi\|_{L^\infty(D)}$, and $\|\varphi\|_{C^r(D)}$ are defined in a similar way.

(iii) We write $\mathcal{D} := \mathcal{D}(\mathbb{R}^d)$ for the space of C^∞ functions of compact support. Given a set $K \subset \mathbb{R}^d$, we write $\mathcal{D}(K)$ for the set of $\varphi \in \mathcal{D}$ such that $\varphi = 0$ on the complement of K. We refer to the members of $\mathcal{D}$ as test functions. We also define

$$\mathcal{D}_r = \{\varphi \in \mathcal{D}(B_1) : \ \|\varphi\|_{C^r} \le 1\},$$

where $B_\delta(a)$ denotes the closed ball of radius δ and center a, and $B_\delta := B_\delta(0)$.

(iv) Let K be a compact subset of $\mathbb{R}^d$, and $r \in \mathbb{N}_0$. Then a linear function $T : \mathcal{D}(K) \to \mathbb{R}$ is called a *distribution* on K of order r, if there exists a constant c such that

$$T(\varphi) \le c\|\varphi\|_{C^r},$$

for every $\varphi \in \mathcal{D}(K)$. The space of such distributions is denoted by $\mathcal{D}'_r(K)$. We also set

$$\mathcal{D}'(K) = \cup_{r=0}^\infty \mathcal{D}'(K).$$

We write $\mathcal{D}' = \mathcal{D}'(\mathbb{R}^d)$ for the set $T : \mathcal{D} \to \mathbb{R}$ such that the restriction of T to $\mathcal{D}(K)$ is in $\mathcal{D}'(K)$. Likewise, we write $\mathcal{D}'_r = \mathcal{D}'_r(\mathbb{R}^d)$ for the set $T : \mathcal{D} \to \mathbb{R}$ such that the restriction of T to $\mathcal{D}(K)$ is in $\mathcal{D}'_r(K)$.

(vi) Given a measurable map $\varphi : \mathbb{R}^d \to \mathbb{R}$, and a scale $\delta > 0$, we set

$$\varphi_x^\delta(y) := \delta^{-d} \varphi\left(\delta^{-1}(y - x)\right).$$

We also write $\varphi_x = \varphi_x^1,\ \varphi^\delta = \varphi_0^\delta$.

(v) Given $\alpha \in (0,1]$, and a compact set K, we define

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

where

$$\langle f,g\rangle = \int_{\mathbb{R}^d} f(x)g(x)\ dx.$$

We write $\mathcal{C}^\alpha_{loc}$ for the set of functions such that $[u]_{\alpha;B_\ell} < \infty$ for every ℓ.

(vi) For $n \in \mathbb{N}_0$ and $\beta \in (0,1]$, we write $C^{\ell,\beta}$ for the set of functions $u : \mathbb{R}^d \to \mathbb{R}$ such that $u \in C^r$, and $\partial^k u \in \mathcal{C}^\beta_{loc}$ for every k with $|k| = n$. We put

$$\|u\|_{C^{\ell,\beta}(K)} = \|u\|_{C^\ell(K)} + \sum_{|k|=n} [\partial^k u]_{\beta;K}.$$

(vii) Given $\alpha > 0$, we define $\mathcal{C}^\alpha_{loc}(\mathbb{R}^d) = \mathcal{C}^\alpha_{loc}$ to be the set $C^{n,\beta}$, where

$$n = \max\{m \in \mathbb{N}_0 :\ n < \alpha\}, \qquad \beta = \alpha - n.$$

We also set

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

where

$$P_{x;u}(y) = \sum_{|k|\le n} \frac{\partial^k u(x)}{k!}(y-x)^k.$$

(viii) Given $\alpha < 0$, we write $r(\alpha)$ for the smallest integer r such that $-\alpha < r$. We then define

$$[u]_{\mathcal{C}^\alpha(K)} = \sup_{x\in K} \sup_{\delta\in(0,1]} \sup_{\varphi\in\mathcal{D}_{r(\alpha)}} \delta^{-\alpha} u(\varphi_x^\delta),$$

We write $\mathcal{C}^\alpha(K)$ for the set of distributions u such that $[u]_{\mathcal{C}^\alpha(K)} < \infty$. We write $C^\alpha_{loc}(\mathbb{R}^d) = C^\alpha_{loc}$, for the set of distributions such that $[u]_{C^\alpha(B_\ell)} < \infty$ for every $\ell \in \mathbb{N}$.

(xi) For $\alpha = n + \beta > 0$, with $\beta \in (0,1)$, we define

$$[u]_{\widehat{\mathcal{C}}^\alpha(K)} = \sup_{x\in K} \sup_{\delta\in(0,1]} \sup_{\varphi\in\mathcal{D}^{(n)}} \delta^{-\alpha} u(\varphi_x^\delta),$$

where $\mathcal{D}^{(n)}$ is the set of $\varphi \in \mathcal{D}$ such that

$$\int \varphi P\ dx = 0,$$

for every polynomial P with $\deg P \le n$. The *Hölder-Zygmund* $\widehat{\mathcal{C}}^\alpha_{loc}$ is the set of u such that $[u]_{\widehat{\mathcal{C}}^\alpha(K)} < \infty$, for every compact set $K \subset \mathbb{R}^d$.

(x) The Hölder spaces $\mathcal{C}^\alpha$ are special cases of Besov spaces; $\mathcal{C}^\alpha_{loc} = \mathcal{B}^\alpha_{\infty,\infty;loc}$. For $\mathcal{B}^\alpha_{p,q;loc}$ we replace uniform bounds in x with L^p norm, and uniform bounds in $\delta \in (0,1]$ with $L^q(\delta^{-1}d\delta)$-norm. □

Theorem A.1 *There are positive constants c_0 and c_1 such that for every $\alpha > 0$,*

$$c_0[u]_{\alpha;B_r} \le [u]_{\mathcal{C}^\alpha(B_r)} \le c_1[u]_{\alpha;B_{r+1}}. \tag{A.1}$$

Proof Evidently,

$$\sup_{\varphi\in\mathcal{D}_0} \langle u - u(x), \varphi^\delta_x \rangle = \sup_{\varphi\in\mathcal{D}_0} \int (u(x+\delta y) - u(x))\varphi(y)\ dy = \int_{B_1} |u(x+\delta y) - u(x)|\ dy. \tag{A.2}$$

As a result,

$$[u]_{\mathcal{C}^\alpha(B_r)} \le |B_1|\ [u]_{r+1,\alpha}.$$

On the other hand, given $x, y \in B_r$, we set $\delta = |x-y|$ and integrate both sides of the inequality

$$|u(x) - u(y)| \le |u(x) - u(z)| + |u(z) - u(y)|,$$

over $z \in B_\delta(x) \cap B_\delta(y)$ to assert

$$\begin{aligned} |B_\delta(x)\cap B_\delta(y)|\ |u(x)-u(y)| &\le \int_{B_\delta(x)\cap B_\delta(y)} [|u(x)-u(z)| + |u(z)-u(y)|]\ dz \\ &\le \int_{B_\delta(x)} |u(x)-u(z)|\ dz + \int_{B_\delta(y)} |u(z)-u(y)|dz \\ &\le 2\delta^{d+\alpha}[u]_{\mathcal{C}^\alpha(B_r)} \le c_2|B_\delta(x)\cap B_\delta(y)|\ \delta^\alpha[u]_{\mathcal{C}^\alpha(B_r)}, \end{aligned}$$

where we have used (A.2) for the last line. Hence

$$[u]_{r,\alpha} \le c_2[u]_{\mathcal{C}^\alpha(B_r)},$$

as desired. □

Note that if a distribution T is of order $r = r_K$ in a compact set K, then $T(\varphi)$ is well-defined for a C^r function whose support is contained in K. We often need to use convolution to approximate a distribution by smooth test functions. For this the following elementary properties of convolution are useful.

$$(\varphi * \psi)^\delta = \varphi^\delta * \psi^\delta, \qquad (\varphi * \psi)_x = \varphi * \psi_x. \tag{A.3}$$

Note that if μ is a measure of compact support, and $\varphi \in \mathcal{D}$, then $\varphi * \mu \in \mathcal{D}$. Note that if we decompose $\mathbb{R}^d$ into dyadic boxes of the form

$$I(t_i^n) := I(t_{i_1}^n, \dots, t_{i_d}^n) = \prod_{r=1}^{d} [t_{i_r}^n, t_{i_r+1}^n), \quad i = (i_1, \dots, i_d) \in \mathbb{Z}^d, \quad t_{i_r}^n = i_r 2^{-n},$$

then

$$(\varphi * \mu)(x) = \int \varphi_y(x)\ \mu(dy) = \lim_{n\to\infty} \sum_{i \in \mathbb{Z}^d} \varphi(x - t_i^n)\ \mu(I(t_i^n)),$$

with the convergence holds with respect to the C^r-topology, for every $r \in \mathbb{N}$. As a result

$$T(\varphi * \mu) = \int T(\varphi_y)\ \mu(dy). \tag{A.4}$$

Given $\varphi \in \mathcal{D}$ and $T \in \mathcal{D}'$, we may define the convolution $T * \varphi \in C^r$ by

$$(T * \varphi)(x) := T(\tilde{\varphi}_x),$$

where $\tilde{\varphi}(x) = \varphi(-x)$. It is not hard to check that $T * \varphi$ is smooth, and

$$\partial^k (T * \varphi)(x) = \int T\left((\partial^k \tilde{\varphi})_x\right) = \left(T * (\partial^k \varphi)\right)(x).$$

Moreover, (A.4) allows us to write

$$\langle T * \tilde{\varphi}, \psi \rangle = T(\varphi * \psi). \tag{A.5}$$

From this, we can readily deduce

$$T(\psi) = \lim_{\delta \to 0} \langle T(\varphi_x^\delta), \psi(x) \rangle, \tag{A.6}$$

for every $(T, \varphi) \in \mathcal{D}' \times \mathcal{D}$, with $\int \varphi = 1$.

Theorem A.2 *For every $\alpha = n + \beta$, with $n \in \mathbb{N}_0$ and $\alpha_0 \in (0,1)$, there exists a constant C_0, such that*

$$[u]_{\widehat{\mathcal{C}}^\alpha(K)} \le [u]_{\mathcal{C}^\alpha(K)} \le C_0 [u]_{\widehat{\mathcal{C}}^\alpha(K)}. \tag{A.7}$$

$\widehat{\mathcal{C}}^\alpha_{loc} = \mathcal{C}^\alpha_{loc}$.

Proof *(Step 1)* Evidently,

$$[u]_{\widehat{\mathcal{C}}^\alpha(K)} \le [u]_{\mathcal{C}^\alpha(K)}.$$

For the reverse inequality, we first assume that $n = 0$. Fix $\rho \in \mathcal{D}$ such that the support of ρ is contained in $B_1(0)$, and $\int \rho \ dx = 1$. For any distribution u, we certainly have

$$u = \lim_{n\to\infty} u_n = u_0 + \sum_{n=0}^{\infty} (u_{n+1} - u_n), \tag{A.8}$$

where

$$u_n(x) = u\left(\rho_x^{2^{-n}\delta}\right).$$

On the other hand, since

$$\rho_x^{2^{-(n+1)}\delta} - \rho_x^{2^{-n}\delta} = (\rho^{1/2} - \rho)_x^{2^{-n}\delta}, \qquad \int (\rho^{1/2} - \rho)\ dx = 0,$$

and $c_0^{-1}(\rho^{1/2} - \rho) \in \mathcal{D}_0$, for $c_0 = 2\|\rho\|_\infty$, we learn,

$$|u_{n+1}(x) - u_n(x)| \le c_0 [u]_{\widehat{\mathcal{C}}^\alpha(K)} \delta^\alpha 2^{-n\alpha},$$

for $x \in K$. From this we learn that if $u \in \widehat{\mathcal{C}}^\alpha$, then the right-hand side of (A.8) is uniformly convergent in K. Hence the distribution u must be a continuous function. Moreover,

$$|\langle u - u(x), \rho_x^\delta\rangle| = |u(x) - \langle u, \rho_x^\delta\rangle| = \left|\sum_{n=0}^{\infty} (u_{n+1}(x) - u_n(x))\right| \le c_0 c_1 [u]_{\widehat{\mathcal{C}}^\alpha(K)} \delta^\alpha,$$

where $c_1 = (1 - 2^{-\alpha})^{-1}$. Now if $\varphi \in \mathcal{D}_0$ with $c_2 := \int \varphi \neq 0$, then we can set $\rho = c_2^{-1}\varphi$ to assert

$$c_2^{-1} |\langle u - u(x), \varphi_x^\delta\rangle| \le 2c_1 \|\rho\|_\infty [u]_{\widehat{\mathcal{C}}^\alpha(K)} \delta^\alpha \le 2c_1 c_2^{-1} [u]_{\widehat{\mathcal{C}}^\alpha(K)} \delta^\alpha.$$

By approximation, we can drop the assumption $\int \varphi \neq 0$. As a result,

$$[u]_{\mathcal{C}^\alpha(K)} \le 2c_1 [u]_{\widehat{\mathcal{C}}^\alpha(K)},$$

as desired.

(Step 2) We now turn our attention to the case $n > 0$. Choose any $k \in \mathbb{N}_0^d$ such that $r = |k| \le n$. Choose any $\varphi \in \mathcal{D}_r$ such that

$$\int \varphi P\ dx = 0,$$

for any polynomial P with $\deg P \le n-r$. We can then argue,

$$\left|(\partial^k u)(\varphi_x^\delta)\right| = \left|u(\partial^k \varphi_x^\delta)\right| = \delta^{-r}\left|u((\partial^k\varphi)_x^\delta)\right| \le \delta^{\alpha-r}[u]_{\widehat{\mathcal{C}}^\alpha(K)},$$

because

$$\int \partial^k \varphi Q\ dx = (-1)^r \int \varphi \partial^k Q\ dx = 0,$$

for any polynomial Q with $\deg Q \le n$. This suggests defining

$$[w]_{\widehat{\mathcal{C}}_\ell^\gamma(K)} = \sup_{x\in K}\sup_{\delta\in(0,1]}\sup_{\varphi\in\mathcal{D}_\ell^{(s)}} \delta^\gamma \left|w(\varphi_x^\delta)\right|,$$

where $\mathcal{D}_\ell^{(s)}$ is the set $\varphi \in \mathcal{D}_\ell$ such that $\int \varphi P = 0$, for any polynomial P with $\deg P \le s$. With this definition, we can now assert

$$[\partial^k u]_{\widehat{\mathcal{C}}_r^{\alpha-|k|}(K)} \le [u]_{\widehat{\mathcal{C}}^\alpha(K)}.$$

To complete the proof, we need to show that we can replace $\widehat{\mathcal{C}}_r^{\alpha-|k|}$ with $\widehat{\mathcal{C}}^{\alpha-|k|} = \widehat{\mathcal{C}}_0^{\alpha-|k|}$. Once this is achieved, then we use the first step to assert that $u \in C^n$, $\partial^k u \in \mathcal{C}^\beta$ for any k with $|k| = n$, and the second inequality in (A.7).

(Step 3) Fix $r \in \mathbb{N}$. Pick $\psi \in \mathcal{D}$ such that the support of ψ is contained in $B_1(0)$, and

$$\int \psi(x) x^k\ dx = \delta_{0,k},$$

for every k with $|k| \le r$ (see Step 4 of the proof of Theorem 4.1 for the construction of such ψ). As in the first step, we always have

$$w = w_0 + \sum_{n=0}^{\infty}(w_{n+1} - w_n),$$

where

$$w_n(x) = w\left(\psi_x^{2^{-n}\delta}\right), \qquad (w_{n+1} - w_n)(x) = w\left(\zeta_x^{2^{-n}\delta}\right),$$

where $\zeta = \psi^{1/2} - \psi$. Observe

$$\int \zeta(x) x^k\ dx = 0,$$

for every k with $|k| \le r$. As in the first step,

$$|w_{n+1}(x) - w_n(x)| \le c_0 [w]_{\widehat{\mathcal{C}}_r^\gamma} 2^{-n\gamma}\delta^\gamma,$$

where $c_0 = \|\zeta\|_{C^r}$. This implies that if $[w]_{\widehat{\mathcal{C}}^\gamma_r} < \infty$ for some $\gamma > 0$, then w is a function. On the other hand,

$$w(\varphi_x^\delta) = w(\varphi_x^\delta * \psi^\delta) + \sum_{n=0}^{\infty} w\left(\varphi_x^\delta * \zeta^{2^{-n}\delta}\right).$$

Clearly,

$$\left|w * \zeta^{2^{-n}\delta}(x)\right| \le c_0 [w]_{\widehat{\mathcal{C}}^\gamma_r} 2^{-n\gamma}\delta^\gamma, \tag{A.9}$$

which yields

$$\sum_{n=0}^{\infty}\left|w\left(\varphi_x^\delta * \zeta^{2^{-n}\delta}\right)\right| \le c_0[w]_{\widehat{\mathcal{C}}^\gamma_r}(1-2^{-\gamma})^{-1}\int |\varphi|\ dx\ \delta^\gamma.$$

Moreover,

$$\int (\varphi * \psi)(x)P(x) = \int \psi(y)\int \varphi(z)P(z+y)\ dz\ dy = 0,$$

for every polynomial P with $\deg P \le r$, and

$$\|\varphi * \psi\|_{C^r} \le \|\psi\|_{C^r}\int |\varphi|\ dx.$$

From this, and (A.9) we deduce

$$[w]_{\widehat{\mathcal{C}}^\gamma} \le c[w]_{\widehat{\mathcal{C}}^\gamma_r},$$

for a constant c. This completes the proof. □

The function spaces we have defined so far can be used to study the regularity of elliptic PDEs. For the parabolic PDEs, we need to mollify our definitions so that we can take advantage of the parabolic scaling $(x,t) \mapsto (\lambda x, \lambda^2 t)$. For this, let us set

$$d_{par}((x,t),(x',t')) = |x-x'| + |t-t'|^{1/2} =: |(x-x', t-t')|_{par}.$$

Note that we are slightly abusing the notation here because $|\cdot|_{par}$ is not a norm.

Definition A.2(i) Given a multi-index $k = (k_1, \dots, k_{d+1}) \in \mathbb{N}_0^{d+1}$, we write

$$|k|_{par} = k_1 + \cdots + k_d + 2k_{d+1}.$$

A parabolic ball is defined by

$$\bar{B}_\delta(a) = \left\{z \in \mathbb{R}^{d+1} :\ |z-a|_{par} < \delta\right\}.$$

Note $|\bar{B}_\delta(a)| = \delta^{d+2}|\bar{B}_1(0)|$.

(ii) Given $r \in \mathbb{N}$, we write C^r_{par} for the set of functions $u : \mathbb{R}^{d+1} \to \mathbb{R}$ such that $\partial^k u$ is continuous for any k with $|k|_{par} \le r$. We also write

$$\|u\|_{C^r_{par}(K)} = \sum_{|k|_{par} \le r} \|\partial^k u\|_{L^\infty(K)}.$$

We simply $\|u\|_{C^r_{par}}$ for $\|u\|_{C^r_{par}(\mathbb{R}^{d+1})}$. We also write $\bar{\mathcal{D}}_r$ for the set of $\varphi \in \mathcal{D}$ such that the support of φ is contained in $\bar{B}_1(0)$, and $\|\varphi\|_{C^r_{par}} \le 1$.

(iii) Given a measurable map $\varphi : \mathbb{R}^{d+1} \to \mathbb{R}$, and a scale $\delta > 0$, we set

$$\bar{\varphi}^\delta_{(x,s)}(y,t) := \delta^{-d-2} \varphi\left(\delta^{-1}(y-x), \delta^{-2}(t-s)\right).$$

We also write $\bar{\varphi}_z = \bar{\varphi}^1_z,\ \bar{\varphi}^\delta = \bar{\varphi}^\delta_0$.

(iv) Given $\alpha \in (0,1]$, and a compact set K, we define

$$[u]_{par;\alpha;K} = \sup_{z,z' \in K} \frac{|u(z) - u(z')|}{|z - z'|^\alpha_{par}}, \qquad [u]_{\mathcal{C}^\alpha_{par}(K)} = \sup_{z \in K} \sup_{\delta \in (0,1]} \sup_{\varphi \in \bar{\mathcal{D}}_0} \frac{\langle u - u(z), \bar{\varphi}^\delta_z \rangle}{\delta^\alpha},$$

where

$$\langle f, g \rangle = \int_{\mathbb{R}^{d+1}} f(z) g(z)\ dz.$$

We write $\mathcal{C}^\alpha_{par;loc}$ for the set of functions such that $[u]_{par;\alpha;B_\ell} < \infty$ for every ℓ.

(v) For $n \in \mathbb{N}_0$ and $\beta \in (0,1]$, we write $C^{\ell,\beta}_{par}$ for the set of functions $u : \mathbb{R}^{d+1} \to \mathbb{R}$ such that $u \in C^r_{par}$, and $\partial^k u \in \mathcal{C}^\beta_{par;loc}$ for every k with $|k|_{par} = n$. We put

$$\|u\|_{C^{\ell,\beta}_{par}(K)} = \|u\|_{C^\ell_{par}(K)} + \sum_{|k|_{par} = n} [\partial^k u]_{par;\beta;K}.$$

(vi) Given $\alpha > 0$, we define $\mathcal{C}^\alpha_{par;loc}(\mathbb{R}^d) = \mathcal{C}^\alpha_{loc}$ to be the set $C^{n,\beta}$, where

$$n = \max\{m \in \mathbb{N}_0 :\ n < \alpha\}, \qquad \beta = \alpha - n.$$

We also set

$$[u]_{\mathcal{C}^\alpha(K)} = \sup_{x \in K} \sup_{\delta \in (0,1]} \sup_{\varphi \in \mathcal{D}_0} \frac{\langle u - P_{x;u}, \varphi^\delta_x \rangle}{\delta^\alpha},$$

where

$$P_{x;u}(y) = \sum_{|k|_{par} \le n} \frac{\partial^k u(x)}{k!} (y - x)^k.$$

(vii) Given $\alpha < 0$, we write $r(\alpha)$ for the smallest integer r such that $-\alpha < r$. We then define

$$[u]_{\mathcal{C}^\alpha_{par}(K)} = \sup_{z\in K} \sup_{\delta\in(0,1]} \sup_{\varphi\in\bar{\varphi}\mathcal{D}_{r(\alpha)}} \delta^{-\alpha} u(\varphi^\delta_x),$$

We write $\mathcal{C}^\alpha_{par}(K)$ for the set of distributions u such that $[u]_{\mathcal{C}^\alpha_{par}(K)} < \infty$. We write $C^\alpha_{par;loc}(\mathbb{R}^d) = C^\alpha_{par;loc}$, for the set of distributions such that $[u]_{C^\alpha_{par}(\bar{B}_\ell)} < \infty$ for every $\ell \in \mathbb{N}$.

(viii) For $\alpha = n + \beta > 0$, with $\beta \in (0,1)$, we define

$$[u]_{\widehat{\mathcal{C}}^\alpha_{par}(K)} = \sup_{z\in K} \sup_{\delta\in(0,1]} \sup_{\varphi\in\bar{\mathcal{D}}^{(n)}} \delta^{-\alpha} u(\varphi^\delta_x),$$

where $\bar{\mathcal{D}}^{(n)}$ is the set of $\varphi \in \mathcal{D}$ such that

$$\int \varphi P \; dx = 0,$$

for every polynomial P with $\deg P \leq n$. The *Hölder-Zygmund* $\widehat{\mathcal{C}}^\alpha_{par;loc}$ is the set of u such that $[u]_{\widehat{\mathcal{C}}^\alpha_{par}(K)} < \infty$, for every compact set $K \subset \mathbb{R}^{d+1}$. □

Theorem A.3 *There are positive constants c_0 and c_1 such that for every $\alpha > 0$,*

$$c_0[u]_{par;\alpha;\bar{B}_r} \leq [u]_{\mathcal{C}^\alpha_{par}(\bar{B}_r)} \leq c_1[u]_{par;\alpha;\bar{B}_{r+1}}. \tag{A.10}$$

The proof of Theorem A.3 is very similar to the proof of Theorem A.1 and is omitted (simply replace the Euclidean balls with the parabolic balls).

B Wavelet Expansion

The classical Fourier representation is designed to express a function in terms of sinusoidal functions $x \mapsto e^{2\pi i\xi x}$:

$$\hat{f}(\xi) = \int f(x)e^{-2\pi i\xi x} \; dx, \qquad f(x) = \check{\hat{f}} = \int \hat{f}(\xi)e^{2\pi i\xi x} \; d\xi.$$

It can be used to solve linear differential equation and translation invariant operator such as convolution. Though it suffers from the drawback the Fourier transform is a nonlocal operation. The theory of wavelets attempts to avoid this drawback by offering a representation of a function in terms of a two parameter family of translates and dilates of a fixed function of compact support. To explain this, we first describe *Mallat's multiresolution analysis* (in short MRA).

Definition B.1(i) Given $n \in \mathbb{Z}$, we set $\Lambda_n := 2^{-n}\mathbb{Z}$. Given a $L^2(\mathbb{R})$ function $\phi : \mathbb{R} \to \mathbb{R}$ with $\int \phi^2 \, dx = 0$, we set

$$\phi_x^{(n)}(y) = 2^{n/2}\phi(2^n(y-x)), \qquad \phi_x := \phi_x^{(0)},$$

for $x \in \Lambda_n$, and $n \in \mathbb{Z}$. We also write $V_n = V_n(\phi) := span\left\{\phi_x^{(n)} : \ x \in \Lambda_n\right\}$. Note

$$V_n = \{f : \ f(x) = g(2^n x) \ \text{ for some } g \in V_0\}.$$

(ii) We say ϕ is a *(father) wavelet* or a *scaling function*, if the following conditions are met:

(1) The family $\{\phi_x : \ x \in \mathbb{Z}\}$ is an orthonormal set.

(2) $V_0 \subset V_1$.

(3) $L^2(\mathbb{R}) = \overline{\cup_n V_n}$.

The second condition is equivalent to the existence of coefficients $\{a_r : \ r \in \mathbb{Z}\}$ such that

(B.1) $$\phi(x) = 2^{1/2}\sum_{r\in\mathbb{Z}} a_r \phi(2x-r).$$

Note

$$a_r = \int 2^{1/2}\phi(2x-r)\phi(x) \, dx, \qquad \sum_r a_r^2 = 1.$$

(2) is also equivalent to the property $V_n \subset V_{n+1}$ for all $n \in \mathbb{Z}$.

(iii) We define *mother wavelet* ψ by

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

where $b_r = (-1)^{-1}a_r$. We also set $\psi_x^{(n)}(y) = 2^{n/2}\psi(2^n(y-x))$.

(iv) Define the periodic functions

$$m_0(\xi) = 2^{-1/2}\sum_r a_r e^{2\pi i r x}, \qquad m_1(\xi) = 2^{-1/2}\sum_r b_r e^{2\pi i r x}.$$

Note that (B.1) and (B.2) mean

(B.3) $$\hat{\phi}(\xi) = m_0(\xi/2)\hat{\phi}(\xi/2), \qquad \hat{\psi}(\xi) = m_1(\xi/2)\hat{\phi}(\xi/2).$$

(v) The orthogonal projection onto V_n is denoted by $\mathcal{P}_n$:

$$\mathcal{P}_n f = \sum_{a\in\Lambda_n} \langle f, \phi_a^{(n)}\rangle \ \phi_a^{(n)}, \qquad \mathcal{Q}_n f := (\mathcal{P}_{n+1} - \mathcal{P}_n) f = \sum_{a\in\Lambda_n} \langle f, \psi_a^{(n)}\rangle \ \psi_a^{(n)}.$$

□

Proposition B.1 **(i)** *For every* ξ*,*

$$\sum_{r\in\mathbb{Z}} |\hat{\phi}(\xi + r)|^2 = \sum_{r\in\mathbb{Z}} |\hat{\psi}(\xi + r)|^2 = 1. \tag{B.4}$$

(ii) *The set* $\{\psi_x^{(n)} : x \in \Lambda_n\}$ *is an orthonormal set. Moreover, if* W_n *denotes their spans, then* $V_{n+1} = V_n \oplus W_n$*. As a consequence, we have the following decomposition:*

$$L^2 = \bigoplus_{n\in\mathbb{Z}} W_n = V_n \oplus W_n \oplus W_{n+1} \oplus \dots. \tag{B.5}$$

Proof(i) For (B.4), observe

$$\delta_{0,\ell} = \langle \phi, \phi_\ell \rangle = \langle \hat{\phi}, \bar{\hat{\phi}}_\ell \rangle = \int |\hat{\phi}(\xi)|^2 e^{2\pi i \ell \xi} \, d\xi = \int_0^1 \left[\sum_{n\in\mathbb{Z}} |\hat{\phi}(\xi + n)|^2 \right] e^{2\pi i \ell \xi} \, d\xi,$$

for every $\ell \in \mathbb{Z}$. Hence the periodic function $\sum_n |\hat{\phi}(\xi + n)|^2$ must be 1. In the same fashion we can show that $\sum_n |\hat{\phi}(\xi + n)|^2 = 1$.

(ii) It suffices to verify the claim of this theorem for $n = 1$. Evidently,

$$\phi_k(x) = 2^{1/2} \sum_r a_{r-2k} \; \phi(2x - r),$$

for every $k \in \mathbb{Z}$, and

$$\langle \psi, \phi_k \rangle = \sum_r b_r \; a_{r-2k} = \sum_r (-1)^r a_{1-r} \; a_{r-2k} = -\sum_{r'} (-1)^{r'} a_{r'-2k} a_{1-r'}.$$

where r' and r are related by $1 - r' = r - 2k$ or $1 - r = r' - 2k$. From this we learn that $\langle \psi, \phi_k \rangle = 0$. As a result, $\psi_\ell \in W_1$ for every $\ell \in \mathbb{Z}$. On the other hand

$$\langle \psi, \psi_\ell \rangle = \sum_r b_r b_{r-2\ell} = \sum_r a_{1-r} a_{1-r+2\ell} = \sum_r a_r a_{r+2\ell} = \langle \phi, \phi_{-\ell} \rangle = \delta_{0,\ell},$$

which confirms that the translates of ψ forms an orthonormal set. It remains to show that the translates of ψ span W_1. For this, it suffices to show

$$\phi^{(1)}(x) = 2^{1/2}\phi(2x) = \sum_{k\in\mathbb{Z}} (c_k \phi(x - k) + d_k \psi(x - k)), \tag{B.6}$$

where $c_k = \langle \phi^{(1)}, \phi_k \rangle$, and $d_k = \langle \phi^{(1)}, \psi_k \rangle$. To see this, observe that if $\tilde{\phi}(x) = \phi(-x)$, then

$$\begin{aligned} c_k =& \langle \hat{\phi}^{(1)}, \check{\phi}_k \rangle = \int \hat{\phi}^{(1)}(\xi) \bar{\hat{\phi}}(\xi) \ e^{2\pi i k\xi} \ d\xi = \frac{1}{2} \int \hat{\phi}(\xi/2) \bar{\hat{\phi}}(\xi/2) \bar{m}_0(\xi/2) \ e^{2\pi i k\xi} \ d\xi \\ =& \int |\hat{\phi}(\xi)|^2 \bar{m}_0(\xi) \ e^{4\pi i k\xi} \ d\xi = \int_0^1 \sum_{n\in\mathbb{Z}} |\hat{\phi}(\xi+n)|^2 \bar{m}_0(\xi) \ e^{4\pi i k\xi} \ d\xi \\ =& \int_0^1 \bar{m}_0(\xi) \ e^{4\pi i k\xi} \ d\xi = 2^{-1/2} a_{2k}. \end{aligned}$$

where we have used (B.3) and (B.4) for the fifth and the last equality. Likewise

$$\begin{aligned} d_k =& \langle \hat{\phi}^{(1)}, \check{\psi}_k \rangle = \int \hat{\phi}^{(1)}(\xi) \bar{\hat{\psi}}(\xi) \ e^{2\pi i k\xi} \ d\xi = \frac{1}{2} \int \hat{\phi}(\xi/2) \bar{\hat{\phi}}(\xi/2) \bar{m}_1(\xi/2) \ e^{2\pi i k\xi} \ d\xi \\ =& \int |\hat{\phi}(\xi)|^2 \bar{m}_1(\xi) \ e^{4\pi i k\xi} \ d\xi = \int_0^1 \sum_{n\in\mathbb{Z}} |\hat{\phi}(\xi+n)|^2 \bar{m}_1(\xi) \ e^{4\pi i k\xi} \ d\xi \\ =& \int_0^1 \bar{m}_1(\xi) \ e^{4\pi i k\xi} \ d\xi = 2^{-1} b_{2k} = 2^{-1/2} a_{1-2k}. \end{aligned}$$

As a result,

$$\sum_k (|c_k|^2 + |d_k|^2) = 2^{1/2} \sum_k (|a_{2k}|^2 + |a_{1-2k}|^2) = \sum_r ||a_r|^2 = \|\phi^1\|_{L^2}^2.$$

From this we can readily deduce (B.6). □

Example B.1(i) *(Haar Basis)* When $\phi = 1\!\!1_{[0,1]}$, then $\psi(x) = 1\!\!1(x \in [0,1/2)) - 1\!\!1(x \in [-1/2, 0)$. Moreover, $f \in V_n$ iff f is constant on dyadic intervals $[i2^{-n}, (i+1)2^{-n}$.

(ii) *(Shannon Basis)* This is basically the Haar basis in the frequency space. In other words, we choose ϕ such that $\hat{\phi} = 1\!\!1_{[-1/2,1/2]}$, so that

$$\phi(x) = \int_{-1/2}^{1/2} e^{2\pi i x\xi} \ d\xi = 2 \ \frac{\sin(\pi x)}{\pi x}.$$

□

A fundamental result of Daubechies [D] guarantees the existence of a regular wavelet basis.

Theorem B.1 *For every $r \in \mathbb{N}$, there exists a scaling function $\phi : \mathbb{R} \to \mathbb{R}$ with the following properties:*

(i) *The function ϕ is of compact support and in C^r.*

(ii) *The corresponding space V_0 includes all polynomials of degree r.*

Instead of an orthogonal basis, it is also useful to search for a biorthogonal basis. Such basis would be natural when we we wish to have a MRA for divergence-free or curl-free vector field. As a warm up, we make the following observation:

Proposition B.2 *Let (ϕ, ψ) be a scaling and wavelet functions of a MRA, and assume that $(\varphi, \varphi^*, \psi^*)$ is a pair of C^1 functions of compact support such that*

$$\phi^(x) = \varphi(x) - \varphi(x-1), \qquad \int_x^{x+1} \varphi(y)\ dy = \varphi^*(x), \qquad \psi^*(x) = \int_{-\infty}^x \psi. \tag{B.7}$$

Then $((\varphi_a, \varphi_a^*) :\ a \in \mathbb{Z})$ is a biorthogonal collection, and

$$\frac{d}{dx}\mathcal{P}_n f = \widehat{\mathcal{P}}_n\left(\frac{df}{dx}\right), \qquad \frac{d}{dx}\mathcal{Q}_n f = \widehat{\mathcal{Q}}_n\left(\frac{df}{dx}\right), \tag{B.8}$$

where

$$\widehat{\mathcal{P}}_n g = \sum_{a \in \Lambda_n} \langle g, \varphi_a^{*(n)}\rangle\ \varphi_a^{(n)}, \qquad \widehat{\mathcal{Q}}_n g = \sum_{a \in \Lambda_n} \langle g, \psi_a^{*(n)}\rangle\ \psi_a^{\prime(n)}.$$

Proof We note that if $\tilde{\varphi}(x) = \int_{-\infty}^x \phi$, then

$$\langle \tau_a\varphi, \varphi^*\rangle = \langle \tau_a\varphi, \tau_{-1}\tilde{\varphi} - \tilde{\varphi}\rangle = \langle \tau_{a+1}\varphi - \tau_a\varphi, \tilde{\varphi}\rangle = -\langle \tau_a\phi', \tilde{\varphi}\rangle = \langle \tau_a\phi, \tilde{\varphi}'\rangle = \langle \tau_a\phi, \tilde{\phi}\rangle = \delta_{0,a},$$

verifying the biorthogonal condition. On the other hand,

$$\begin{aligned}
\frac{d}{dx}\mathcal{P}_n f &= \sum_{a \in \Lambda_n} \left\langle f, \phi_a^{(n)}\right\rangle\ \frac{d}{dx}\phi_a^{(n)} = \sum_{a \in \Lambda_n} \left\langle f, \phi_a^{(n)}\right\rangle\ \left(\varphi_a^{(n)} - \varphi_{a+2^{-n}}^{(n)}\right) \\
&= \sum_{a \in \Lambda_n} \left\langle f, \phi_a^{(n)} - \phi_{a-2^{-n}}^{(n)}\right\rangle\ \varphi_a^{(n)} = -\sum_{a \in \Lambda_n} \left\langle f, \frac{d}{dx}\phi_a^{*(n)}\right\rangle\ \varphi_a^{(n)} \\
&= \sum_{a \in \Lambda_n} \left\langle \frac{d}{dx}f, \phi_a^{*(n)}\right\rangle\ \varphi_a^{(n)} = \widehat{\mathcal{P}}_n\left(\frac{df}{dx}\right).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\frac{d}{dx}\mathcal{Q}_n f &= \sum_{a \in \Lambda_n} \left\langle f, \psi_a^{(n)}\right\rangle\ \frac{d}{dx}\psi_a^{(n)} = -\sum_{a \in \Lambda_n} \left\langle f, \frac{d}{dx}\psi_a^{*(n)}\right\rangle\ \frac{d}{dx}\psi_a^{(n)} \\
&= \sum_{a \in \Lambda_n} \left\langle \frac{d}{dx}f, \psi_a^{*(n)}\right\rangle\ \frac{d}{dx}\psi_a^{(n)} = \widehat{\mathcal{Q}}_n\left(\frac{df}{dx}\right).
\end{aligned}$$

□

We may use the pair (ϕ,ψ) to build a similar multiresolution decomposition for $L^2(\mathbb{R}^d)$:
(i) Set $\Lambda_n^d = 2^{-n}\mathbb{Z}^d$. The collection of the maps $\phi_x^{d,n}:\mathbb{R}^d\to\mathbb{R}; x=(x_1,\dots,x_d)\in\Lambda_n^d$, of the form

$$\phi_x^{d,n}(y)=\phi_x^{d,n}(y_1,\dots,y_d)=\prod_{i=1}^d \phi_{x_i}^{(n)}(y_i),$$

forms an orthonormal set in $L^2(\mathbb{R}^d)$. The span of these maps are denoted by $V_n(\mathbb{R}^d)=V_n$.
(ii) Let Ψ denote the set of functions $\eta:\mathbb{R}^d\to\mathbb{R}$, defined by

$$\eta(x)=\prod_{i=1}^d \zeta_i(x_i),$$

where $\zeta_i\in\{\phi,\psi\}$ for each i, and that at least one of ζ_i equals to ψ. Given $\eta\in\Psi$ and $x\in\Lambda_n^d$, we define

$$\eta_x^{d,n}(y)=2^{nd/2}\eta(2^n(y-x)),$$

The collection $\{\eta_x^{d,n}:\ x\in\Lambda_n^d,\eta\in\Psi\}$ is an orthonormal set and its span is denoted by $W_n(\mathbb{R}^d)=W_n$. The decomposition (B.5) continues to be valid in higher dimensions. In particular, any $f\in L^2(\mathbb{R}^d)$ can be decomposed as

$$f=\sum_{x\in\Lambda_n^d}\langle f,\phi_x^{d,n}\rangle\phi_x^{d,n}+\sum_{m=n}^{\infty}\sum_{\eta\in\Psi}\sum_{x\in\Lambda_m^d}\langle f,\eta_x^{d,m}\rangle\eta_x^{d,m}, \tag{B.9}$$

where $\langle f,g\rangle$ is a short-hand for the inner product $\int fg\ dx$.

Let us note that as an immediate consequence of Theorem 2.1**(iv)**, we have

$$\int \eta_x^{d,n}(y)P(y)\ dy=0, \tag{B.10}$$

for every polynomial P of degree r.

Proposition B.3 *Let $(\phi,\psi,\varphi,\varphi^*,\psi^*)$ be as in Proposition B.2, and set*

$$\Phi(x)=(\varphi(x,1),\dots,\varphi(x,d),\qquad \Phi^*(x)=(\varphi^*(x,1),\dots,\varphi^*(x,d),$$

where

$$\varphi(x,i)=\varphi(x_i)\prod_{j\neq i}\phi(x_j),\qquad \varphi^*(x,i)=\varphi^*(x_i)\prod_{j\neq i}\phi(x_j).$$

Then Then $((\Phi_a,\Phi_a^*):\ a\in\mathbb{Z})$ is a biorthogonal collection, and

$$\nabla\mathcal{P}_n f=\widehat{\mathcal{P}}_n(\nabla f):=\left(\widehat{\mathcal{P}}_n^i(f_{x_i})\right), \tag{B.11}$$

where

$$\widehat{\mathcal{P}}_n^i g=\sum_{a\in\Lambda_n^d}\langle g,\varphi_a^{*(n)}(\cdot,i)\rangle\ \varphi_a^{(n)}(\cdot,i).$$

C Hopf Algebra

In this chapter, we give an overview of Hopf algebras.

Definition C.1(i) Given a field k, by a *k-algebra* A, we mean a k-vector space that is equipped with an associative product. There are two equivalent way of describing a product, either as a bilinear map $q : A \times A \to A$, or a linear map $m : A \otimes A \to A$, which is related to q by $m(a \otimes b) = q(a, b)$. Throughout, we adopt the latter. Sometimes we write $a \cdot_m b$ for $m(a \otimes b)$. By associativity, we mean such that $(a \cdot_m b) \cdot_m c = a \cdot_m (b \cdot_m c)$. We say that the algebra A is unital of it possesses a unit element $\mathbb{1}$. By an slight abuse of notation, we also write $\mathbb{1}$ for the map k to A that is defined by $\mathbb{1}(c) = c\mathbb{1}$.

(ii) Let A be an algebra with a product m. We may use (A, m) to define an algebra $(A \otimes A, m_2)$ with

(C.1) $$m_2\left((a_1 \otimes b_1) \otimes (a_2 \otimes b_2)\right) = m(a_1 \otimes a_2) \otimes m(b_1 \otimes b_2).$$

If (A, m) is unital with the unit $\mathbb{1}_A$, then $(A \otimes A, m_2)$ is also unital with $\mathbb{1}_{A\otimes A} = \mathbb{1}_A \otimes \mathbb{1}_A$. Note that we may write

$$m_2 = (m \otimes m) \circ (id \otimes \tau \otimes id),$$

where $\tau(a \otimes b) = b \otimes a$. □

Example C.1 Given a k-vector space, the tensor space

$$T(V) = \oplus_{n\in\mathbb{N}_0} V^{\otimes n},$$

is a unital algebra with the product $m(a, b) = a \otimes b$ (by $V^{\otimes 0}$ we mean the field k). If $\dim V = \ell$, and $\{e_1, \dots, e_\ell\}$ is a basis for V, then the set of $e_{i_1,\dots,i_k},\ i_1, \dots, i_k \in \{1, \dots, \ell\}$ forms a basis for $T(V)$, where $e_{i_1,\dots,i_k} = e_{i_1} \otimes \dots \otimes e_{i_k}$. Evidently,

$$e_{i_1,\dots,i_k} \otimes e_{i'_1,\dots,i'_k} = e_{i_1,\dots,i_k,i'_1,\dots,i'_k}.$$

We may regard $a = (i_1, \dots, i_k)$ as a word with letters in $\{1, \dots, \ell\}$. □

We next discuss *coalgebras*. To motivate the definition, imagine that A^* is a dual of A with a pairing $\langle\cdot,\cdot\rangle : A^* \times A \to \mathbb{R}$. This pairing induces a pairing

$$\langle\cdot,\cdot\rangle' : (A^* \otimes A^*) \times (A \otimes A) \to \mathbb{R},$$

such that

$$\langle f \otimes g, a \otimes b\rangle' = \langle f, a\rangle\langle g, b\rangle.$$

Now if we have a multiplication on A^*, say $\hat{m}$, we may attempt to turn $\hat{m}$ to a suitable operation on A:

(C.2) $$\langle \hat{m}(f \otimes g), a\rangle = \langle f \otimes g, \Delta a\rangle',$$

i.e., $\Delta = \hat{m}^*$. If we write

$$\Delta a = \sum_i b^i \otimes \hat{b}^i,$$

then

$$\langle \hat{m}(f \otimes g), a\rangle = \sum_i \langle f, b^i\rangle\langle g, \hat{b}^i\rangle.$$

Definition C.2(i) Given a field k, by a *k-coalgebra* A, we mean a k-vector space that is equipped with an associative coproduct Δ. By the latter, we mean that $\Delta : A \to A \otimes A$ is a linear map, such that if we define $\hat{m}$ by (C.2), then the pair $(A^*, \hat{m})$ is an algebra. The associativity of $\hat{m}$ yields a property that is referred to as the *coassociativity* of Δ. This can be expressed directly with no reference to the dual space:

$$(\Delta \otimes id) \circ \Delta = (id \otimes \Delta) \circ \Delta : A \to A \otimes A \otimes A.$$

We say that the coalgebra A is counital of it possesses a counit element $1\!\!1'$. By this we mean $1\!\!1' : A \to k$ is a linear map such that the maps $(id \otimes 1\!\!1') \circ \Delta$ (respectively $(1\!\!1' \otimes id) \circ \Delta$) is an isomorphism of $A \otimes k$ (respectively $k \otimes A$) and A. That is,

$$[(id \otimes 1\!\!1') \circ \Delta](a) = a \otimes 1, \qquad [(1\!\!1' \otimes id) \circ \Delta](a) = 1 \otimes a.$$

Here 1 is the unit element of k. Equivalently, if $\Delta a = \sum_i a^i \otimes b^i$, then

$$\text{(C.3)} \qquad \sum_i 1\!\!1'(a^i) \otimes b^i = \sum_i 1\!\!1'(a^i) b^i = a, \qquad \sum_i a^i \otimes 1\!\!1'(b^i) = \sum_i 1\!\!1'(b^i) a^i = a.$$

(ii) We can use Δ to define a coproduct on $A \otimes A$ by

$$\Delta_2 : A \otimes A \to A \otimes A \otimes A \otimes A, \qquad \Delta_2 = (id \otimes \tau \otimes id) \circ (\Delta \otimes \Delta).$$

Also if A is counital, we set

$$1\!\!1'_{A \otimes A} = 1\!\!1'_A \otimes 1\!\!1'_A : A \otimes A \to k \otimes k \equiv k.$$

(iii) Let $(A; m, 1\!\!1)$ be an algebra, and $(C; \Delta, 1\!\!1')$ be a coalgebra, and let $L_k(C, A) = L(C, A)$ be the space of linear maps from A to C. We may define a product on $L(C, A)$ by

$$F \star G = m \circ (F \otimes G) \circ \Delta.$$

We refer to $\star$ as a *convolution*; note

$$\text{(C.4)} \qquad \Delta a = \sum_i a^i \otimes b^i \quad \Longrightarrow \quad (F \star G)(a) = \sum_i m(F(a^i) \otimes G(b^i)).$$

Moreover, $\mathbb{1} \circ \mathbb{1}'$ is a unit element for the product $\star$: If $\Delta a = \sum_i a^i \otimes b^i$, then by (C.4), and (C.3)

$$(F \star (\mathbb{1} \circ \mathbb{1}'))(a) = \sum_i m(F(a^i) \otimes \mathbb{1}'(b^i)\mathbb{1}) = \sum_i \mathbb{1}'(b^i)F(a^i)$$
$$= F\left(\sum_i \mathbb{1}'(b^i)a^i\right) = F(a).$$

(iv) If $A = k$ in **(iii)**, we may identify $L_k(C, k)$ with C^*, the dual of C, where the pairing is $\langle f, a\rangle = f(a)$. The corresponding $\star = \Delta^*$ is denoted by $\star_\Delta$. □

Example C.2 We give a simple example of a coproduct on $T(V)$. To avoid confusion between the product $\otimes$ on $T(V)$ and our coproduct, we also use $\boxtimes$ for the tensor product, so that $\Delta : T(V) \to T(V) \boxtimes T(V)$. Assume that $\{e_1, \ldots, e_\ell\}$ is a basis for V, so that for $a = (i_1, \ldots, i_n)$,

$$e_a = e_{i_1} \otimes \cdots \otimes e_{i_n}.$$

We define a coproduct by

$$\Delta e_c = \sum_{(a,b)=c} e_a \boxtimes e_b.$$

□

Definition C.3(i) Assume that $(A; m, \mathbb{1})$ is a unital algebra, and $(A; \Delta, \mathbb{1}')$ is a counital coalgebra. Then $(A; m, \mathbb{1}; \Delta, \mathbb{1}')$ is a *bialgebra*, if Δ is an algebra homomorphism between $(A, m, \mathbb{1}_A)$, and $(A \otimes A, m_2, \mathbb{1}_{A\otimes A})$:

$$\text{(C.5)} \qquad \Delta(m(a \otimes b)) = m_2(\Delta(a) \otimes \Delta(b)), \qquad \Delta(\mathbb{1}_A) = \mathbb{1}_A \otimes \mathbb{1}_A.$$

Equivalently, m is a coalgebra homomorphism from $(A \otimes A, \Delta_2, \mathbb{1}'_{A\otimes A})$ to $(A, \Delta, \mathbb{1}'_A)$.

(ii) We say $(H; m, \mathbb{1}; \Delta, \mathbb{1}'; S)$ is a *Hopf algebra*, if $(H; m, \mathbb{1}; \Delta, \mathbb{1}')$ is a bialgebra, and S satisfies

$$\text{(C.6)} \qquad m \circ (id \otimes S) \circ \Delta = m \circ (S \otimes id) \circ \Delta = \mathbb{1} \circ \mathbb{1}' =: \mathbb{1}_\star.$$

Equivalently, $id \star S = S \star id = \mathbb{1}_\star$. We refer to such a map S as an *antipode*. Note,

$$\Delta a = \sum_i a^i \otimes \hat{a}^i \quad \Longrightarrow \sum_i Sa^i \cdot_m \hat{a}^i = \sum_i a^i \cdot_m S\hat{a}^i = (\mathbb{1}' \circ \mathbb{1})(a) = \mathbb{1}'(a)\mathbb{1}.$$

(iii) Let $(H; m, \mathbb{1}; \Delta, \mathbb{1}'; S)$ be Hopf algebra. We write $G(H)$ for the set of *characters* of H. By a character we mean a linear $g : H \to k$ such that $g(h_1 \cdot_m h_2) = g(h_1)g(h_2)$, and

$g(\mathbb{1}) = 1$. The latter condition is equivalent to $q \neq 0$ because the former implies that $g(h) = g(h \cdot_m \mathbb{1}) = g(h)g(\mathbb{1})$.

(iv) We say a bialgebra $(H; m, \mathbb{1}; \Delta, \mathbb{1}'; S)$ is *graded and connected* (or simply graded if there is no danger of confusion), if

$$H = \oplus_{n \in \mathbb{N}_0} H_n, \quad S : H_n \to H_n, \quad m : H_i \otimes H_j \to H_{i,j}, \quad \Delta : H_n \to \oplus_{i+j=n} H_i \otimes H_j,$$

and H_0 is spanned by $\mathbb{1}$. □

Example C.3(i) Let k be a field, and let $(G, \cdot)$ be a group. We write kG for the vector space freely generated by the elements of G. Define linear maps $m : (kG) \otimes (kG) \to kG$ and $\Delta : kG \to (kG) \otimes (kG)$ such that

$$m(g_1 \otimes g_2) = g_1 \cdot g_2, \quad \Delta(g) = g \otimes g,$$

for every $g_1, g_2, g \in G$. We can readily show that $(kG; m, 1; \Delta, \mathbb{1}')$ is a bialgebra, where 1 is the unit element of G, and $\mathbb{1}' : kG \to k$ is defined by

$$\mathbb{1}'(\sum_i c_i g_i) = \sum_i c_i,$$

for $c_i \in k, g_i \in G$. To see this, first observe that the corresponding product m_2 on $(kG) \otimes (kG)$ satisfies

$$m_2(g_1 \otimes g_2, h_1 \otimes h_2) = (g_1 \cdot h_1) \otimes (g_2 \cdot h_2).$$

As a result,

$$\Delta(g \cdot h) = (g \cdot h) \otimes (g \cdot h) = m_2(g \otimes g, h \otimes h) = m_2(\Delta g, \Delta h),$$

which implies that Δ is a homomorphism. Moreover, if we define a linear map $S : kG \to kG$ so that $S(g) = g^{-1}$, then S is an antipode because

$$m(S \otimes id)\Delta(g) = m(S \otimes id)(g \otimes g) = g^{-1} \cdot g = 1 = \mathbb{1} \circ \mathbb{1}'(g).$$

(ii) Let A^* be the space of smooth functions $f : \mathbb{R}^d \to \mathbb{R}$ with the pointwise multiplication. The differential operators $(\partial/\partial x_i : \; i = 1, \ldots, d)$ generate an algebra A that consists of the linear span of operators $(\partial^k : \; k \in \mathbb{N}_0^d)$, with the composition of operators playing the role of multiplication:

$$m(\partial^k \otimes \partial^\ell) = \partial^k \circ \partial^\ell = \partial^{k+\ell}.$$

Using the pairing

$$\langle f, D \rangle = (Df)(0),$$

we can turn the pointwise multiplication to a coproduct Δ on A. From

$$\langle fg, \partial^k\rangle = \partial^k(fg)(0) = \sum_{s+\ell=k} (\partial^s f)(0)(\partial^\ell g)(0) = \sum_{s+\ell=k} \langle f\otimes g, \partial^s\otimes\partial^\ell\rangle',$$

we learn

$$\Delta(\partial^k) = \sum_{s+\ell=k} \partial^s\otimes\partial^\ell.$$

It is straightforward to check that $(A;\circ, id;\Delta, id')$ is a bialgebra, where $id'(\partial^k) = \delta_{0,k}$. Moreover, if we define a linear $S: A\to A$, by $S(\partial^k) = (-1)^{|k|}\partial^k$, then

$$[m\circ(id\otimes S)\circ\Delta](\partial^k) = [m\circ(id\otimes S)]\sum_{s+\ell=k}\partial^s\otimes\partial^\ell = \sum_{s+\ell=k}(-1)^\ell\partial^k = 0,$$

whenever $k\neq 0$.

(iii) Let V be a k-vector space with a basis $\{e_1,\dots,e_\ell\}$. Recall that if $I = \{1,\dots,\ell\}$, then

$$\{e_a : a\in\emptyset\cup I\cup\dots\cup I^n\dots\},\qquad e_\emptyset = 1,\qquad e_{(i_1,\dots,i_n)} = e_{i_1}\otimes\dots\otimes e_{i_n},$$

is a basis for $T(V)$. Note that $H = T(V)$ is graded with $H_n = T_n(V)$ which is spanned by $\{e_a : \ a\in I^n\}$. We consider a Hopf algebra $(T(V);\sqcup\!\sqcup, \mathbb{1};\Delta,\mathbb{1}';S)$. Here the *shuffle product* $\sqcup\!\sqcup$ is defined by

$$e_a \sqcup\!\sqcup e_b = \sum_{c\in Sh(a,b)} e_c,$$

where $Sh(a,b)$, $a = (i_1,\dots,i_k)$, $b = (j_1,\dots,j_{k'})$ is the set of permutations c of the index set $(a,b) = (i_1,\dots,i_k,j_1,\dots,j_{k'})$ which preserves the original ordering of a and b. For example,

$$e_{i,j}\sqcup\!\sqcup e_{i',j'} = e_{i,j,i',j'} + e_{i,i',j,j'} + e_{i,i',j',j} + e_{i',i,j,j'} + e_{i',i,j',j} + e_{i',j',i,j}.$$

Note that the cardinality of the set $Sh(a,b)$ is $\binom{|a|+|b|}{|a|}$, where $|a|$ denotes the length of a. The unit is $\mathbb{1} = e_\emptyset = 1$. The coproduct is defined as

$$\Delta(e_a) = \Delta(e_{(i_1,\dots,i_n)}) = e_a\boxtimes\mathbb{1} + \mathbb{1}\boxtimes e_a + \sum_{k=1}^{n-1} e_{(i_1,\dots,i_k)}\boxtimes e_{(i_{k+1},\dots,i_n)}. \tag{C.7}$$

As for the counit, we simply have $\mathbb{1}'(e_a) = \delta_{a,\emptyset}$. Finally $S(e_{(i_1,\dots,i_n)}) = (-1)^n S(e_{(i_n,\dots,i_1)})$, or more generally

$$S(v_1\otimes\dots\otimes v_n) = (-1)^n S(v_n\otimes\dots\otimes v_1).$$

To verify (C.5), observe that if $a\in I^m$, and $b\in I^n$, then

$$\Delta(e_a\sqcup\!\sqcup e_b) = \sum_{c\in Sh(a,b)}\Delta(e_c) = \sum_{c\in Sh(a,b)}\sum_{i=0}^{m+n} e_{c^i}\boxtimes e_{\hat{c}^i}, \tag{C.8}$$

where c^i and $\hat{c}^i$ represent the terms that appear on the right-hand side of (C.7). On the other hand, $\sqcup\!\sqcup$ induces a product $\sqcup\!\sqcup_2$ which is isomorphic to the shuffle product on $V \boxtimes V$. By definition,

$$(e_a \boxtimes e_{a'}) \sqcup\!\sqcup_2 (e_b \boxtimes e_{b'}) = (e_a \sqcup\!\sqcup e_b) \boxtimes (e_{a'} \sqcup\!\sqcup e_{b'}).$$

Hence

$$\begin{aligned} \Delta(e_a) \sqcup\!\sqcup_2 \Delta(e_b) &= \left(\sum_{i=0}^{m} e_{a^i} \boxtimes e_{\hat{a}^i} \right) \sqcup\!\sqcup_2 \left(\sum_{j=0}^{n} e_{b^j} \boxtimes e_{\hat{b}^j} \right) \\ &= \sum_{i=0}^{m} \sum_{j=0}^{n} (e_{a^i} \sqcup\!\sqcup e_{b^j}) \boxtimes (e_{\hat{e}^i} \sqcup\!\sqcup e_{\hat{b}^j}) \\ &= \sum_{i=0}^{m} \sum_{j=0}^{n} \sum_{c_{ij} \in Sh(a^i, b^j)} \sum_{\hat{c}_{ij} \in Sh(\hat{a}^i, \hat{b}^j)} e_{\hat{c}_{ij}} \boxtimes e_{\hat{c}_{ij}}. \end{aligned} \tag{C.9}$$

It is not hard to show that there is one-to-one correspondence between the right-hand side of (C.8) and (C.9).

We think of I as the set of alphabet, and a as a word. When there is no danger of confusion we write a for e_a. Also, we write $a_j = (i_1, \ldots, i_j)$, $\hat{a}_j = (i_{j+1}, \ldots, i_n)$, when $a = (i_1, \ldots, i_n)$, and write $a\ell$ for $(i_1, \ldots, i_n, \ell)$. Note that with these conventions,

$$(ak) \sqcup\!\sqcup (b\ell) = (a \sqcup\!\sqcup (b\ell))\, k + ((ak) \sqcup\!\sqcup b)\, \ell. \tag{C.10}$$

We now show that S is an antipode. For this, we show that if $a \neq \emptyset$, then $m \circ (S \boxtimes id) \circ \Delta(a) = 0$. We may verify this by induction on $|a|$. The case $|a| = 1$ is strightforward. When $|a| = n + 2 \geq 2$, we write $a = kb\ell =: a'\ell =: ka''$ with $k, \ell \in I$. Observe

$$\begin{aligned} (m \circ (S \boxtimes id) \circ \Delta)(a) =& m \circ (S \boxtimes id) \left(a \boxtimes \mathbb{1} + \mathbb{1} \boxtimes a + \sum_{j=0}^{n} \left((kb_j) \boxtimes (\hat{b}_j \ell) \right) \right) \\ =& m \left((Sa) \boxtimes \mathbb{1} + \mathbb{1} \boxtimes a + \sum_{j=0}^{n} \left(S(kb_j) \boxtimes (\hat{b}_j \ell) \right) \right) \\ =& Sa + a + \sum_{j=0}^{n} \left(S(kb_j) \sqcup\!\sqcup (\hat{b}_j \ell) \right) \\ =& Sa + a + \sum_{j=0}^{n} \left(S(kb_j) \sqcup\!\sqcup \hat{b}_j \right) \ell - \sum_{j=0}^{n} \left((Sb_j) \sqcup\!\sqcup (\hat{b}_j \ell) \right) k \\ =& -(Sa'')k + a'\ell + \left(\sum_{j=0}^{n} (Sa'_j) \sqcup\!\sqcup \hat{a}'_j \right) \ell - \left(\sum_{j=0}^{n} (Sa''_j) \sqcup\!\sqcup \hat{a}''_j \right) k \\ =& ((m \circ (S \boxtimes id) \circ \Delta)(a'))\, \ell - ((m \circ (S \boxtimes id) \circ \Delta)(a''))\, k = 0, \end{aligned}$$

where we used $S(kb_j) = -(Sb_j)k$ and (C.13) for the fourth equality, and we used the induction hypothesis for the last equality. □

Proposition C.1 *When* $(H; m, \mathbb{1}; \Delta, \mathbb{1}'; S)$ *is a Hopf algebra, then* $(H^*; \Delta^*, m^*, \mathbb{1}'^*; m^*, \mathbb{1}^*; S^*)$ *is also a Hopf algebra. Given* $g \in H^*$*, set* $\Gamma_g = \Lambda_g^* \in L(H) = L(H, H)$*, where* $\Lambda_g(f) = f \bullet g$*.*

(i) *The set of characters* $(G(H), \Delta^*)$ *is a group, with* $g^{-1} = g \circ S$*, and the unit* $\mathbb{1}'^*$*.*

(ii) *The map* $\Gamma : (H^*, \bullet) \to (L(H), \circ)$ *is a homomorphism i.e.,*

$$\Gamma_{g_1 \bullet g_2} = \Gamma_{g_1} \circ \Gamma_{g_2}. \tag{C.11}$$

Moreover, $\Gamma_g = (id \otimes g)\Delta$*.*

(iii) *Define*

$$Z(H) = \{h \in H : \ g(h) = 0 \ \ \textit{for all } g \in G(H)\}.$$

Then

$$G_g(h_1 \cdot_m h_2) - \Gamma_g(h_1) \cdot_m G_g(h_2) \in Z(H). \tag{C.12}$$

Proof(i) To ease the notation we write

$$g_1 \bullet g_2 := \Delta^*(g_1 \otimes g_2), \quad h_1 \cdot h_2 := m(h_1, h_2), \quad k_1 \cdot_2 k_2 = m_2(k_1, k_2).$$

Take $h_1, h_2 \in H$, and write

$$\Delta h_1 = \sum_i h_1^i \otimes \hat{h}_1^i, \quad \Delta h_2 = \sum_i h_2^i \otimes \hat{h}_2^i.$$

Assume that $g_1, g_2 \in G(H)$. We have

$$\begin{aligned}
\langle g_1 \bullet g_2, h_1 \cdot h_2 \rangle =& \langle g_1 \otimes g_2, \Delta(h_1 \cdot h_2) \rangle' = \langle g_1 \otimes g_2, \Delta(h_1) \cdot_2 \Delta(h_2) \rangle' \\
=& \sum_{i,j} \langle g_1 \otimes g_2, (h_1^i \otimes \hat{h}_1^i) \cdot_2 (h_2^j \otimes \hat{h}_2^j) \rangle' \\
=& \sum_{i,j} \langle g_1 \otimes g_2, (h_1^i \cdot h_2^j) \otimes (\hat{h}_1^i \cdot \hat{h}_2^j) \rangle' \\
=& \sum_{i,j} \langle g_1, h_1^i \cdot h_2^j \rangle \ \langle g_2, \hat{h}_1^i \cdot \hat{h}_2^j \rangle \\
=& \sum_{i,j} \langle g_1, h_1^i \rangle \langle g_1, h_2^j \rangle \langle g_2, \hat{h}_1^i \rangle \langle g_2, \hat{h}_2^j \rangle \\
=& \sum_{i,j} \langle g_1 \otimes g_2, h_1^i \otimes \hat{h}_1^i \rangle' \ \langle g_1 \otimes g_2, h_2^j \otimes \hat{h}_2^j \rangle' \\
=& \langle g_1 \otimes g_2, \Delta h_1 \rangle' \ \langle g_1 \otimes g_2, \Delta h_2 \rangle' = \langle g_1 \bullet g_2, h_1 \rangle \ \langle g_1 \bullet g_2, h_2 \rangle,
\end{aligned}$$

which implies that $g_1 \bullet g_2 \in G(H)$. On the other hand, if $g \in G(H)$, and $\Delta h = \sum_i h^i \otimes \hat{h}^i$,

$$\begin{aligned}\langle g \bullet (g \circ S), h\rangle &= \langle g \otimes (g \circ S), \Delta h\rangle' = \sum_i \langle g \otimes (g \circ S), h^i \otimes \hat{h}^i\rangle' \\ &= \sum_i \langle g, h^i\rangle \langle g \circ S, \hat{h}^i\rangle = \sum_i \langle g, h^i\rangle \langle g, S(\hat{h}^i)\rangle \\ &= \sum_i \langle g, h^i \cdot S(\hat{h}^i)\rangle = \left\langle g, \sum_i h^i \cdot S(\hat{h}^i)\right\rangle = \mathbb{1}'(h)\langle g, \mathbb{1}\rangle = \mathbb{1}'(h),\end{aligned}$$

as desired.

(ii) We certainly have

$$\langle f, (\Gamma_{g_1} \circ \Gamma_{g_2})h\rangle = \langle (\Lambda_{g_2} \circ \Lambda_{g_1})f, h\rangle = \langle (f \bullet g_1 \bullet g_2, h\rangle = \langle \Lambda_{g_1 \bullet g_2} f, h\rangle = \langle f, \Gamma_{g_1 \bullet g_2} h\rangle,$$

proving (C.11). As for (C.12), observe that if $\Delta h = \sum_i h^i \otimes \hat{h}^i$, then

$$\begin{aligned}\langle f, \Gamma_g(h)\rangle = \langle \Lambda_g(f), h\rangle = \langle f \bullet g, h\rangle = \langle f \otimes g, \Delta h\rangle' &= \sum_i f(h^i) g(\hat{h}^i) \\ = f\left(\sum_i g(\hat{h}^i) h^i\right) = f\left((id \otimes g) \sum_i h^i \otimes \hat{h}^i\right) &= \langle f, (id \otimes g)\Delta h\rangle,\end{aligned}$$

as desired.

(iii) If g_1, g_2, h_1 and h_2 are as in **(i)**, then

$$\begin{aligned}\langle g_1, \Gamma_{g_2}(h_1) \cdot \Gamma_{g_2}(h_2)\rangle &= \left\langle g_1, \left(\sum_i g_2(\hat{h}_1^i) h_1^i\right) \cdot \left(\sum_j g_2(\hat{h}_2^j) h_2^j\right)\right\rangle \\ &= \sum_{i,j} g_2(\hat{h}_1^i) g_2(\hat{h}_2^j)\ \langle g_1, h_1^i \cdot h_2^j\rangle \\ &= \sum_{i,j} \langle g_2, \hat{h}_1^i\rangle\ \langle g_2, \hat{h}_2^j\rangle\ \langle g_1, h_1^i\rangle\ \langle g_1, h_2^j\rangle \\ &= \langle g_1 \bullet g_2, h_1 \cdot h_2\rangle = \langle g_1, \Gamma_{g_2}(h_1 \cdot h_2)\rangle.\end{aligned}$$

Here for we used our calculation in part **(i)** for the last equality. □

Proposition C.2 *If $(H; m, \mathbb{1}; \Delta, \mathbb{1}')$, with $H = \oplus_{n \in N_0} H_n$ is a graded bialgebra, then there exists a unique $S : H \to S$ such that $(H; m, \mathbb{1}; \Delta, \mathbb{1}'; S)$ is a Hopf algebra. Moreover, if $h \in H_n$, then*

$$Sh = \sum_{m=0}^{n} (\mathbb{1}_\star - id)^{\star m}. \tag{C.13}$$

D Spectral Gap and Logarithmic Sobolev Inequality

Consider a probability measure on $\mathbb{R}^N$ of the form

$$\mu(dx) = Z^{-1}e^{-H(x)}\ dx,$$

and consider the inner product

$$\langle f, g\rangle = \int fg\ d\mu,$$

on the Hilbert space $L^2(\mu)$. Consider a diffusion with the generator

$$\mathcal{L} = -\sum_i \partial_i^* \partial_i,$$

where ∂_i is a differential operator, and ∂_i^* is the adjoint of ∂_i with respect to the inner product $\langle\cdot,\cdot\rangle$. The operator $-\mathcal{L}$ is non-negative definite, and

$$\mathcal{D}(f) = \|f\|_1^2 = \langle \mathcal{L}f, f\rangle = \sum_i (\partial_i f)^2\ d\mu.$$

In the case of Ginzburg-Landau model, we choose

$$H(x) = \sum_{i=1}^N V(x_i),$$

and consider a periodic lattice $\mathbb{Z}_N = \{0, 1, \ldots, N\}$, with $0 = N$, and $\partial_i = D_{i,i+1}$.

Using and idea of Bakry and Emery, we can prove a Poincare-type bound

$$\int f^2\ d\mu - \left(\int f\ d\mu\right)^2 \leq c\mathcal{D}(f), \tag{D.1}$$

for a reversible Markov process by bounding from above the Sobolev's H^1 norm by the H^2-norm. To motive this idea, observe that if $\|\cdot\|_s$ denotes the H_s Sobolev norm, and T_t for the semigroup $e^{t\mathcal{L}}$, then

$$\frac{d}{dt}\|T_t f\|_0^2 = 2\langle T_t f, \mathcal{L}T_t f\rangle = -\|T_t f\|_1^2, \qquad \frac{d}{dt}\|T_t f\|_1^2 = -2\langle \mathcal{L}T_t f, \mathcal{L}T_t f\rangle = -\|T_t f\|_2^2.$$

From this it is not hard to see that a bound of the form

$$\|g\|_1^2 \leq c\|g\|_2^2, \tag{D.2}$$

would lead to an exponential convergence of $T_t f$ to $\int f\ d\mu$ in L^2 sense. Alternatively, we may use the spectral representation of the symmetric operator $\mathcal{L}$ to show that (D.3) and (D.2) are equivalent. A mollification of the above calculation leads to a logarithmic Sobolev inequality.

Theorem D.1 *If $D^2V(x) \geq c^{-1}$, then*

$$\mathcal{H}(f) = \int f \log f \; d\mu \leq 4N^2 c\mathcal{D}(f). \tag{D.3}$$

Proof. Let us write $T^t = e^{tL}$ for the semigroup associated with the generator cL. To ease the notation, we write

$$\nabla := (\partial_1, \ldots, \partial_{N-1}).$$

This allows us to write

$$\mathcal{L} = \nabla \cdot \nabla - (\nabla H) \cdot \nabla.$$

Define

$$\begin{aligned}
2\Gamma_1(f,g) :&= \mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f = 2\nabla f \cdot \nabla g, \\
2\Gamma_2(f,g) :&= \mathcal{L}\Gamma_1(f,g) - \Gamma_1(\mathcal{L}f, g) - \Gamma_1(f, \mathcal{L}g) = 2\sum_{i,j} \partial_{ij} f \partial_{ij} g + 2(\nabla^2 H)\ \nabla f \cdot \nabla g,
\end{aligned}$$

$$\mathcal{L}f/f = \mathcal{L}\log f + |\nabla \log f|^2, \quad \int \Gamma_1(f,g)\ d\mu = \int \nabla f \cdot \nabla g \ d\mu = -\int f\mathcal{L}g \ d\mu,$$

where $\partial_{ij} = \partial_i \circ \partial_j$. Now if f is a probability density with respect to μ and $f_t = T_t f$, $h(t) = \int f_t \log f_t d\mu$, then

$$\begin{aligned}
h'(t) &= \int (\log f_t)\mathcal{L}f_t \ d\mu = -\int \Gamma_1(f_t, \log f_t)\ d\mu, \\
h''(t) &= -\int \left(\Gamma_1(\mathcal{L}f_t, \log f_t) + \Gamma_1(f_t, \mathcal{L}f_t/f_t)\right)\ d\mu \\
&= \int \big(\mathcal{L}f_t \cdot \mathcal{L}\log f_t - \Gamma_1(f_t, \mathcal{L}\log f_t) - \Gamma_1(f_t, |\nabla \log f_t|^2)\big)\ d\mu \\
&= -\int \big(2f_t\Gamma_1(\log f_t, \mathcal{L}\log f_t) + \Gamma_1(f_t, |\nabla \log f_t|^2)\big)\ d\mu \\
&= \int \big(-2f_t\Gamma_1(\log f_t, \mathcal{L}\log f_t) + f_t\mathcal{L}|\nabla \log f_t|^2)\big)\ d\mu \\
&= \int f_t\Gamma_2(\log f_t, \log f_t)\ d\mu \geq \int (\nabla^2 H)\nabla f_t \cdot \nabla f_t \ f_t^{-1} d\mu.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\partial_i H(x) = V'(x_i) - V'(x_{i+1}), \quad \partial_{ii} H(x) = V''(x_i) + V''(x_{i+1}), \\
&\partial_i \partial_{i+1} H(x) = \partial_{i+1}\partial_i H(x) = -V''(x_{i+1}),
\end{aligned}$$

and $\partial_i \partial_j H(x) = 0$ if $|i-j| > 1$. Hence, if $\sum_i g_i = 0$, then

$$\begin{aligned}\int (\nabla^2 H)\nabla g \cdot g &= \sum_i (V''(x_i) + V''(x_{i+1}))g_i^2 - 2V''(x_{i+1})g_i g_{i+1} \\ &= \sum_i V''(x_{i+1})(g_{i+1}-g_i)^2 \geq c^{-1}\sum_i (g_{i+1}-g_i)^2 \\ &\geq 2c^{-1}N^2 \sum_i g_i^2,\end{aligned}$$

because

$$(g_j - g_0)^2 \leq N \sum_i g_i^2, \quad g_0 = N^{-1}\sum_i (g_0 - g_i).$$

As a result,

$$\int (\nabla^2 H)\nabla g \cdot g \geq 2c^{-1}N^2 \int \Gamma_1(f_t, f_t) f_t^{-1} \; d\gamma = -c^{-1}h'(t).$$

Hence

$$\int |\nabla f|^2/f \; d\gamma = -h'(0) \geq h'(t) - h'(0) \geq 2c^{-1}N^2(h(0) - h(t)),$$

and this implies (D.3) provided that we can show that $\lim_{t\to\infty} h(t) = 0$ for a subsequence. To see this, first observe that if $g_t = \sqrt{f_t}$, then $\int g_t^2 d\gamma = 1$ and $\int_0^\infty \int |\nabla g_t|^2 \; d\mu dt < \infty$. Hence for some $t_n \to \infty$, we have that $\int |\nabla g_{t_n}|^2 \; d\gamma \to 0$ as $n \to \infty$. From this, we deduce that $g_{t_n} \to 1$ in $\mathcal{L}^2(\mu)$ by Rellich's theorem. Hence $f_t \to 0$ almost everywhere along a subsequence. Note that if we assume that f is bounded, then $\{f_t\}$ is uniformly bounded in t and we may use the Bounded Convergence Theorem to deduce that $\lim_{t\to\infty} h(t) = 0$ for a subsequence. This implies LSI in the case of the bounded f. The general f can be treated by a truncation. For example, for every ℓ, choose a smooth non-decreasing function ϕ_ℓ such that $\phi_\ell(f) = f$, for $f \leq \ell$, $\phi_\ell(f) = \ell + 1$, for $f \geq \ell + 2$, $\phi' \leq 1$ everywhere, and $\phi_\ell(f) \geq (\ell+1)f/(\ell+2)$, for $f \leq \ell + 2$. Given a density function f, we set $f^\ell = \phi_\ell(f)$ and apply LSI to f^ℓ. We then send $\ell \to \infty$ to establish LSI for arbitrary f. □

E Central Limit Theorem for Markov Processes

Let $\mathbf{x} = (x(t) : \; t \in \mathbb{R})$ be a continuous time Markov process with a polish state space E, the transition kernel $p_t(x, dy)$, and the infinitesimal generator $\mathcal{L}$:

$$T_t f(x) = e^{t\mathcal{L}} f(x) = \mathbb{E}^{x(0)=x} f(x(t)) = \int f(y) \; p_t(x, dy).$$

Assume that π is an ergodic invariant measure for the process $x(t)$. By the classical ergodic theorem,

$$\lim_{t\to\infty} t^{-1}\int_0^t \delta_{x(s)}\ ds = \pi. \tag{E.1}$$

We are interested in the question of central limit theorem for the convergence in (E.1). In other words, we wish to study

$$\lim_{\varepsilon\to 0} m_\varepsilon := \lim_{\varepsilon\to\infty} \varepsilon^{-1/2}\left[\varepsilon\int_0^{t/e} \delta_{x(s)}\ ds - t\pi\right]. \tag{E.2}$$

Let us write $\mathcal{L} = \mathcal{A} + \mathcal{S}$, with $\mathcal{S}$ symmetric and $\mathcal{A}$ antisymmetric. Define

$$\mathcal{D}(f) := \|f\|_1^2 := \langle -\mathcal{L}f, f\rangle = \langle -\mathcal{S}f, f\rangle.$$

This is well defined for $f \in \mathcal{D} := \mathcal{D}(\mathcal{L}) \cap \mathcal{D}(\mathcal{L}^*)$, where $\mathcal{D}$ denotes the core (so that the graph of the operator is closed). The completion of $(\mathcal{D}, \|\cdot\|_1)$ is denoted by $\mathcal{H}_1$. Note that $\|\cdot\|_1$ is only a semi-norm. We can turn it to a norm by first defining equivalent relation $f \equiv g$ iff $\|f - g\|_1 = 0$, and then taking the equivalent classes. (In the case of ergodic invariant measure, $\|h\| = 0$, means that h is a constant.) We also set $\mathcal{H}_0 = L^2(\pi)$, and write $\|\cdot\|_0$ for the L^2-norm. The dual space $\mathcal{H}_{-1}$ is also Hilbert space with the norm

$$\|f\|_{-1}^2 = \sup_{g\in\mathcal{D}}\left(2\langle f, g\rangle - \|g\|_1^2\right).$$

Note that for $h \in \mathcal{D}$, we always have $f = \mathcal{S}h \in \mathcal{H}_{-1}$, and that $\|f\|_{-1} = \|h\|_1$. One can show that the set of such f is dense in $\mathcal{H}_{-1}$. Indeed the map $\mathcal{S} : \mathcal{H}_1 \to \mathcal{H}_{-1}$ is an isometry.

A standard trick for CLT in our context is that if V is in the range of $\mathcal{L}$, i.e. $-\mathcal{L}u = V$ for some $u \in \mathcal{D}$, then we can express

$$X(t) := \int_0^t V(x(s))\ ds = M(t) - u(x(t)) + u(x(0)),$$

for a martingale M. With a mild regularity on u, we can guarantee

$$\varepsilon^{1/2}X(t/\varepsilon) = \varepsilon^{1/2}M(t/\varepsilon) + O(\varepsilon^{1/2},$$

which reduces the question of CLT for X to that of M. Since in general we do not expect to have a nice solution u for our Poisson-like equation, we switch from $(-\mathcal{L})^{-1}$ to the resolvent $(\lambda - \mathcal{L})^{-1}$. In other words, we find u_λ such that

$$\lambda u_\lambda - \mathcal{L}u_\lambda = V. \tag{E.3}$$

From this we deduce

$$\lambda\|u_\lambda\|_0^2 + \|u_\lambda\|_{H^1}^2 = \langle u_\lambda, V\rangle \le \|u_\lambda\|_{H^1}\|V\|_{H^{-1}}.$$

Hence

$$\lambda\|u_\lambda\|_0^2 + \|u_\lambda\|_{H^1}^2 \le \|V\|_{H^{-1}}^2. \tag{E.4}$$

We then establish a CLT for X by showing the the corresponding martingales $(M^\lambda:\ \lambda > 0)$ form a tight sequence as $\lambda \to 0$. Though for the tightness we need more than (E.8).

Theorem E.1 *Let $V \in C_b(E)$ be such that $\int V\ d\pi = 0$, and define the family $u_\lambda,\ \lambda > 0$ as in* (E.7)*.*

(i) *If $\lambda\|u_\lambda\|_0^2 \to 0$, in low λ limit, and there exists $u \in \mathcal{H}_1$ such that*

$$\lim_{\lambda\to 0}\|u_\lambda - u\|_1 = 0. \tag{E.5}$$

Then the process

$$X^\varepsilon(t) = \varepsilon^{1/2}\int_0^{t/\varepsilon} V(x(s))\ ds,$$

converges to a Brownian B motion (with possibly degenerate variance).

(ii) *(Kipnis-Varadhan) When $\mathcal{A} = 0$, and $\|V\|_{-1} < \infty$, the conditions of part* **(i)** *are always met, and*

$$\mathbb{E}B(t)^2 = 2t\|V\|_{H^{-1}}^2.$$

(iii) *More generally, assume that there exists a constant C_0 such that*

$$\langle f, -\mathcal{L}g\rangle \le C_0\|f\|_1\|g\|_1, \tag{E.6}$$

for all $f, g \in \mathcal{D}$. The the conditions of part **(i)** *are met whenever $\|V\|_{-1} < \infty$.*

Proof(i)

(ii) When $\mathcal{A} = 0$, we have

$$\pi(dx)p_t(x, dy) = \pi(dy)p_t(y, dx).$$

This is equivalent to saying that the bilinear form

$$\mathcal{D}(f, g) = -\int f\mathcal{L}g\ d\pi,$$

is symmetric. Let us write $\mathbb{E}^\pi$ for the expected value when $x(0)$ is distributed according to the measure π. We can write

$$\begin{aligned}\mathbb{E}^\pi\left[t^{-1/2}\int_0^t V(x(s))\ ds\right]^2 =&2\mathbb{E}^\pi\ t^{-1}\iint_0^t V(x(s_1))V(x(s_2))\ 1\!\!1(s_2\ge s_1)\ ds_1ds_2\\ =&2\mathbb{E}^\pi\ t^{-1}\iint V(x(s))V(x(s+\theta))1\!\!1\left(s\in[0,t],\ s+\theta\in[0,t],\ \theta\ge 0\right)\ dsd\theta\\ =&2\mathbb{E}^\pi\ t^{-1}\iint V(x(0))V(x(\theta))1\!\!1\left(s\in[0,t],\ s+\theta\in[0,t],\ \theta\ge 0\right)\ dsd\theta\\ =&2\mathbb{E}^\pi\ t^{-1}\int_0^t (t-\theta)V(x(0))V(x(\theta))\ d\theta = 2\int_0^t\left(1-\frac{s}{t}\right)\langle V,e^{s\mathcal{L}}V\rangle\ ds.\end{aligned}$$

From this and the monotone convergence theorem we deduce

$$\lim_{t\to\infty}\mathbb{E}^\pi\left[t^{-1/2}\int_0^t V(x(s))\ ds\right]^2 = 2\int_0^\infty \langle V,e^{s\mathcal{L}}V\rangle\ ds = 2\langle V,(-\mathcal{L})^{-1}V\rangle =: 2\|V\|^2_{H^{-1}}, \tag{E.7}$$

provided that $\|V\|_{H^{-1}}<\infty$. For such a function, we expect that a CLT to hold as the following result of Kipnis and Varadhan [KV] confirms.

Observe that if there exists a function $u\in C_b(E)$ in the domain of the definition $\mathcal{L}$ such that $-\mathcal{L}u=V$, then

$$\|V\|^2_{H^{-1}} = -\langle \mathcal{L}u,u\rangle = \mathcal{D}(u) =: \|u\|^2_{H^1}<\infty,$$

provided that $\mathcal{D}(u)<\infty$. In this case one can establish Theorem 6.1 by observing that if

$$M(t)=u(x(t))-u(x(0))-\int \mathcal{L}u(x(s))\ ds,\qquad N(t)=M(t)^2-\int_0^t(\mathcal{L}u^2-2u\mathcal{L}u)(x(s))\ ds,$$

then the process $M(t)$ and $N(t)$ are martingales, and

$$\varepsilon^{1/2}\int_0^{t/\varepsilon}V(x(s))\ ds = \varepsilon^{1/2}M(t/\varepsilon)+E_\varepsilon(t) =: M_\varepsilon(t)+E_\varepsilon(t), \tag{E.8}$$

with $E_\varepsilon(t)=\varepsilon^{1/2}(u(x(0))-u(x(t)))=O(\varepsilon^{1/2})$. Note that by (E.8), we only need to establish a CLT for M_ε, and with the aid of Doob's inequality we can establish the tightness of the family $\{M_\varepsilon:\ \varepsilon>0\}$, and if M is a limit point of this family, then the processes $M(t)$, and

$$N(t):=M(t)^2-t\|u\|^2_{H^1}=M(t)^2-2t\|V\|^2_{H^{-1}},$$

are martingales, because by the ergodic theorem,

$$\begin{aligned}\lim_{\varepsilon\to 0}\varepsilon\int_0^{t/\varepsilon}(\mathcal{L}u^2-2u\mathcal{L}u)(x(s))\ ds =&\lim_{\varepsilon\to 0}\mathbb{E}\ \varepsilon\int_0^{t/\varepsilon}(\mathcal{L}u^2-2u\mathcal{L}u)(x(s))\ ds\\ =&-2\lim_{\varepsilon\to 0}\mathbb{E}\ \varepsilon\int_0^{t/\varepsilon}(u\mathcal{L}u)(x(s))\ ds = 2t\mathcal{D}(u).\end{aligned}$$

(Step 2) We wish to show that if $\|V\|_{H^1}<\infty$, then there exists u such that $-\mathcal{L}u=V$, and $\mathcal{D}(u)<\infty$. Note that for $\lambda>0$, we can always invert the uniformly positive operator $\lambda-\mathcal{L}$:

$$u_\lambda := (\lambda-\mathcal{L})^{-1}V=\int_0^\infty e^{-\lambda t}T_tV\ dt.$$

On the other hand from $\lambda u_\lambda-\mathcal{L}u_\lambda=V$, we learn

$$\lambda\|u_\lambda\|_{L^2}^2+\|u_\lambda\|_{H^1}^2=\langle u_\lambda,V\rangle\le\|u_\lambda\|_{H^1}\|V\|_{H^{-1}}.$$

In particular

$$\|u_\lambda\|_{H^1}\le\|V\|_{H^{-1}},\qquad \lambda\|u_\lambda\|_{L^2}^2\le\|V\|_{H^{-1}}.$$

Hence $\lambda u_\lambda\to 0$ strongly in L^2, and, along a subsequence $u_\lambda\to u$ weakly in H^1. As a consequence,

$$\|u\|_{H^1}^2\le\liminf_{\lambda\to 0}\|u_\lambda\|_{H^1}^2\le\liminf_{\lambda\to 0}\langle u_\lambda,V\rangle=\langle u,V\rangle=\|u\|_{H^1}^2.$$

From this we learn that $u_\lambda\to u$ strongly in H^1 along a subsequence. □

We assume that the process $\mathbf{x}$ is stationary i.e., $x(t)$ is distributed according to π.

Theorem E.2 *Assume that* $F:E\times[0,T]\to\mathbb{R}$ *is a continuous function such that* $F(\cdot,t)\in\mathcal{D}$ *for every* $t\in[0,T]$. *Then*

$$\mathbb{E}^\pi\sup_{t\in[0,T]}\left(\int_0^t\mathcal{S}F(s,x(s))\ ds\right)^2\preceq\mathbb{E}^\pi\int_0^T\|F(s,\cdot)\|_1^2\ ds. \tag{E.9}$$

Proof We note that the reversed process $(x(T-s):\ s\in[0,T])$ is also a stationary Markov process with the generator $\mathcal{L}^*=-\mathcal{A}+\mathcal{S}$. For a sufficiently nice function $F(s,x)$, the processes

$$\begin{aligned}M(t)=&F(x(t),t)-F(x(0),0)-\int_0^t(\partial_s+\mathcal{L})F(s,x(s))\ ds,\\ N(t)=&M(t)^2-\int_0^t(\mathcal{L}F^2-2F\mathcal{L}F)(s,x(s))\ ds\end{aligned}$$

are Martingale with respect to the (forward) filtration $(\mathcal{F}_s : \ s \in [0,T])$, where $\mathcal{F}_s$ is the σ-algebra generated by $(x(\theta) : \ \theta \in [0,s])$. Similarly the processes

$$\widehat{M}(t) = F(x(T-t), T-t) - F(x(T), T) - \int_0^t (-\partial_s + \mathcal{L}^*) F(T-s, x(T-s)) \, ds,$$
$$\widehat{N}(t) = \widehat{M}(t)^2 - \int_0^t (\mathcal{L}^* F^2 - 2F\mathcal{L}^* F)(s, x(s)) \, ds$$

are Martingales with respect to the (backward) filtration $(\widehat{\mathcal{F}}_s : \ s \in [0,T])$, where $\mathcal{F}_s$ is the σ-algebra generated by $(x(T-\theta) : \ \theta \in [0,s])$. Evidently

$$\widehat{M}(T-t) - \widehat{M}(T) = F(x(t), t) - F(x(0), 0) + \int_0^t (-\partial_s + \mathcal{L}^*) F(s, x(s)) \, ds.$$

As a result,

$$\widehat{M}(T-t) - \widehat{M}(T) - M(t) = 2 \int_0^t \mathcal{S} F(s, x(s)) \, ds,$$

which leads to the bound

$$\mathbb{E}^\pi \sup_{t \in [0,T]} \left(\int_0^t \mathcal{S} F(s, x(s)) \, ds \right)^2 \le 3 \mathbb{E}^\pi \sup_{t \in [0,T]} \left(2\widehat{M}(t)^2 + M(t)^2 \right).$$

From this and Doob's inequality we deduce (E.9). □

Theorem E.3 *Assume that* $V : E \times [0,T] \to \mathbb{R}$ *is a continuous function such that* $V(\cdot, t) \in L^2(\pi) \cap \mathcal{H}_1$ *for every* $t \in [0,T]$. *Then*

$$\mathbb{E}^\pi \sup_{t \in [0,T]} \left(\int_0^t V(s, x(s)) \, ds \right)^2 \preceq \mathbb{E}^\pi \int_0^T \|V(s, \cdot)\|_{-1}^2 \, ds. \tag{E.10}$$

Proof

F Martingales

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\widehat{\mathcal{F}} = (\mathcal{F}_n : n \in \mathbb{N})$ be a filtration i.e., each $\mathcal{F}_n$ is a sub σ-algebra of $\mathcal{F}$, and $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for every $n \in \mathbb{N}$. A collection$X_n : \Omega \to \mathbb{R}$ is called a sequence of martingale differences with respect to $\widehat{\mathcal{F}}$ if $(M_n := X_1 + \cdots + X_n : n \in \mathbb{N})$ is a $\widehat{\mathcal{F}}$ martingale (with $M_0 = 0$). In other words, $X_n \in \mathcal{F}_n$, and $\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) = 0$ for and $n \in \mathbb{N}$. Here $\mathcal{F}_0 = \{, \Omega\}$ is the trivial σ-algebra. It is not hard to show

$$\mathbb{E}(X_1 + \cdots + X_n)^2 = \mathbb{E}(X_1^2 + \cdots + EX_n^2).$$

because $\mathbb{E}X_iX_j = 0$ when $i \neq j$. We can prove more:

Theorem F.1 *There exists a constant C_p such that for every $p \geq 1$ and every martingale difference $(X_n : n \in \mathbb{N})$,*

$$\text{(F.1)}\quad \mathbb{E}(X_1 + \cdots + X_n)^p \leq C_p \mathbb{E}\big(X_1^2 + \cdots + X_n^2\big)^{p/2} \leq C_p \left(\big(\mathbb{E}X_1^p\big)^{2/p} + \cdots + \big(\mathbb{E}X_1^p\big)^{2/p}\right)^{p/2}.$$

G Parabolic Regularity

Definition For $\alpha \in \mathbb{R}$ and $p \geq 1$, the Banach space $\big(H^{\alpha,p}, \|\cdot\|_{\alpha,p}$ is defined by

$$H^{\alpha,p} := (I-\Delta)^{-\alpha/2}(L^p), \qquad \|u\|_{\alpha,p} = \big\|(I-\Delta)^{-\alpha/2}u\big\|_p,$$

where $\|\cdot\|_p$ is the usual L^p-norm. The space $H^{\alpha,p}_{loc}$ is the set of all the distribution u such that $u1\!\!1_U \in H^{\alpha,p}$, for every bounded open set U. For $\alpha \in (0,2)$ and $p > 1$,

$$\|u\|_{\alpha,p} \sim \|u\|_p + \big\|\Delta^{\alpha/2}u\|_p,$$

Where,

$$\Delta^{\alpha/2}u(x) = P.V. \int_{\mathbb{R}^d} \frac{f(x+y)-f(x)}{|y|^{d+\alpha}}\ dy.$$

where P.V. stands for Cauchy's principle value. For $u : \mathbb{R}^d \times [0,T]$, the corresponding $H^{\alpha,p}$ is defined as

$$\|u\|^p_{\alpha,p} = \int_0^T \|u(\cdot,t)\|^p_{\alpha,p}\ dt.$$

Theorem G.1 *Let $\lambda > 0, \alpha \in \mathbb{R}, p > 1$, and $T > 0$. For any $u^0 \in H^{2+\alpha,p}$ and $f \in H^{\alpha,p}$, there is a unique solution $u \in H^{2+\alpha,p}$, to the initial-value problem*

$$u_t = \Delta u - \lambda u + f, \qquad u(x,0) = u^0(x). \tag{G.1}$$

Moreover, for any $\theta \in [0,2]$, there is a constant $C = C(\theta,p,d,T)$ such that for all $\lambda > 1$,

$$\|u\|_{\alpha,p} \leq C\big(\lambda^{-\frac{1}{p}}\|u^0\|_{\theta+\alpha,p} + \lambda^{\frac{\theta}{2}-1}\|f\|_{\alpha,p}\big).$$

Proof Let $P_t = e^{t\Delta}$. We wish to obtain a bound on $\|u\|_{\alpha+2,p}$, which holds uniformly in $\lambda \geq 0$. For such a bound, observe that by Duhamel's formula,

$$u(x,t) = e^{-\lambda t}(P_t u^0)(x) + \int_0^t e^{\lambda s}P_s f(t-s,x)\ ds. \tag{G.2}$$

Observe that using

$$\Delta \int_0^t (P_s h)\ ds = P_t h - h,$$

we deduce

$$\left\|\int_0^t (P_s h)\ ds\right\|_{\alpha+2,p} \preceq \|h\|_{\alpha,p}.$$

Hence, for every $\lambda \geq 0$,

$$\|u\|_{\alpha+2,p} \preceq \|u^0\|_{\alpha+2,p} + \|f\|_{\alpha,p}.$$

Moreover, for $\theta \in (0, 2)$, we have

$$\|P_s h\|_{\theta,p} \preceq s^{-\theta/2}\|h\|_p. \tag{G.3}$$

Using this, and (G.2),

$$\begin{aligned}
\|u\|_{\alpha,p}^p \leq & \int_0^T e^{-\lambda tp}\|P_t u^0\|_{\alpha,p}\, dt + \int_0^T \left(\int_0^t e^{\lambda s}\|P_s f(t-s,\cdot)\|_{\theta+\alpha,p}\, ds\right)^p dt \\
\leq & \|P_t u^0\|_{\alpha,p} \int_0^T e^{-\lambda tp}\, dt + \int_0^T \left(\int_0^T e^{\lambda s} s^{-\theta/2}\|f(t-s,\cdot)\|_{\alpha,p}\, ds\right)^p dt \\
\preceq & \lambda^{-1}\|u^0\|_{\theta+\alpha,p}^p + \lambda^{\frac{p\theta}{2}-p}\|f\|_{\alpha,p}\big).
\end{aligned}$$

For the last inequality, we used Holder's inequality, and

$$\int_0^T e^{\lambda s} s^{-\theta/2}\, ds = \lambda^{\theta/2-1}\int_0^{\lambda T} e^s s^{-\theta/2}\, ds \preceq \lambda^{\theta/2-1}.$$

□

References

[AK] S. Albeverio, and S. Kusuoka, The invariant measure and the flow associated to the Φ_3^4-quantum field model. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 20 (2020), 1359–1427.

[BG] N. Barashkov, and M. Gubinelli, A variational method for Φ_3^4. Duke Math. J. **169** (2020), 3339–3415.

[CSZ] F. Caravenna, R. Sun, and N. Zygouras, The two-dimensional KPZ equation in the entire subcritical regime. Ann. Probab. **48** (2020), no. 3, 1086–1127.

[CZ] F. Caravenna, and L. Zambotti, Hairer's reconstruction theorem without regularity structures. EMS Surv. Math. Sci. **7** (2020), 207–251.

[CK] A. Connes and D. Kreimer, Hopf algebras, renormalization and noncommutative geometry. Comm. Math. Phys. **199** (1998) 203–242.

[DG] A. Dunlap, and Y. Gu, A forward-backward SDE from the 2D nonlinear stochastic heat equation. arXiv:2010.03541

[FH] P. K. Friz, and M. Hairer. A course on rough paths. With an introduction to regularity structures. Second edition. Universitext. Springer 2020.

[G1] M. Gubinelli, Controlling rough paths. J. Funct. Anal. **216** (2004) 86–140.

[G2] M. Gubinelli, Ramification of rough paths. J. Differential Equations **248** (2010) 693–721.

[GIP] M. Gubinelli, P. Imkeller, and N. Perkowski, Paracontrolled distributions and singular PDEs. Forum Math. Pi 3 (2015), e6, 75 pp.

[H1] M. Hairer, Solving the KPZ equation. Ann. of Math. **178** (2013), 559–664.

[H2] M. Hairer, A theory of regularity structures. Invent. Math. **198** (2014), 269–504.

[H3] M. Hairer and C. Labbé, The reconstruction theorem in Besov spaces. J. Funct. Anal. bf 273 (2017), 2578–2618.

[HK] M. Hairer and D. Kelly, Geometric versus non-geometric rough paths. Ann. Inst. H. Poincaré Probab. Statist. **51** (2015) 207–251.

[LV] T. Lyons, and N. Victoir, An extension theorem to rough paths. Ann. Inst. H. Poincaré Anal. Non Linéaire **24** (2007), 835–847.

[M] D. Manchon, Hopf algebras, from basics to applications to renormalization. arXiv:math/0408405 9 (2004).