Energy solutions and generators of singular SPDEs

BY LUKAS GRÄFNER, NICOLAS PERKOWSKI

May 12, 2023

Abstract

Energy solutions are a probabilistic theory for singular SPDEs with tractable (quasi-) invariant measures. The prototypical example is the stochastic Burgers/KPZ equation with its white noise invariant measure. Energy solutions were introduced by Gonçalves-Jara [4] and later Gubinelli-Jara [7] and they are based on methods from hydrodynamic limits such as replacement lemmas and martingale estimates. More recently, we understood how to use chaos decompositions to construct and control infinitesimal generators in this setting, which leads to a (weak) well-posedness theory of energy solutions. Compared to pathwise approaches like regularity structures, this requires only relatively soft estimates and the method applies to some scaling (super-)critical equations.

In these lectures, we will start with the guiding example of a diffusion in a singular divergence-free vector field, where we can understand the main ideas of energy solutions without many technicalities and we can already see some (super-)critical problems. Then we will discuss a relatively general and abstract construction of infinitesimal generators, semigroups, and energy solutions based on chaos expansions and infinite-dimensional analysis. Finally we will study applications to singular SPDEs.

Table of contents

1 Introduction

Disclaimer: This is a work in progress and in particular there are many references missing and we should give more credit to other works.

Singular SPDEs are nonlinear stochastic partial differential equation with very irregular noise. For example, if we derive an SPDE as a mesoscopic model for fluctuations in a random system, then the noise in the equation will typically a space-time white noise. (As long as the microscopic system does not have correlations that persist over infinite distances.)

The bulk of these lectures will be quite abstract, and to motivate the following abstract considerations let us first consider some examples that we will be able to treat.

The example which started the theory of singular SPDEs is interface growth. In the pictures below you see different growing interfaces. In a 1986 landmark paper in physics by Kardar, Parisi and Zhang [14] it was conjectured that the fluctuations in such interface growth can, in a certain regime, be modelled by an SPDE which now is called the KPZ equation: $h: \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$,

$$\partial_t h = \Delta h + |\nabla h|^2 + \xi,$$

where ξ is a space-time white noise, i.e. a centered generalized Gaussian process with $\mathbb{E}[\xi(f)\xi(g)] = \langle f, g \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R}^d)}$. This is a singular SPDE, because the noise makes the solution irregular and only in $d = 1$ it is even a function, in higher dimensions it could only be a generalized function (Schwartz distribution)[1.1]. And even in $d = 1$, which corresponds to the pictures below (two-dimensional phases, one-dimensional interface), h is non-smooth and $x \mapsto h(t, x)$ is only as regular as a Brownian motion and therefore $|\nabla h|^2$ makes no sense.

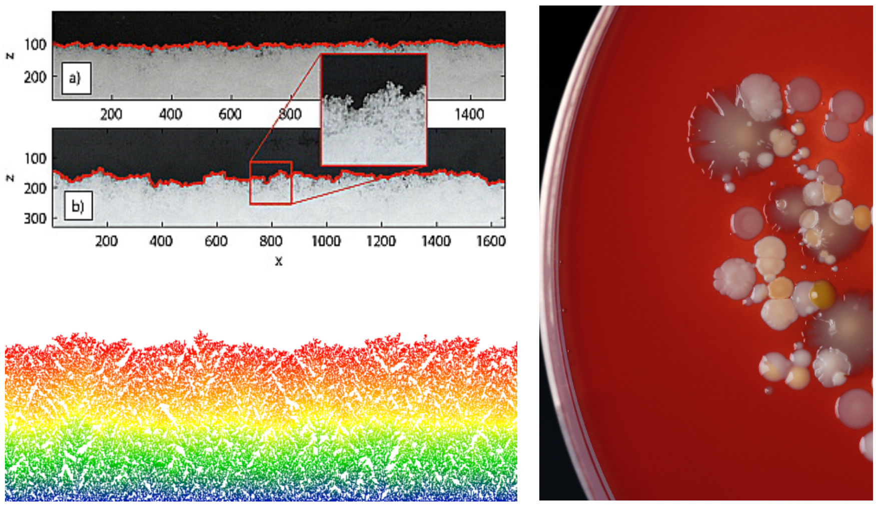

Figure 1.1. Growing interfaces. Image credit: Löwe et al., Geophys. Res. Letters, Vol. 34, L21507, 2007 (upper left), Nils Berglund (lower left), iStockphoto.com/rudigobbo (right)

In that case there is a simple trick to make sense of h: If we define $w = e^h$ ("Cole-Hopf transformation"), then w formally solves the stochastic heat equation

$$\partial_t w = \Delta w + w\xi,$$

which is linear and well-posed as an Itô SPDE. Therefore, we can simply define $h := \log w$ (luckily w is strictly positive for positive initial conditions) and this gives us the right object to work with. But in this way we do not get an equation for h.

The first widely visible[1.2] breakthrough in singular SPDEs was a work by Hairer [13] in which he solved the KPZ equation using rough path integrals (a pathwise version of stochastic integration). The key point is that the roughness in $h(t, \cdot)$ is in the space variable, and therefore there is no direction of information and Itô techniques are not useful. While the pathwise approach does not care about that. This inspired a lot of follow-up research, for example Hairer's regularity structures [12] and paracontrolled distributions by Gubinelli-Imkeller-Perkowski [6] extend rough path integration to higher dimensions, which is necessary to treat equations with higher-dimensional space variables.

1.1. In fact it is expected/in some cases shown that in $d \geqslant 3$ there is no nontrivial solution to the SPDE, and h is Gaussian and a solution of $\partial_t h = \nu \Delta h + \sigma \xi$ for some effective parameters $\nu, \sigma > 0$. The physically most relevant case $d = 2$ (three-dimensional phases) is more subtle and finer details of the equation should determine whether solutions are Gaussian or not.

1.2. There was a previous work by Hairer [11] where he first demonstrated the usefulness of rough path techniques for singular SPDEs and which layed the foundation for the KPZ paper. This is a beautiful and groundbreaking work, and the main reason why it did not get the same attention as the KPZ paper is that the SPDE treated here is not as famous as the KPZ equation.

By now singular SPDEs are a flourishing area of research and hundreds of papers developed and consolidated the field since those early days. We do not go into detail and simply stress that most works in the area follow the original, pathwise philosophy: To solve a singular SPDE, we freeze a realization of the noise $\xi(\omega)$ (together with a finite number of "trees", i.e. nonlinear functionals, built from $\xi(\omega)$), and then proceed to solve the SPDE with deterministic arguments. At the moment this is the only general approach we have for solving many singular SPDEs in a unified framework. But for some equations there is an alternative approach, based on probabilistic techniques such as martingales. Here we present this alternative approach. Let us give some examples of SPDEs where this is applicable:

Example 1.1.

i. Stochastic Burgers equation: The derivative of the KPZ equation solves the stochastic Burgers equation $u: \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$

$$\partial_t u = \partial_{xx} u + \partial_x u^2 + \partial_x \xi.$$

We can derive this model as a mesoscopic fluctuation scaling limit from microscopic models for local differences in interface growth. The simplest model is the (weakly asymmetric) simple exclusion process on $\mathbb{Z}$, i.e. a system of particles which perform continuous time independent random walks with rate p (resp. $1-p$) of jumping to the right (resp. the left), but which are not allowed to jump on top of each other; each site has at most one particle.

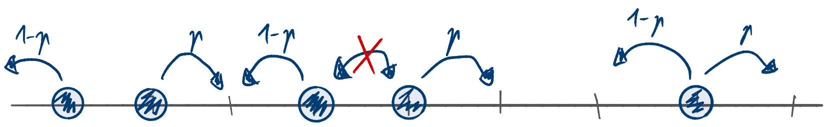

Figure 1.2. Simple exclusion process

One motivation for studying this particle system is that it corresponds to a simple interface model: We can imagine a piecewise linear curve $(h(t,k): t \geqslant 0, k \in \mathbb{Z})$ over $\mathbb{Z}$, such that $h(t,k+1) - h(t,k) = 1$ if there is a particle at site k, and $h(t,k+1) - h(t,k) = -1$ otherwise. Then local maxima become local minima with rate p, and local minima become local maxima with rate $1-p$.

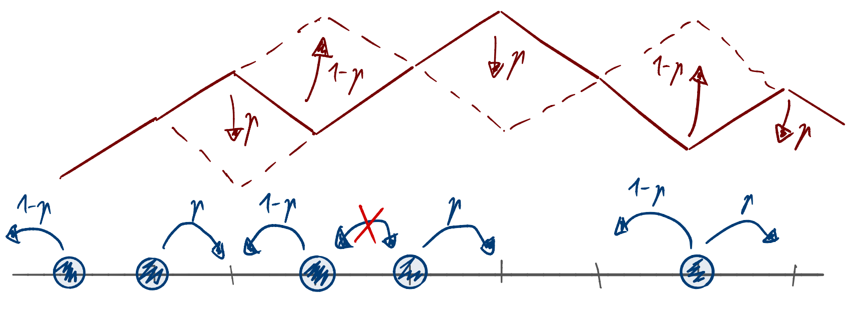

Figure 1.3. Simple exlusion as interface model

The oscillations of this random interface could for example be a toy model for the interfaces in Figure 1.1. In that case the up and down motion is not symmetric, so we would expect $p < \frac{1}{2}$ ($p = 0$ would correspond to growth only, and $p > 0$ would for example allow some melting of the snow). The large scale behavior for fixed $p \in \left(0, \frac{1}{2}\right)$ is described by the KPZ fixed point, a complicated stochastic process which can only be described by explicit formulas for its transition probabilities, but for which we do not know any differential equation [16, 18].

But if the random walk in the exclusion process is symmetric, i.e. $p=\frac{1}{2}$, then on large scales the particle system converges to the SPDE

$$\partial_t u = \frac{1}{2}\partial_{xx} u + \partial_x \xi,$$

on $\mathbb{R}_+ \times \mathbb{R}$, where ξ is a space-time white noise (roughly speaking $\xi(t,x)$ is independent of $\xi(s,y)$ whenever $(t,x)\neq(s,y)$). This equation is called the *infinite-dimensional Ornstein-Uhlenbeck process*. If we take a small perturbation around the symmetric jump rates and consider $p=\frac{1}{2}+\lambda\varepsilon$, with $\lambda\in\mathbb{R}$ and $\varepsilon\to 0$ as we scale out, then the scaling limit is given by the *stochastic Burgers equation*

$$\partial_t u = \frac{1}{2}\partial_{xx} u + \lambda\partial_x u^2 + \partial_x \xi,$$

which is singular because u is only a generalized function and therefore u^2 is not classically defined. In that case the scaling limit for the interface is the *KPZ equation*

$$\partial_t h = \frac{1}{2}\partial_{xx} h - (\partial_x h)^2 + \xi.$$

Using the Cole-Hopf transform, this convergence was established well before the start of singular SPDEs [2].

ii. Fractional, multi-component Burgers equation: If particles in the exclusion process are allowed to do long range jumps in such a way that the rescaled random walk of a single non-interacting particle would converge to an α-stable Lévy process, then the scaling limit of the particle system is a nonlocal stochastic Burgers equation $u:\mathbb{R}_+\times\mathbb{R}\to\mathbb{R}$ (again interpreted as a distribution in the space variable)

$$\partial_t u = -(-\Delta)^{\theta} u + \partial_x u^2 + \sqrt{2}(-\Delta)^{\theta/2}\xi,$$

where again ξ is a space-time white noise and $\theta\in\left[\frac{3}{4},1\right]$, see [5]. For $\theta>\frac{3}{4}$ we need to consider a weakly asymmetric regime, while for $\theta=\frac{3}{4}$ this limit arises from a fixed strength of asymmetry. For $\theta<1$ this equation does not have a Cole-Hopf transform. For $\theta=\frac{3}{4}$ it is scaling invariant and therefore out of the range of pathwise theories, which are crucially based on the fact that nonlinearities vanish on small scales and can be controlled by the linear terms – this is called *subcriticality*.

If there are more than one particle type, say red, blue and green particles which interact differently, then we could expect a multi-component fractional stochastic Burgers equation

$$\partial_t u = -(-\Delta)^{\theta} u + \partial_x (u\cdot\Gamma u) + \sqrt{2}(-\Delta)^{\theta/2}\xi,$$

where $\Gamma\in\mathbb{R}^{d\times d\times d}$ is a tensor coupling the different components, and ξ is a vector-valued space-time white noise.

iii. Stochastic surface quasi-geostrophic equation, regularization by noise: The surface quasi-geostrophic equation is a popular model in fluid dynamics, describing for example the evolution of the temperature in a fluid. It is given by

$$\begin{aligned}\partial_t\theta + u\cdot\nabla\theta &= 0,\\ u &= \nabla^{\perp}(-\Delta)^{-1/2}\theta\end{aligned}$$

on $\mathbb{R}_+\times\mathbb{T}^2$ or $\mathbb{R}_+\times\mathbb{R}^2$ and where $\nabla^{\perp}=(\partial_2,-\partial_1)$. To the best of our knowledge, the well-posedness of this equation remains a challenging open problem and one of the best results is the well-posedness of the critical model with fractional viscosity [15]:

$$\partial_t\theta + u\cdot\nabla\theta = (-\Delta)^{1/2}\theta.$$

This equation is formally scaling invariant, but of course the viscosity has a regularizing effect and it adds energy dissipation, while the original equation formally conserves the energy $\int \theta^2$. In particular, while the inviscous equation formally has an invariant measure given by the law of the Gaussian white noise, the viscous equation does not preserve this measure. We could regularize the equation differently, by adding an injection of energy on top of the dissipation, so that formally the energy is preserved:

$$\begin{aligned} \partial_t \theta + u \cdot \nabla \theta &= (-\Delta)\theta + \sqrt{2}(-\Delta)^{1/2}\xi, \\ u &= \nabla^{\perp}(-\Delta)^{-1/2}\theta, \end{aligned}$$

for a space-time white noise ξ. Now the equation is scaling invariant, and also the energy is preserved. Strictly speaking the energy $\int \theta^2$ is infinite at each time, but the "energy measure" formally given by $e^{-\int \theta^2}\mathrm{d}\theta$ is preserved, which is the white noise measure.

iv. To some extent Landau-Lifshitz Navier-Stokes. To do.

We will see that, at least i.-iii., can be interpreted as infinite-dimensional stochastic differential equations with a drift given by an infinite-dimensional drift. To get an idea how to solve those, we first treat a conceptually simpler, finite-dimensional example.

Example 1.2. (Diffusion in the curl of the GFF) Let ξ be a periodic Gaussian free field on $\mathbb{R}^2$, i.e. the centered Gaussian process with covariance $\mathbb{E}[\xi(f)\xi(g)] = \langle (-\Delta)^{-1/2} f, (-\Delta)^{-1/2} g\rangle$. The recent works To Do! consider the diffusion on $\mathbb{R}^2$

$$\mathrm{d}X_t = \nabla^{\perp}\rho * \xi(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t,$$

where $\rho \in C_c^{\infty}$ is a mollifier. They show that on large scales, X behaves super-diffusively, roughly speaking

$$\mathbb{E}[|X_t|^2] \simeq t\sqrt{\log t}, \qquad t \to \infty.$$

Translating this to a small scale problem, this shows that taking the truncation away, there exists no limit: the sequence of processes

$$\mathrm{d}X_t^{\varepsilon} = \nabla^{\perp}\rho_{\varepsilon} * \xi(X_t^{\varepsilon})\mathrm{d}t + \sqrt{2}\mathrm{d}W_t,$$

where $\rho_{\varepsilon}(x) = \varepsilon^{-2}\rho(\varepsilon^{-1}\cdot)$, is not tight and does not converge in distribution. While on large scales the issue are the long range correlations in the GFF, on small scales the issue is its irregularity: We have (locally) $\xi \in \mathcal{C}^{-1-\kappa} = B_{\infty,\infty}^{-1-\kappa}$ for all $\kappa > 0$ but not for $\kappa = 0$. The law of the diffusion is formally equivalent to the Kolmogorov backward equation

$$\partial_t u_{\varepsilon} = \Delta u_{\varepsilon} + b_{\varepsilon} \cdot \nabla u_{\varepsilon}, \qquad b_{\varepsilon} = \nabla^{\perp}\rho_{\varepsilon} * \xi.$$

We know from regularity structures that this equation is scaling subcritical exactly if (locally) $b_{\varepsilon} \in \mathcal{C}^{-1+\kappa}$ (i.e. then u_{ε} is a perturbation of the heat equation on small scales). Our example shows that in this case the subcriticality condition is a sharp obstacle to well-posedness and if it is only slightly violated it may be impossible to construct a limit of $(X^{\varepsilon})_{\varepsilon > 0}$.

But what if we regularize this problem slightly, say with a Fourier multiplier $(1 + |x|^2)^{-\kappa}$ for some $\kappa > 0$? In that case $b_{\varepsilon} \in \mathcal{C}^{-1+2\kappa-} \subset \mathcal{C}^{-1+\kappa}$ locally, so the dynamics are subcritical and we should be able to solve the equation using regularity structures. But for $\kappa \to 0$ the number of trees we need to construct converges to ∞.

2 Lecture 1: Diffusion with distributional drift

We want to solve the d-dimensional SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}B_t, \qquad X_0 \sim \mu,$$

where b is a distribution with $\operatorname{div} b = 0$. We will assume that b is periodic with respect to integer shifts. We will interpret the SDE in the weak sense, and define the drift by a limit:

$$X_t = X_0 + \lim_{n \to \infty} \int_0^t b_n(X_s)\mathrm{d}s + B_t,$$

for smooth approximations b_n of b. We will construct such solutions and show weak uniqueness if X additionally satisfies some "energy estimates/admissibility condition" and an "incompressibility condition". The uniqueness argument will rely on duality with the Kolmogorov backward equation, for which we will construct sufficiently good solutions.

2.1 Construction of energy solutions to the SDE

To construct energy solutions, we go through the usual construction of solutions to martingale problems: We replace b by a smooth approximation and prove tightness. Let us start with an auxiliary observation. As usual we identify functions on $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ with periodic functions on $\mathbb{R}^d$.

Lemma 2.1. *Let X solve*

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

where B is a Brownian motion and where $b \in C^\infty(\mathbb{T}^d)$ is divergence free. Let $Y_t = X_t \bmod \mathbb{Z}^d$. Then Y is an ergodic $\mathbb{T}^d$-valued Markov process with generator $\mathcal{L} = \frac{1}{2}\Delta + b\cdot\nabla$ and with unique invariant measure the Lebesgue measure on $\mathbb{T}^d$. If Y_0 is stationary, then for any $T > 0$ the time-reversed process $\hat{Y}_t$ has the generator $\mathcal{L}^* = \frac{1}{2}\Delta - b\cdot\nabla$.

Proof. Y is Markov because b is periodic and therefore predicting the future behavior of $(X_{t+s})_{s\geqslant 0} \bmod \mathbb{Z}^d$ requires only knowledge of $X_t \bmod \mathbb{Z}^d$. To see that Y has the generator $\mathcal{L} = \frac{1}{2}\Delta + b\cdot\nabla$ we can apply Itô's formula to X. Next, we have by integration by parts and because b is divergence free:

$$\begin{aligned}
\int_{\mathbb{T}^d} \mathcal{L}f(x)g(x)\mathrm{d}x &= \int_{\mathbb{T}^d} \left(\frac{1}{2}\Delta + b\cdot\nabla\right) f(x)g(x)\mathrm{d}x \\
&= \int_{\mathbb{T}^d} \left(\frac{1}{2}\Delta f(x) + \nabla\cdot(bf)(x)\right) g(x)\mathrm{d}x \\
&= \int_{\mathbb{T}^d} f(x)\left(\frac{1}{2}\Delta g(x) - \nabla\cdot(bg)(x)\right)\mathrm{d}x \\
&= \int_{\mathbb{T}^d} f(x)\mathcal{L}^* g(x)\mathrm{d}x,
\end{aligned}$$

where $\mathcal{L}^* = \frac{1}{2}\Delta - b\cdot\nabla$. With $g \equiv 1$ we see that the Lebesgue measure is indeed invariant. It then follows from general results on Markov processes[2.1] that for Lebesgue distributed initial condition the time-reversed process has the generator $\mathcal{L}^*$.

Uniqueness of the invariant distribution holds because the diffusion is irreducible since we have additive noise. Here is a shorter PDE-theoretic argument: The Kolmogorov forward equation, which describes the evolution of the probability distribution of Y, is

$$\partial_t \rho = \mathcal{L}^* \rho = \left(\frac{1}{2}\Delta - b\cdot\nabla\right)\rho.$$

From basic regularity theory we get that $\rho(t_0)$ has a smooth density for any $t_0 > 0$. Using the equation on $[t_0, \infty)$ and differentiating the L^2-norm of $\rho - 1$, where by abuse of notation ρ also is the density, we get with the $L^2(\mathbb{T}^d)$ inner product $\langle\cdot,\cdot\rangle$

$$\begin{aligned}
\partial_t \int (\rho - 1)^2 &= 2\left\langle (\rho - 1), \left(\frac{1}{2}\Delta - b\cdot\nabla\right)\rho \right\rangle \\
&= 2\left\langle (\rho - 1), \left(\frac{1}{2}\Delta - b\cdot\nabla\right)(\rho - 1) \right\rangle \\
&= -\int |\nabla(\rho - 1)|^2,
\end{aligned}$$

2.1. See tutorial.

where we used that $b\cdot\nabla$ is antisymmetric wrt. Lebesgue measure (as we have seen above), and therefore $\langle f, b\cdot\nabla f\rangle = 0$. By the Poincaré inequality we have

$$-\int |\nabla(\rho-1)|^2 \leqslant |2\pi|^2 \int |\rho-1|^2,$$

so by Gronwalls inequality

$$\int (\rho(t)-1)^2 \leqslant e^{-|2\pi|^2(t-t_0)} \int (\rho(t_0)-1)^2,$$

which converges to 0 for $t\to\infty$. □

The time-reversal gives a very useful bound for expectations of the form $\mathbb{E}[|\int_0^T f(Y_s)\mathrm{d}s|^p]$. This is based on the so called *Itô trick*:

Lemma 2.2. (Itô trick) *Let Y be a periodic diffusion with generator $\frac{1}{2}\Delta + b\cdot\nabla$, where $b\in C^\infty(\mathbb{T}^d)$ is divergence free. Assume that Y_0 is uniformly distributed on $\mathbb{T}^d$. Then we have for all $T>0$, for all $f\in C^2(\mathbb{T}^d)$ and for all $p\geqslant 2$*

$$\mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t \Delta f(Y_s)\mathrm{d}s\right|^p\right] \lesssim T^{\frac{p}{2}} \|\nabla f\|_{L^p(\mathbb{T}^d)}^p.$$

Proof. Let $\hat{Y}_t = Y_{T-t}$. We get from the martingale problem

$$\begin{aligned}
f(Y_t) &= f(Y_0) + \int_0^t \mathcal{L}f(s, Y_s)\mathrm{d}s + M_t^f,\\
f(\hat{Y}_T) &= f(\hat{Y}_{T-t}) + \int_{T-t}^T \mathcal{L}^* f(\hat{Y}_s)\mathrm{d}s + \hat{M}_T^f - \hat{M}_{T-t}^f,
\end{aligned}$$

where M^f is a martingale and $\hat{M}^f$ is a martingale in the backward filtration (the filtration generated by $\hat{Y}$). We can rewrite the equation involving $\hat{Y}$ as

$$\begin{aligned}
f(Y_0) &= f(Y_t) + \int_{T-t}^T \mathcal{L}^* f(Y_{T-s})\mathrm{d}s + \hat{M}_T^f - \hat{M}_{T-t}^f\\
&= f(Y_t) + \int_0^t \mathcal{L}^* f(Y_s)\mathrm{d}s + \hat{M}_T^f - \hat{M}_{T-t}^f.
\end{aligned}$$

Adding these two equations and using that $\mathcal{L} + \mathcal{L}^* = \Delta$, we get

$$\int_0^t \Delta f(Y_s)\mathrm{d}s = M_t^f + \hat{M}_T^f - \hat{M}_{T-t}^f.$$

The quadratic variations are

$$\langle M^f\rangle_t = \int_0^t |\nabla f(Y_s)|^2\mathrm{d}s, \qquad \langle \hat{M}^f\rangle_t = \int_0^t |\nabla f(\hat{Y}_s)|^2\mathrm{d}s,$$

which follows for example by writing $f(Y_t) = f(X_t)$ and applying Itô's formula to $f(X_t)$, and similarly for $\hat{Y}$. Thus, we have by Burkholder-Davis Gundy, Minkowski's inequality and stationarity

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t \Delta f(Y_s)\mathrm{d}s\right|^p\right] &\lesssim \mathbb{E}\left[\left(\int_0^T |\nabla f(Y_s)|^2\mathrm{d}s\right)^{p/2}\right]\\
&\quad + \mathbb{E}\left[\left(\int_0^T |\nabla f(Y_{T-s})|^2\mathrm{d}s\right)^{p/2}\right]\\
&\lesssim \left(\int_0^T \mathbb{E}[|\nabla f(Y_s)|^p]^{2/p}\mathrm{d}s\right)^{p/2}\\
&= T^{\frac{p}{2}} \|\nabla f\|_{L^p}^p.
\end{aligned}$$

□

Corollary 2.3. *Let $(b_n)_{n\in\mathbb{N}}\subset C^\infty(\mathbb{T}^d)$ be divergence free such that $\sup_n \|b_n\|_{\mathcal{C}^{-1+\kappa}}<\infty$ for some $\kappa>0$, and let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $\mathbb{T}^d$ such that $\sup_{n\in\mathbb{N}}\|\frac{\mathrm{d}\mu_n}{\mathrm{d}\lambda}\|_{L^2(\mathbb{T}^d)}<\infty$ with the Lebesgue measure λ on $\mathbb{T}^d$.*

Let $Y_0^n\sim\mu_n$ and let Y^n be the periodic diffusion with generator $\frac{1}{2}\Delta+b_n\cdot\nabla$. Then $(Y^n)_{n\in\mathbb{N}}$ is tight in $C(\mathbb{R}_+,\mathbb{T}^d)$. Moreover, we have uniformly in n for all $f\in L^2(\mathbb{T}^d)$

$$\mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t f(Y_s^n)\mathrm{d}s\right|\right]\lesssim T^{1/2}\|(-\Delta)^{-1/2}f\|_{L^2(\mathbb{T}^d)}.$$

Proof. For simplicity we assume that $\mathcal{F}b_n(0)=0$ for each n, i.e. that b_n has vanishing zero Fourier mode[2.2]. In that that we can construct $(-\Delta)^{-1}b_n$ by Fourier analysis:

$$\mathcal{F}((-\Delta)^{-1}b_n)(k)=\frac{1}{|2\pi k|^2}\mathcal{F}b_n(k).$$

The Itô trick yields under the assumption $\frac{\mathrm{d}\mu_n}{\mathrm{d}\lambda}\equiv 1$, bounding $\nabla(-\Delta)^{-1/2}\lesssim 1$:

$$\begin{aligned}\mathbb{E}\left[\left|\int_s^t b_n(Y_r^n)\mathrm{d}r\right|^p\right] &\lesssim (t-s)^{\frac{p}{2}}\|\nabla(-\Delta)^{-1}b_n\|_{L^p(\mathbb{T}^d)}^p\\ &\lesssim (t-s)^{\frac{p}{2}}\|(-\Delta)^{-1/2}b_n\|_{L^p(\mathbb{T}^d)}^p\\ &\lesssim (t-s)^{\frac{p}{2}}\|(-\Delta)^{-1/2}b_n\|_{L^\infty(\mathbb{T}^d)}^p\\ &\lesssim (t-s)^{\frac{p}{2}}\|b_n\|_{\mathcal{C}^{-1+\kappa}}^p,\end{aligned}$$

where the last step follows from regularity estimates for Δ in Besov spaces. For non-stationary Y_0^n we use the Cauchy-Schwarz inequality:

$$\begin{aligned}\mathbb{E}_{\mu_n}\left[\left|\int_s^t b_n(Y_r^n)\mathrm{d}r\right|^p\right] &= \int\mathbb{E}_y\left[\left|\int_s^t b_n(Y_r^n)\mathrm{d}r\right|^p\right]\frac{\mathrm{d}\mu_n}{\mathrm{d}\lambda}(y)\mathrm{d}y\\ &\leqslant \left\|\frac{\mathrm{d}\mu_n}{\mathrm{d}\lambda}\right\|_{L^2(\mathbb{T}^d)}^{1/2}\left(\int\mathbb{E}_y\left[\left|\int_s^t b_n(Y_r^n)\mathrm{d}r\right|^p\right]^2\mathrm{d}y\right)^{1/2}\\ &\leqslant \left\|\frac{\mathrm{d}\mu_n}{\mathrm{d}\lambda}\right\|_{L^2(\mathbb{T}^d)}^{1/2}\left(\int\mathbb{E}_y\left[\left|\int_s^t b_n(Y_r^n)\mathrm{d}r\right|^{2p}\right]\mathrm{d}y\right)^{1/2}\\ &= \left\|\frac{\mathrm{d}\mu_n}{\mathrm{d}\lambda}\right\|_{L^2(\mathbb{T}^d)}^{1/2}\mathbb{E}_\lambda\left[\left|\int_s^t b_n(Y_r^n)\mathrm{d}r\right|^{2p}\right]^{1/2}\\ &\lesssim (t-s)^{\frac{p}{2}}\|(-\Delta)^{-1}b_n\|_{\mathcal{C}^{-1+\kappa}}^p.\end{aligned}$$

Since $p\geqslant 2$ is arbitrary, tightness follows from Kolmogorov's continuity criterion. Also, the arguments did not depend on the fact that we plugged in b_n and therefore the claim for general f holds by the same steps. □

Theorem 2.4. (Existence of energy solutions) *Let (b_n), (μ_n), (X^n) as in the previous lemma. Assume that b_n converges in $\mathcal{C}^{-1+\kappa}$ to some b, that μ_n converges weakly to some μ, and let Y be a weak limit of the (Y^n). Then:*

i. Y solves the martingale problem with generator $\frac{1}{2}\Delta+b\cdot\nabla$ in a limiting sense: For all $f\in C^2(\mathbb{T}^d)$ the process

$$f(Y_t)-f(Y_0)-\lim_{n\to\infty}\int_0^t\left(\frac{1}{2}\Delta+b_n\cdot\nabla\right)f(Y_s)\mathrm{d}s$$

is a martingale.

ii. Y is incompressible: *For all bounded and measurable $f:\mathbb{T}^d\to\mathbb{R}$ we have*

$$\mathbb{E}[|f(Y_t)|]\lesssim\|f\|_{L^2(\mathbb{T}^d)}.$$

2.2. Otherwise we would solve $(-\Delta)^{-1}(b_n-\mathcal{F}b_n(0))$

iii. Y is admissible / satisfies an energy estimate*: For all $f \in L^2(\mathbb{T}^d)$*

$$\mathbb{E}\left[\sup_{t \leqslant T}\left|\int_0^t f(Y_s)\mathrm{d}s\right|\right] \lesssim T^{1/2}\|(-\Delta)^{-1/2} f\|_{L^2(\mathbb{T}^d)}.$$

We call such Y an energy solution *of the SDE $\mathrm{d}Y_t = b(Y_t)\mathrm{d}t + \mathrm{d}B_t$ (interpreted periodically). In that case we can extend the map $I: L^2(\mathbb{T}^d) \to L^1(\Omega, C([0,T]))$,*

$$I(f)_t = \int_0^t f(Y_s)\mathrm{d}s,$$

continuously to H^{-1}, and we denote the extension with the same symbol I.

Proof.

i. Let $f \in C^2$ and $0 \leqslant s < t$ and let $G: C([0,s], \mathbb{T}^d)$ be continuous and bounded. We have to show that $\lim_{n\to\infty} \int_0^t \left(\frac{1}{2}\Delta + b_n \cdot \nabla\right) f(Y_s)\mathrm{d}s$ exists and that

$$\mathbb{E}\left[\left(f(Y_t) - f(Y_s) - \lim_{n\to\infty}\int_s^t \left(\frac{1}{2}\Delta + b_n\cdot\nabla\right) f(Y_r)\mathrm{d}r\right) G((Y_r)_{r\in[0,s]})\right] = 0.$$

For simplicity we assume that $b_n \cdot \nabla f$ has no zero Fourier mode[2.3], for all n. The Itô trick yields

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t \left(\frac{1}{2}\Delta + b_n\cdot\nabla\right) f(Y_s^n)\mathrm{d}s - \int_0^t \left(\frac{1}{2}\Delta + b_m\cdot\nabla\right) f(Y_s^n)\mathrm{d}s\right|\right]\\
&\lesssim T^{1/2}\|(-\Delta)^{-1/2}((b_n - b_m)\cdot\nabla f)\|_{L^2(\mathbb{T}^d)}\\
&\lesssim T^{1/2}\|b_n - b_m\|_{\mathcal{C}^{-1+\kappa}}\|f\|_{C^2}.\\
&\lesssim T^{1/2}\|b - b_m\|_{\mathcal{C}^{-1+\kappa}}\|f\|_{C^2}.
\end{aligned}$$

Since b_m is smooth and bounded, the map $x \mapsto \int_0^t b_m \cdot \nabla f(x(s))\mathrm{d}s$ is a continuous bounded map in the uniform topology and we obtain

$$\int_0^t \left(\frac{1}{2}\Delta + b_m\cdot\nabla\right) f(Y_s^n)\mathrm{d}s \overset{n\to\infty}{\longrightarrow} \int_0^t \left(\frac{1}{2}\Delta + b_m\cdot\nabla\right) f(Y_s)\mathrm{d}s,$$

weakly. Therefore, Fatou's lemma for weak convergence yields

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t \left(\frac{1}{2}\Delta + b_m\cdot\nabla\right) f(Y_s)\mathrm{d}s - \int_0^t \left(\frac{1}{2}\Delta + b_{m'}\cdot\nabla\right) f(Y_s)\mathrm{d}s\right|\right]\\
&\leqslant \liminf_{n\to\infty}\mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t \left(\frac{1}{2}\Delta + b_m\cdot\nabla\right) f(Y_s^n)\mathrm{d}s - \int_0^t \left(\frac{1}{2}\Delta + b_{m'}\cdot\nabla\right) f(Y_s^n)\mathrm{d}s\right|\right]\\
&\lesssim \|b_m - b_{m'}\|_{\mathcal{C}^{-1+\kappa}}
\end{aligned}$$

and thus $\int_0^t \left(\frac{1}{2}\Delta + b_m\cdot\nabla\right) f(Y_s)\mathrm{d}s$ is Cauchy in L^1 and the L^1-limit exists. Moreover, by L^1-convergence

$$\begin{aligned}
&\mathbb{E}\left[\left(f(Y_t) - f(Y_s) - \lim_{m\to\infty}\int_s^t \left(\frac{1}{2}\Delta + b_m\cdot\nabla\right) f(Y_r)\mathrm{d}r\right) G((Y_r)_{r\in[0,s]})\right]\\
&= \lim_{m\to\infty}\mathbb{E}\left[\left(f(Y_t) - f(Y_s) - \int_s^t \left(\frac{1}{2}\Delta + b_m\cdot\nabla\right) f(Y_r)\mathrm{d}r\right) G((Y_r)_{r\in[0,s]})\right]\\
&= \lim_{m\to\infty}\lim_{n\to\infty}\mathbb{E}\left[\left(f(Y_t^n) - f(Y_s^n) - \int_s^t \left(\frac{1}{2}\Delta + b_m\cdot\nabla\right) f(Y_r^n)\mathrm{d}r\right) G((Y_r^n)_{r\in[0,s]})\right]\\
&= \lim_{m\to\infty}\lim_{n\to\infty}\mathbb{E}\left[\left(\int_s^t (b_n - b_m)\cdot\nabla f(Y_r^n)\mathrm{d}r\right) G((Y_r^n)_{r\in[0,s]})\right] + 0,
\end{aligned}$$

2.3. As discussed before, the contribution from the zero Fourier mode can be easily bounded by "brute force" and we use the Itô trick to handle the rest.

where we used that Y^n solves the martingale problem for $\frac{1}{2}\Delta + b_n \cdot \nabla$. Now the claim follows from another application of the Itô trick.

ii. and iii. follow from Fatou's lemma. □

Note that we only constructed energy solutions for near-stationary initial conditions with density in L^2. This can be relaxed to densities with finite entropy, but for deterministic initial conditions we would need an additional argument and in that case we would have to find new versions of the admissibility and incompressibility conditions.

Our next goal is to prove uniqueness of energy solutions. This will be based on duality with the Kolmogorov backward equation.

2.2 Solving the Kolmogorov backward equation

To prove uniqueness for the energy solution X, we want to solve the Kolmogorov backward equation

$$\partial_t u = \frac{1}{2}\Delta u + b \cdot \nabla u, \qquad u(0) = \varphi,$$

for a sufficiently rich class of initial conditions φ. For simplicity we only consider periodic φ, which means we will only show uniqueness in law of $Y = X \bmod \mathbb{Z}^d$. We will explain later how we could adapt the arguments to get uniqueness in law of X.

For $b \in \mathcal{C}^{-1+\kappa}(\mathbb{T}^d)$ our first goal is to make sense of the operator $\mathcal{L} = \frac{1}{2}\Delta + b \cdot \nabla$ on a suitable subset of $L^2(\mathbb{T}^d)$. We will work with L^2-Sobolev spaces

$$H^\alpha = H^\alpha(\mathbb{T}^d) = B^\alpha_{2,2}(\mathbb{T}^d) = W^{\alpha,2}(\mathbb{T}^d).$$

To construct $\mathcal{L}$, it is useful to work with paraproducts. We will not go into the details of paraproducts here, but just use the following facts:

Lemma 2.5. *The product of smooth functions $(f, g) \mapsto fg$ can be decomposed as a sum of three operators,*

$$fg = f \otimes g + f \odot g + g \otimes f,$$

which are continuous bilinear maps satisfying the following bounds[2.4] *for $\alpha \in \mathbb{R}$ and $\beta < 0$*

$$\|f \otimes g\|_{H^\alpha} \lesssim \|f\|_{L^2} \|g\|_{\mathcal{C}^\alpha}, \qquad \|f \otimes g\|_{H^{\beta+\alpha}} \lesssim \|f\|_{\mathcal{C}^\beta} \|g\|_{H^\alpha},$$

and, if $\alpha + \beta > 0$,

$$\|f \odot g\|_{H^{\alpha+\beta}} \lesssim \|f\|_{H^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

We call $\otimes$ the paraproduct and $\odot$ the resonant product.

You can find the construction of the paraproduct and the proof of these bounds in Chapter 2 of [1]. Equipped with this, we define now $\mathcal{L}: H^1 \to H^{-1+\kappa}$ via

$$\mathcal{L} f := \mathcal{L}_0 f + \mathcal{G} f := \underbrace{\frac{1}{2}\Delta f}_{\mathcal{L}_0 f} + \underbrace{b \otimes \nabla f + b \otimes \nabla f + \nabla \cdot (b \odot f)}_{\mathcal{G} f}, \qquad f \in H^1(\mathbb{T}^d).$$

Lemma 2.6. *We have*

$$\begin{aligned} \|\mathcal{L}_0 f\|_{H^{-1}} &\lesssim \|f\|_{H^1}, \\ \|\mathcal{G} f\|_{H^{-1+\kappa}} &\lesssim \|b\|_{\mathcal{C}^{-1+\kappa}} \|f\|_{H^1}. \end{aligned}$$

2.4. Actually the first bound only holds if we replace H^α on the left hand side by $B^\alpha_{2,\infty}$ or by $H^{\alpha-\varepsilon}$ for some arbitrarily small $\varepsilon > 0$. But for $\kappa > 0$ we anyways have some "spare regularity", so to simplify the presentation we ignore this small loss of regularity.

Proof. The bound for $\mathcal{L}_0$ is clear. We use the bounds for paraproduct and resonant product as follows:

$$\begin{aligned}\|b \varolessthan \nabla f\|_{H^{-1+\kappa}} &\lesssim \|b\|_{\mathcal{C}^{-1+\kappa}}\|\nabla f\|_{L^2} \lesssim \|b\|_{\mathcal{C}^{-1+\kappa}}\|f\|_{H^1},\\ \|\nabla f \varolessthan b\|_{H^{-1+\kappa}} &\lesssim \|\nabla f\|_{L^2}\|b\|_{\mathcal{C}^{-1+\kappa}} \lesssim \|f\|_{H^1}\|b\|_{\mathcal{C}^{-1+\kappa}},\\ \|\nabla\cdot(b \odot f)\|_{H^{-1+\kappa}} &\lesssim \|b \odot f\|_{H^\kappa} \lesssim \|b\|_{\mathcal{C}^{-1+\kappa}}\|f\|_{H^1}.\end{aligned}$$

□

Theorem 2.7. (Construction of the semigroup / solution of the Kolmogorov backward equation) [2.5]*Let $b \in \mathcal{C}^{-1+\kappa}(\mathbb{T}^d)$ and $\mathcal{L} = \frac{1}{2}\Delta + b\cdot\nabla = \mathcal{L}_0 + \mathcal{G}$. Define $\mathcal{D}(\mathcal{L}) = \mathcal{L}^{-1}L^2 = \{f \in H^1 : \mathcal{L}f \in L^2\}$. Then $\mathcal{D}(\mathcal{L})$ is dense and the closed operator $(\mathcal{D}(\mathcal{L}), \mathcal{L})$ generates a contraction semigroup $(S_t)_{t \geqslant 0}$ on $L^2(\mathbb{T}^d)$.*

Proof.

1. A priori bound: Let $(b_n)_{n\in\mathbb{N}} \subset C^\infty(\mathbb{T}^d)$ be a sequence of divergence free smooth approximations of b. Define the operators

$$\mathcal{L}_n = \mathcal{L}_0 + \mathcal{G}_n, \qquad \mathcal{G}_n f := b_n \varolessthan \nabla f + b_n \varogreaterthan \nabla f + \nabla\cdot(b_n \odot f) = b_n \cdot \nabla f,$$

   and the bounded resolvent operator

$$(1 - \mathcal{L}_n)^{-1} : L^2 \to H^1$$

   (we take this fact for granted but could also derive it from the following estimation with an additional regularization argument; recall that b_n is smooth). Testing $u_n := (1 - \mathcal{L}_n)^{-1} f$ against f, we get

$$\begin{aligned}\langle u_n, f\rangle &= \langle u_n, (1 - \mathcal{L}_n)u_n\rangle\\ &= \|u_n\|^2 + \|(-\mathcal{L}_0)^{1/2}u_n\|^2 - \langle u_n, \mathcal{G}_n u_n\rangle\\ &= \|u_n\|^2 + \|(-\mathcal{L}_0)^{1/2}u_n\|^2\\ &= \|u_n\|_{H^1}^2\end{aligned}$$

   Bounding

$$|\langle u_n, f\rangle| \leqslant \frac{1}{2}\|u_n\|_{H^1}^2 + \frac{1}{2}\|f\|_{H^{-1}}^2,$$

   we get uniformly in n

$$\|u_n\|_{H^1} \leqslant \|f\|_{H^{-1}}.$$

2. Construction of a domain $\tilde{\mathcal{D}}(\mathcal{L})$ for the resolvent: Let (e_k) be an orthonormal basis in L^2. By the a priori bound and a diagonal sequence argument we can extract a subsequence, denoted by abuse of notation again by (n), such that

$$(1 - \mathcal{L}_n)^{-1}e_k \to f_k := R_1 e_k$$

   weakly in H^1 (and strongly in $H^{1-\varepsilon}$ for all $\varepsilon > 0$, but we will not need this). We have for all $g \in C^\infty$:

$$\begin{aligned}\langle e_k, g\rangle &= \lim_n \langle (1 - \mathcal{L}_n)(1 - \mathcal{L}_n)^{-1}e_k, g\rangle\\ &= \lim_n \langle (1 - \mathcal{L}_n)^{-1}e_k, (1 - \mathcal{L}_n^*)g\rangle\\ &= \lim_n \langle (1 - \mathcal{L}_n)^{-1}e_k, (1 - \mathcal{L}^*)g\rangle + \langle (1 - \mathcal{L}_n)^{-1}e_k, (\mathcal{L}_n^* - \mathcal{L}^*)g\rangle.\end{aligned}$$

   Since $(1 - \mathcal{L}^*)g \in H^{-1}$, and since $((1 - \mathcal{L}_n)^{-1}e_k)_n$ converges weakly in H^1 to f we get that the first term converges to

$$\langle f_k, (1 - \mathcal{L}^*)g\rangle = \langle (1 - \mathcal{L})f_k, g\rangle.$$

[2.5]. This result is a consequence of classical perturbation theory. We set up the proof so carefully because later we will use similar arguments to derive a stronger result which does not follow from perturbation theory.

The remainder is bounded by

$$\begin{aligned}\langle(1-\mathcal{L}_n)^{-1}e_k,(\mathcal{L}_n^*-\mathcal{L}^*)g\rangle &\leqslant \|(1-\mathcal{L}_n)^{-1}e_k\|_{H^1}\|(\mathcal{L}_n^*-\mathcal{L}^*)g\|_{H^{-1}}\\ &\lesssim \|e_k\|_{H^{-1}}\|b_n-b\|_{\mathcal{C}^{-1+\kappa}}\|g\|_{H^1},\end{aligned}$$

which converges to 0. Therefore,

$$(1-\mathcal{L})f_k=e_k.$$

By linearity, this convergence extends to the span of the (e_k) and the resulting map $\sum_k a_k e_k \mapsto \sum_k a_k f_k$ is linear – note that here we only allow sums with finitely many non-zero a_k. The extension to infinite series follows by an approximation argument, using the a priori bound $\|(1-\mathcal{L}_n)^{-1}f\| \leqslant \|f\|$ which extends by Fatou's lemma to the limit. We write $R_1 f = \lim_{n\to\infty}(1-\mathcal{L}_n)^{-1}f$, again with weak convergence in H^1. We have thus constructed a space

$$\tilde{\mathcal{D}}(\mathcal{L})=\{R_1 f: f\in L^2\},$$

such that every $R_1 f\in\tilde{\mathcal{D}}(\mathcal{L})$ satisfies $(1-\mathcal{L})R_1 f=f\in L^2$.

3. Existence of the semigroup: We apply the Lumer-Phillips theorem (see below): We showed that $1-\mathcal{L}$ is surjective, and $\mathcal{L}$ is dissipative because for $f=R_1 g\in\tilde{\mathcal{D}}(\mathcal{L})$

$$\begin{aligned}\langle f,\mathcal{L}f\rangle &= \|f\|^2-\langle f,(1-\mathcal{L})f\rangle\\ &= \|f\|^2-\langle f,g\rangle\\ &\leqslant \liminf_{n\to 0}(\|f_n\|^2-\langle f_n,g\rangle),\end{aligned}$$

where $f_n := (1-\mathcal{L}_n)^{-1}g$ and we used weak lower semi-continuity of the norm. Now $g=(1-\mathcal{L}_n)f_n$ and therefore $\langle f_n,g\rangle=\|f_n\|^2+\|(-\mathcal{L}_0)^{1/2}f_n\|^2$, so finally

$$\langle f,\mathcal{L}f\rangle\leqslant\liminf_{n\to\infty}(-\|(-\mathcal{L}_0)^{1/2}f_n\|^2)\leqslant 0$$

and $\mathcal{L}$ is dissipative. Therefore, the Lumer-Phillips theorem shows that $(\tilde{\mathcal{D}}(\mathcal{L}),\mathcal{L})$ is closed and generates a contraction semigroup.

4. $\tilde{\mathcal{D}}(\mathcal{L})=\mathcal{D}(\mathcal{L})$ is the maximal domain: We clearly have

$$\mathcal{L}^{-1}L^2=(1-\mathcal{L})^{-1}L^2.$$

If we can show that $(1-\mathcal{L})$ is injective on H^1, then for $f\in(1-\mathcal{L})^{-1}L^2$ with $(1-\mathcal{L})f=g$ we must have $f=R_1 g$ and thus $f\in\tilde{\mathcal{D}}(\mathcal{L})$. For that purpose we note that actually

$$\langle f,\mathcal{L}f\rangle=-\|(-\mathcal{L}_0)^{1/2}f\|^2,\qquad f\in\tilde{\mathcal{D}}(\mathcal{L}).$$

Indeed, we can approximate $f\in H^1\supset\tilde{\mathcal{D}}(\mathcal{L})$ by smooth functions (f_n) and b by smooth divergence free functions b_n , so that we can perform the integration by parts rigorously, and then we use the fact that $\mathcal{G}$ maps H^1 boundedly to $H^{-1+\kappa}$ to pass to the limit.

Thus, if $f\in H^1$ satisfies $(1-\mathcal{L})f=0$, then

$$0 = \langle(1-\mathcal{L})f,f\rangle=\|f\|_{L^2}^2+\|(-\mathcal{L}_0)^{1/2}f\|_{L^2}^2,$$

and therefore $f=0$. □

Theorem. (Lumer-Phillips) *Let A be a linear operator defined on a linear subspace $D(A)$ of a Hilbert space H. Assume that*

i. A is dissipative, i.e. $\langle Ah,h\rangle\leqslant 0$ for all $h\in D(A)$, and

ii. $(\lambda-A)$ is surjective for some $\lambda>0$.

Then A is closed and it generates a contraction semigroup.

We have thus solved the Kolmogorov backward equation $(\partial_t-\mathcal{L})u=0$, $u(0)=f$ by $u(t)=S_t f$. By definition, $u\in C^1(\mathbb{R}_+,L^2)\cap C(\mathbb{R}_+,\mathcal{D}(\mathcal{L}))$, where $\mathcal{D}(\mathcal{L})$ is equipped with the norm $\|f\|_{L^2}+\|\mathcal{L}f\|_{L^2}$.

2.3 Duality of energy solutions and Kolmogorov backward equation

Here we combine the previous results to prove the uniqueness in law and Markov property of energy solutions. Let us first connect energy solutions with the operator $\mathcal{L}$:

Lemma 2.8. *Let $b\in\mathcal{C}^{-1+\kappa}$ and let $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}\in L^2(\mathbb{T}^d)$. Let Y be an energy solution of $\mathrm{d}Y_t=b(Y_t)\mathrm{d}t+\mathrm{d}B_t$, with $Y_0\sim\mu$, and let $(\mathcal{D}(\mathcal{L}),\mathcal{L})$ be the operator constructed in Theorem 2.5. Then Y solves the martingale problem for $\mathcal{L}$: For all $f\in\mathcal{D}(\mathcal{L})$ the process*

$$f(Y_t)-f(Y_0)-\int_0^t\mathcal{L}f(Y_s)\mathrm{d}s$$

is a martingale.

Proof. Let $f_\varepsilon=f*\rho_\varepsilon$ for a mollifier. Then $f_\varepsilon\in C^2$ and therefore by definition of energy solutions

$$f_\varepsilon(Y_t)=f_\varepsilon(Y_0)+\lim_{n\to\infty}\int_0^t\left(\frac{1}{2}\Delta+b_n\cdot\nabla\right)f_\varepsilon(Y_s)\mathrm{d}s+M_t^{f_\varepsilon}$$

is a martingale. Since $f_\varepsilon\in C^2$, we have that $\left(\frac{1}{2}\Delta+b_n\cdot\nabla\right)f_\varepsilon=(\mathcal{L}_0+\mathcal{G}_n)f_\varepsilon$ converges to $\mathcal{L}f_\varepsilon$ in H^{-1} as $n\to\infty$, and therefore $I\left(\left(\frac{1}{2}\Delta+b_n\cdot\nabla\right)f_\varepsilon\right)$ converges uniformly to $I(\mathcal{L}f_\varepsilon)$. Since $f\in\mathcal{D}(\mathcal{L})\subset H^1$ and $\mathcal{G}\colon H^1\to H^{-1}$ is continuous, we have that $\mathcal{L}f_\varepsilon$ converges in H^{-1} to $\mathcal{L}f=0$, and therefore $I(\mathcal{L}f_\varepsilon)_t$ converges in L^1 to $I(\mathcal{L}f)_t=\int_0^t\mathcal{L}f(Y_s)\mathrm{d}s$. Of course, also $f_\varepsilon(Y_s)\to f(Y_s)$ for $s=0,t$. Since the L^1-limit of martingales is a martingale, the claim follows. □

Lemma 2.9. *If Y is incompressible and it solves the martingale problem for $\mathcal{L}$, then for all $f\in C(\mathbb{R}_+,\mathcal{D}(\mathcal{L}))\cap C^1(\mathbb{R}_+,L^2)$ the process*

$$f(t,Y_t)-f(0,Y_0)-\int_0^t(\partial_s+\mathcal{L})f(s,Y_s)\mathrm{d}s$$

is a martingale.

Proof. (Sketch) We use time discretization, $t_k^n=\frac{k}{n}t$ and

$$\begin{aligned}f(t,Y_t)-f(0,Y_0) &= \sum_{k=0}^{n-1}\left(f(t_{k+1}^n,Y_{t_{k+1}^n})-f(t_k^n,Y_{t_k^n})\right)\\ &= \sum_{k=0}^{n-1}\left(f(t_{k+1}^n,Y_{t_k^n})-f(t_k^n,Y_{t_k^n})\right)+\sum_{k=0}^{n-1}\left(f(t_{k+1}^n,Y_{t_{k+1}^n})-f(t_{k+1}^n,Y_{t_k^n})\right),\end{aligned}$$

and use the fundamental theorem of calculus for the first term and the martingale problem for the second term. Then use incompressibility to show L^1-convergence as $n\to\infty$, and that the L^1-limit of martingales is a martingale. □

Theorem 2.10. *Let $b\in\mathcal{C}^{-1+\kappa}$ and let $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}\in L^2(\mathbb{T}^d)$. Let Y be an energy solution of $\mathrm{d}Y_t=b(Y_t)\mathrm{d}t+\mathrm{d}B_t$, with $Y_0\sim\mu$. Then the law of Y is unique and the finite-dimensional distributions are given by*

$$\mathbb{E}[f_1(Y_{t_1})\cdots f_n(Y_{t_n})] = \int_{\mathbb{T}^d}(S_{t_1}f_1S_{t_2-t_1}f_2\ldots S_{t_n-t_{n-1}}f_n)(x)\mu(\mathrm{d}x),$$

where (S_t) is the semigroup generated by $\mathcal{L}=\frac{1}{2}\Delta+b\cdot\nabla$. In particular, Y is a Markov process.

Proof. Markov property and uniqueness in law follow from the claimed identity. And the identity can be shown recursively in n, with the induction base and induction step using basically the same argument. So we only treat the case $n=1$. With our preparations this is now easy: For $f_1\in\mathcal{D}(\mathcal{L})$ consider $f(t,x)=S_{t_1-t}f_1(x)$. Then $f\in C(\mathbb{R}_+,\mathcal{D}(\mathcal{L}))\cap C^1(\mathbb{R}_+,L^2)$ and therefore the martingale problem with time-dependent functions gives

$$0=\mathbb{E}\left[f(t_1,Y_t)-f(0,Y_0)-\int_0^{t_1}(\partial_s+\mathcal{L})f(s,Y_s)\mathrm{d}s\right]=\mathbb{E}[f_1(Y_{t_1})-S_{t_1}f_1(Y_0)].$$

□

Remark 2.11.

i. We only proved uniqueness in law for the periodic diffusion $Y = X \bmod \mathbb{Z}^d$. To get uniqueness in law for X we have to work with non-periodic initial conditions for the Kolmogorov backward equation/semigroup and see it as an equation on $\mathbb{R}^d$. This is possible and the construction in Theorem 2.5 works verbatim. But the energy/admissibility estimate for non-periodic test functions is a bit subtle, because now X is not stationary any more and we do not have such nice formulas for the time-reversed process. But we still get a weaker version of the Itô trick and under the assumption $b \in \mathcal{C}^{-1+\kappa}$ we can adapt all our arguments to get also uniqueness in law for the non-periodic process X.

ii. The construction of X and proving its weak well-posedness can also be done without relying on stationarity, simply by Schauder estimates for the heat semigroup and by the paraproduct estimates. If we would naively treat this equation in regularity structures we might think that we need to construct $O\left(\frac{1}{\kappa}\right)$ tress for $\kappa \to 0$ to make sense of the dynamics, but using the Leibniz rule trick $b \odot \nabla u = \nabla \cdot (b \odot u)$ we can bypass this. This would lead to stronger results, but to set the stage for the following SPDE results and the extension discussed below, we preferred to use simple arguments based on energy estimates. Also, the supercritical extension sketched below really relies on the approach taken here.

iii. We could easily adapt the arguments to push the regularity of b to divergence free[2.6]

$$b \in \mathcal{C}^{-1,\log}, \qquad \mathcal{C}^{\alpha} = \left\{ f \in \mathcal{S}' : \|\Delta_j f\|_{L^\infty} \lesssim \frac{2^j}{j} \right\}.$$

Then $\mathcal{C}^{-1,\log} \not\subset \mathcal{C}^{-1+\kappa}$ for any $\kappa > 0$ and if we count regularity on a power scale this looks critical. (Of course it is still subcritical but only with logarithmic correction from criticality).

iv. To construct the semigroup, we used the "free" a priori bound $\|(1-\mathcal{L})^{-1} f\|_{H^1} \leqslant \|f\|_{H^{-1}}$. But we have another bound for free:

$$|(1-\mathcal{L})^{-1} f(x)| \leqslant \int_0^\infty e^{-t} |S_t f(x)| \mathrm{d}x \leqslant \|f\|_\infty,$$

i.e. $\|(1-\mathcal{L})^{-1} f\|_\infty \leqslant \|f\|_\infty$. Interpolating these bounds, we get well-posedness for divergence free

$$b \in B_{p,\infty}^{-\gamma,\log} = \left\{ f \in \mathcal{S}' : \|\Delta_j f\|_{L^p} \lesssim \frac{2^{j\gamma}}{j} \right\},$$

as long as $\gamma \leqslant 1$ and

$$p \geqslant \frac{2}{1-\gamma}.$$

The key property we have to verify is that then there still exists a function space $\mathcal{X}$ such that $\mathcal{L}$ is a bounded linear map from $\mathcal{X}$ to H^{-1}. And here we can take $\mathcal{X} = H^1 \cap L^\infty$ and use interpolation. Note that critical regularity would be $B_{p,\infty}^{-\gamma}$ for

$$p > \frac{d}{1-\gamma},$$

so in $d \geqslant 3$ this means we get well-posedness for supercritical equations. This interpolation of the L^∞ and H^1 regularity is inspired by [19].

2.4 Recap: What did we need

While we were studying a concrete problem on concrete function spaces, many of the arguments were quite robust. Let us collect all the ingredients we used so far:

i. Itô trick: We need a stationary Markov process with invariant measure μ and generator $\mathcal{L} = \mathcal{L}_0 + \mathcal{G}$, where $\mathcal{L}_0$ is symmetric and $\mathcal{G}$ is antisymmetric on $L^2(\mu)$. Then we can estimate for initial conditions with L^2 density with respect to the invariant measure

$$\mathbb{E}\left[\sup_{t \leqslant T} \left| \int_0^t f(Y_s) \mathrm{d}s \right| \right] \lesssim T^{1/2} \|(-\mathcal{L}_0)^{-1/2} f\|_{L^2(\mathbb{T})}.$$

2.6. Δ_j is the Littlewood-Paley block. If you don't know what that is, just ignore this part.

There are also similar L^p estimates, which we only needed for tightness. Those are a bit more tedious to state.

ii. Construction of the semigroup: We need again a generator $\mathcal{L}=\mathcal{L}_0+\mathcal{G}$, where $\mathcal{L}_0$ is symmetric and $\mathcal{G}$ is antisymmetric on $L^2(\mu)$. We used that

$$\mathcal{G}: H^1=(1-\mathcal{L}_0)^{1/2}L^2(\mu)\to(1-\mathcal{L}_0)^{-1/2}L^2(\mu)=H^{-1}$$

is a bounded operator, and similarly for $\mathcal{L}_0$. Most steps of the construction actually work if $\mathcal{G}$ is a continuous operator from H^1 to a bigger space, say $H^{-\gamma}$ for some $\gamma>0$. In that case the closure of $(\tilde{\mathcal{D}}(\mathcal{L}),\mathcal{L})$ still generates a contraction semigroup. We used $\gamma=1$ only to show that $\tilde{\mathcal{D}}(\mathcal{L})=\mathcal{D}(\mathcal{L})$ is the maximal domain.

iii. Duality of energy solution and Kolmogorov backward equation: Here we needed that for smooth $f^{(\varepsilon)}$ we have $\mathcal{G}_n f^{(\varepsilon)}\to\mathcal{G}f^{(\varepsilon)}$ in H^{-1} (which basically means that $\mathcal{G}$ has to be a continuous operator from H^γ for some $\gamma>0$ to H^{-1} and which is a very mild requirement) and that $\mathcal{G}f^{(\varepsilon)}\to\mathcal{G}f$ in H^{-1} for smooth approximations $f^{(\varepsilon)}$ of f; this last step means that $\mathcal{G}$ should be a continuous operator from a function space containing $\mathcal{D}(\mathcal{L})$ to H^{-1}.

Therefore, if $\mathcal{G}: H^1\to H^{-1}$ is bounded, all our analysis works. This is also what perturbation theory would suggest, because then we can treat $\mathcal{G}$ as a perturbation of $\mathcal{L}_0$. Unfortunately we will soon see that this is not satisfied for any of our SPDE examples.

3 Lecture 2: White noise invariant measure and Fock space

Here we will set up function spaces for singular SPDEs on which we can carry over and strengthen/extend the previous finite-dimensional arguments. These function spaces will be centered around the white noise and its chaos representation isometry to the Fock space.

3.1 Why white noise?

To analyze the (fractional) stochastic Burgers equation, we will work under the white noise measure and with its Fock space representation. Depending on the problem, this could be replaced by another measure, maybe Gaussian with another covariance or non-Gaussian. In any case we need a reference measure which in some sense corresponds to the dynamics: On infinite-dimensional spaces, measures are typically mutually singular with respect to each other, consider for example the law of B and of σB for $\sigma\in\mathbb{R}\setminus\{-1,1\}$, where B is a Brownian motion[3.1]. If our dynamics have a tractable invariant or quasi-invariant measure (meaning that if we start absolutely continuous with respect to a quasi-invariant measure, we stay absolutely continuous), then this is a natural reference measure.

But this does not yet answer why white noise should be invariant for the stochastic Burgers equation. To understand this, we again study a finite-dimensional example first:

Example 3.1. Let Σ be an invertible covariance matrix on $\mathbb{R}^d$ and consider an ODE with smooth vector field f,

$$\dot{x}(t)=f(x(t)),$$

such that

$$\nabla\cdot\Big(f(x)e^{-\frac{1}{2}\langle x,\Sigma^{-1}x\rangle}\Big)=0.$$

In that case we have with the $\mathcal{N}(0,\Sigma)$ distribution μ for all sufficiently nice test functions φ

$$\int_{\mathbb{R}^d} f(x)\cdot\nabla\varphi(x)\mu(\mathrm{d}x)=-\frac{1}{Z}\int_{\mathbb{R}^d}\varphi(x)\nabla\cdot\Big(f(x)e^{-\frac{1}{2}\langle x,\Sigma^{-1}x\rangle}\Big)\mathrm{d}x=0,$$

where Z is a normalization constant, and therefore μ is an invariant measure for the ODE. Note that our condition is always satisfied if f is divergence free,

$$\nabla\cdot f=0,\qquad\text{and}\qquad\langle f(x),\Sigma^{-1}x\rangle=0.$$

3.1. Exercise: Show that $\mathrm{law}((B_t)_{t\in[0,1]})$ and $\mathrm{law}((\sigma B_t)_{t\in[0,1]})$ are indeed mutually singular. Hint: consider the quadratic variation.

The second condition is equivalent to $E(x) := \frac{1}{2}\langle x, \Sigma^{-1} x\rangle$ being preserved by the dynamics:

$$\partial_t E(x(t)) = \langle \dot{x}(t), \Sigma^{-1} x(t)\rangle = \langle f(x(t)), \Sigma^{-1} x(t)\rangle = 0.$$

Example 3.2. Consider now the inviscid Burgers equation

$$\partial_t u = \partial_x u^2.$$

For simplicity we work on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, but all arguments extend to $\mathbb{R}$. Then the energy $E(u) = \frac{1}{2}\int u^2$ is formally preserved:

$$\partial_t E(u) = \langle u, \partial_t u\rangle_{L^2(\mathbb{T})} = \langle u, \partial_x u^2\rangle_{L^2(\mathbb{T})} = \left\langle 1, \frac{2}{3}\partial_x u^3 \right\rangle_{L^2(\mathbb{T})} = 0$$

by the periodic boundary conditions. The divergence free condition is a bit more difficult to formulate in this setting. Formally we can take the functional derivative D_x (differentiating in the direction of δ_x) and get

$$\int_{\mathbb{T}} D_x \partial_x u^2(x) \mathrm{d}x = 2\int_{\mathbb{T}} \partial_x (u(x)\delta(x)) \mathrm{d}x = 0,$$

because the integral of any derivative over $\mathbb{T}$ is zero. Therefore, the measure

$$\frac{1}{Z} e^{-\frac{1}{2}\langle u, u\rangle_{L^2}} \mathrm{d}u$$

is formally preserved. How should we interpret this measure? On $\mathbb{R}^d$ we have for $X \sim \frac{1}{Z} e^{-\langle x, x\rangle_{\mathbb{R}^d}} \mathrm{d}x$ that X is centered Gaussian with

$$\mathbb{E}[\langle X, x\rangle_{\mathbb{R}^d} \langle X, y\rangle_{\mathbb{R}^d}] = \langle x, y\rangle_{\mathbb{R}^d}, \qquad x, y \in \mathbb{R}^d.$$

Therefore, we would expect that $\eta \sim \frac{1}{Z} e^{-\frac{1}{2}\langle u, u\rangle_{L^2}} \mathrm{d}u$ is an L^2-valued centered Gaussian with

$$\mathbb{E}[\langle \eta, f\rangle_{L^2} \langle \eta, g\rangle_{L^2}] = \langle f, g\rangle_{L^2}, \qquad f, g \in L^2.$$

This is nearly correct, except that η is almost surely not L^2-valued but instead it takes values in a larger space of distributions ($H^{-1/2-\kappa}$ works for example). Such η is called a *white noise.*

We could imagine other Gaussian invariant measures for nonlinearities which conserve other quadratic energies. Indeed, many variations of Euler's nonlinearities (e.g. Euler, surface quasi-geostrophic or Leray α nonlinearities) formally have invariant Gaussian measures.

Definition 3.3. *Let H be a separable Hilbert space. A white noise on H is a centered Gaussian process $(\eta(h))_{h \in H}$ with covariance*

$$\mathbb{E}[\eta(g)\eta(h)] = \langle g, h\rangle.$$

Such a process always exists and one can show that it is equivalently characterized as a linear isometry from H to $L^2(\Omega)$ such that $\eta(h)$ is centered Gaussian for each $h \in H$.

Now we made sense of the candidate invariant measure for the inviscid Burgers equation. But that argument was purely formal and actually it is not clear if there are (even non-unique) weak solutions to the inviscid Burgers equation with invariant white noise distribution. The issue is that the white noise is only a generalized function and therefore $\partial_x u^2$ is not well-defined and the equation is very singular. We will see how to make sense of $\partial_x u^2$ as a distribution over the white noise space, but this is too singular to control the solutions.

Therefore, we have to add additional terms to the equation. And indeed we started with the goal of solving the stochastic Burgers equation[3.2]

$$\partial_t u = \Delta u + \partial_x u^2 + \sqrt{2}(-\Delta)^{1/2}\xi.$$

3.2. Here we replaced $\partial_x \xi$ by $(-\Delta)^{1/2}\xi$. But those two processes have the same law, and since we only care about the martingale formulation of the equation, this replacement has no influence for us.

Example 3.4. Let A be a symmetric, negative definite Fourier multiplier, i.e. $\mathcal{F}(A u)(k) = a(k)\mathcal{F}u(k)$ for some symmetric function $a: \mathbb{Z} \to (-\infty, 0]$. Then the equation

$$\partial_t v = A v + \sqrt{2} A^{1/2} \xi$$

preserves the white noise measure: Indeed, consider the Fourier basis $(e_k)_{k \in \mathbb{Z}}$ and note that

$$\partial_t \hat{v}_t(k) = v_t(A e_k) + \sqrt{2} \xi_t(A^{1/2} e_k) = a(k) \hat{v}_t(k) + \sqrt{-2 a(k)}\, \hat{\xi}_t(k).$$

This shows that, given the right initial condition, for each $t \geqslant 0$ the family $(\hat{v}_t(k))_k$ consists of independent complex-valued standard normal variables, up to the constraint $\hat{v}_t(-k) = \overline{\hat{v}_t(k)}$. But this is just the distribution of the white noise.

3.2 Chaos expansion and Fock space

From now on we focus on the white noise, but the following considerations work in principle for any Gaussian measure. Let η be a white noise on $\mathbb{T}$. Our next goal is to define function spaces on $L^2(\mu)$ for $\mu = \mathrm{law}(\eta)$ which resemble the Sobolev spaces from the finite-dimensional setting. We will achieve this with the chaos expansion. We start with some preparations:

Definition 3.5. (Symmetric functions) *For $n \in \mathbb{N}$ let $L_s^2(\mathbb{T}^n) \subset L^2(\mathbb{T}^n)$ be the space of* symmetric *functions in $L^2(\mathbb{T}^n)$, which are such that $\varphi(x_1, \ldots, x_n) = \varphi(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ for any permutation $\sigma \in \Sigma_n$ of $\{1, \ldots, n\}$. There is a canonical symmetrization map $\Pi: L^2(\mathbb{T}^n) \to L_s^2(\mathbb{T}^n)$,*

$$\Pi\varphi(x_1, \ldots, x_n) := \frac{1}{n!} \sum_{\sigma \in \Sigma_n} \varphi(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

For $\varphi \in L^2(\mathbb{T}^n)$ we write

$$\|\varphi\|_{L_s^2(\mathbb{T}^n)} := \|\Pi\varphi\|_{L^2(E^n)}.$$

Note that for $\varphi \in L^2(\mathbb{T}^n)$ we have by the triangle inequality for the $L^2(E^n)$-norm:

$$\|\varphi\|_{L_s^2(\mathbb{T}^n)} = \|\Pi\varphi\|_{L^2(\mathbb{T}^n)} \leqslant \frac{1}{n!} \sum_{\sigma \in \Sigma_n} \|\varphi(x_{\sigma(1)}, \ldots, x_{\sigma(n)})\|_{L^2(\mathbb{T}^n)} = \frac{1}{n!} n! \|\varphi\|_{L^2(\mathbb{T}^n)} = \|\varphi\|_{L^2(\mathbb{T}^n)}.$$

There exists a Brownian motion B such that

$$\eta(\varphi) = \int_0^1 \varphi(x) \mathrm{d}B_x, \qquad \varphi \in L^2(\mathbb{T}).$$

Indeed, it suffices to set $B_x = \eta(1_{[0,x]})$ (continuous modification). From now on we will always use B to denote this Brownian motion.

Definition 3.6. (Wiener-Itô integral) *For $\varphi \in L^2(\mathbb{T}^n)$ we define the n*-th Wiener Itô integral *as the iterated Itô integral*

$$W_n(\varphi) := W_n(\Pi\varphi) := n! \int_0^1 \int_0^{x_n} \cdots \int_0^{x_2} \Pi\varphi(x_1, \ldots, x_n) \mathrm{d}B_{x_1} \cdots \mathrm{d}B_{x_n}.$$

The factor $n!$ is explained by the fact that we are integrating over the arbitrary ordering $x_1 < x_2 < \cdots < x_n$ and there are $n!$ possible orderings which all would give the same integral.

Lemma 3.7. *The Wiener-Itô integral is a (multiple of a) linear isometry from $L_s^2(\mathbb{T}^n)$ to $L^2(\Omega)$:*

$$\|W_n(\varphi)\|_{L^2(\Omega)}^2 = \mathbb{E}[W_n(\varphi)^2] = n! \|\varphi\|_{L_s^2(E^n)}^2 = n! \|\Pi\varphi\|_{L^2(E^n)}^2.$$

Proof. This follows by repeated application of Itô's isometry. □

Definition 3.8. *We write $\mathcal{H}_n \subset L^2(\Omega)$ for the image of W_n, and we call $\mathcal{H}_n$ the n*-th Wiener-Itô chaos.

Note that by closedness of $L_s^2(\mathbb{T}^n)$ also each space $\mathcal{H}_n$ is closed and that $\mathcal{H}_n$ and $\mathcal{H}_m$ are orthogonal subspaces of $L^2(\Omega)$, which again follows by repeated use of Itô's isometry. Our next goal is to show the chaos representation

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) = \bigoplus_{n \geqslant 0} \mathcal{H}_n,$$

where

$$\mathcal{F} = \sigma(\eta(\varphi) : \varphi \in L^2(\mathbb{T})).$$

We will only sketch the argument.

Definition 3.9. *We define the* Hermite polynomials *recursively via*

$$H_0(x) = 1, \qquad H_n(x) = x H_{n-1}(x) - H_{n-1}'(x).$$

The first few Hermite polynomials are

$$H_0(x) = 1, \qquad H_1(x) = x, \qquad H_2(x) = x^2 - 1,$$
$$H_3(x) = x^3 - 3x, \qquad H_4(x) = x^4 - 6x^2 + 3.$$

One can show (exercise!) that $H_n' = n H_{n-1}$ and that the map $H_n(x, t) := t^{n/2} H_n\left(\frac{x}{\sqrt{t}}\right)$ solves the backward heat equation $\left(\partial_t + \frac{1}{2}\Delta\right) H_n(x, t) = 0$ with initial condition $H_n(x, 0) = \lim_{t \to 0} H_n(x, t) = x^n$. This leads to the following result:

Lemma 3.10. *Let M be a continuous local martingale with $M_0 = 0$. Then*

$$H_n(M_t, \langle M \rangle_t) = n \int_0^t H_{n-1}(M_s, \langle M \rangle_s) d M_s.$$

Proof. We apply Ito's formula to $H_n(M_t, \langle M \rangle_t)$: Since $H_n(0, 0) = 0$ and $\left(\partial_t + \frac{1}{2}\partial_x^2\right) H_n \equiv 0$, we get

$$\begin{aligned} H_n(M_t, \langle M \rangle_t) &= \int_0^t \partial_x H_n(M_s, \langle M \rangle_s) d M_s + \int_0^t \left(\partial_t + \frac{1}{2}\partial_x^2\right) H_n(M_s, \langle M \rangle_s) d\langle M \rangle_s \\ &= n \int_0^t H_{n-1}(M_s, \langle M \rangle_s) d M_s. \end{aligned}$$

□

Corollary 3.11. *For $\varphi \in L^2(\mathbb{T})$ we have with $\varphi^{\otimes n}(x_1, \ldots, x_n) := \varphi(x_1) \cdots \varphi(x_n)$:*

$$W_n(\varphi^{\otimes n}) = H_n(\eta(\varphi), \|\varphi\|_{L^2(\mathbb{T})}^2).$$

Proof. Consider the continuous martingale $M_t^{\varphi} = \eta(1_{[0, t]} \varphi)$. Then we get by repeated application of Lemma 3.10:

$$\begin{aligned} H_n(\eta(\varphi), \|\varphi\|_{L^2(\mathbb{T})}^2) &= H_n(M_1^{\varphi}, \langle M^{\varphi} \rangle_1) \\ &= n! \int_0^1 \int_0^{t_n} \cdots \int_0^{t_2} \mathrm{d} M_{t_1}^{\varphi} \mathrm{d} M_{t_2}^{\varphi} \cdots \mathrm{d} M_{t_n}^{\varphi} \\ &= W_n(\varphi^{\otimes n}). \end{aligned}$$

□

Corollary 3.12. *We have the chaos representation property*

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) = \bigoplus_{n \geqslant 0} \mathcal{H}_n,$$

where

$$\mathcal{F} = \sigma(\eta(\varphi) : \varphi \in L^2(\mathbb{T})).$$

In particular, every random variable $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ can be represented uniquely as

$$X = \sum_{n=0}^{\infty} W_n(\varphi_n), \qquad \varphi_n \in L_s^2(\mathbb{T}^n),$$

and

$$\mathbb{E}[X^2] = \sum_{n=0}^{\infty} n! \|\varphi_n\|_{L_s^2(\mathbb{T}^n)}^2$$

Proof. It suffices to apply Corollary 3.11 and to note that the monomial x^n can be written as a linear combination of $H_k(x, t)$ with $k \leqslant n$. Therefore, any random variable which is orthogonal to $\bigoplus_{n=0}^{\infty} \mathcal{H}_n$ is orthogonal to all polynomial of $\eta(\varphi)$, for all $\varphi \in L^2(\mathbb{T})$. See Theorem 1.1.1 of [17] for details. $\square$

The following result is very useful, but we will only apply it once in our lectures, and therefore we skip the proof in the lectures:

Theorem. (Gaussian hypercontractivity) *For all $p \in (0, \infty)$ there exists a constant $C_p > 0$ such that for all $n \in \mathbb{N}_0$ and all $\varphi \in L_s^2(\mathbb{T}^n)$:*

$$\mathbb{E}[|W_n(\varphi)|^p] \leqslant C_p^n (n!)^{p/2} \|\varphi\|_{L_s^2(\mathbb{T}^n)}^p = C_p^n \mathbb{E}[|W_n(\varphi)|^2]^{p/2}.$$

Proof. For $p < 2$ there is nothing to show, so let $p \geqslant 2$. By the Burkholder-Davis-Gundy inequality, together with the Minkowski inequality $\|\int_{\mathbb{T}} (\ldots) \mathrm{d}z\|_{L^{p/2}(\Omega)} \leqslant \int_{\mathbb{T}} \|\ldots\|_{L^{p/2}(\Omega)} dz$, we have

$$\begin{aligned}
\mathbb{E}[|W_n(\varphi)|^p] &\leqslant \mathbb{E}\left[\sup_{t \geqslant 0} |W_n(\varphi 1_{[0,t]}^{\otimes n})|^p\right] \leqslant C_p \mathbb{E}\left[\left(n^2 \int_{\mathbb{T}} W_{n-1}\big(\varphi(x, \cdot) 1_{[0,s]}^{\otimes (n-1)}\big)^2 dx\right)^{p/2}\right] \\
&\leqslant C_p \left(n^2 \int_{\mathbb{T}} \mathbb{E}\big[\big|W_{n-1}\big(\varphi(x_1, \cdot) 1_{[0,s_1]}^{\otimes (n-1)}\big)\big|^p\big]^{2/p} dx_1\right)^{p/2} \\
&\leqslant C_p^2 \left(n^2 (n-1)^2 \int_{\mathbb{T}} \int_{\mathbb{T}} \mathbb{E}\big[\big|W_{n-2}\big(\varphi(x_1, x_2, \cdot) 1_{[0,s_2]}^{\otimes (n-2)}\big)\big|^p\big]^{2/p} 1_{s_2 \leqslant s_1} dx_2 dx_1\right)^{p/2} \\
&\leqslant \cdots \leqslant C_p^n \left((n!)^2 \int_{\mathbb{T}} \ldots \int_{\mathbb{T}} |\varphi(x_1, \ldots, x_n)|^2 1_{s_n \leqslant \cdots \leqslant s_1} dx_n \ldots dx_1\right)^{p/2} \\
&= C_p^n \left(n! \int_{\mathbb{T}} \ldots \int_{\mathbb{T}} |\varphi(x_1, \ldots, x_n)|^2 dx_n \ldots dx_1\right)^{p/2},
\end{aligned}$$

where in the last step we used that φ is symmetric in its n arguments. The right hand side equals $C_p^n (n!)^{p/2} \|\varphi\|_{L^2(\mathbb{T}^n)}^p$, and this completes the proof. $\square$

4 Lecture 3: Energy solutions, the generator, and its semigroup

4.1 Sobolev spaces

From now on we assume that $\mathcal{F} = \sigma(\eta(\varphi) : \varphi \in L^2(\mathbb{T}))$. Then

$$L^2(\Omega) \simeq \Gamma L^2 := \bigoplus_{n \geqslant 0} L_s^2(\mathbb{T}^n),$$

and the space on the right hand side is called the (bosonic) Fock space. From now on we mostly identify $L^2(\Omega)$ with ΓL^2 and we interpret elements as ΓL^2 as random variables on Ω. Let us define two operators:

$$\begin{aligned}
\mathcal{N} \varphi_n &= n \varphi_n, && \text{number operator,} \\
\mathcal{L}_0 \varphi_n &= \Delta \varphi_n := (\partial_{x_1 x_1} + \cdots + \partial_{x_n x_n}) \varphi_n, && \text{Laplacian.}
\end{aligned}$$

Definition 4.1. (Sobolev type spaces) *For $\alpha, \beta \in \mathbb{R}$ we define*

$$\mathcal{H}^\alpha_\beta := \{\varphi = (\varphi_n)_{n\in\mathbb{N}_0}: (1-\mathcal{L}_0)^{\alpha/2}(1+\mathcal{N})^\beta \varphi \in \Gamma L^2\},$$

$$\begin{aligned}\|\varphi\|^2_{\mathcal{H}^\alpha_\beta} &:= \|(1-\mathcal{L}_0)^{\alpha/2}(1+\mathcal{N})^\beta\varphi\|^2 \\ &= \sum_{n=0}^\infty n!(1+n)^{2\beta}\|(1-\Delta)^{\alpha/2}\varphi_n\|^2_{L^2_s} \\ &= \sum_{n=0}^\infty n!(1+n)^{2\beta}\|\varphi_n\|^2_{H^\alpha_s}.\end{aligned}$$

So α measures the regularity of the kernels (φ_n), while β measures the decay of the kernels as $n\to\infty$. For $\alpha, \beta \geqslant 0$ this is a subspace of ΓL^2 and therefore elements of $\mathcal{H}^\alpha_\beta$ correspond to random variables. But if $\alpha < 0$ or $\beta < 0$ we have only "distributions on random variables", i.e. an element $\varphi \in \mathcal{H}^\alpha_\beta$ cannot be evaluated for $\omega\in\Omega$ and instead we can only make sense of the expectation $\mathbb{E}[X\varphi]$ for "nice" random variables X.

Using these function spaces, we would like to do a similar analysis as for the finite-dimensional singular diffusion example from the first lecture. Of course, that now there are two regularity indices suggests already that we will have to do something more complicated, but we will worry about that later.

4.2 Fock space representation of the generator, bounds

We start by considering the Ornstein-Uhlenbeck generator $\mathcal{L}_0$, i.e. the generator of

$$\partial_t v = \Delta v + \sqrt{2}(-\Delta)^{1/2}\xi.$$

Lemma 4.2. (Fock space representation of the OU generator) *For $\varphi\in\mathcal{H}^2_0$ we have $\varphi \in \mathcal{D}(\mathcal{L}_0)$ and*

$$(\mathcal{L}_0\varphi)_n = \Delta\varphi_n := (\partial_{x_1x_1} + \cdots + \partial_{x_nx_n})\varphi.$$

Proof. It suffices to prove this identity for $W_n(\varphi_n)$ with $\varphi_n \in H^2_s(\mathbb{T}^n)$. And by an approximation argument we may take $\varphi_n = \varphi^{\otimes n}$ with $\varphi\in C^2(\mathbb{T})$ such that $\|\varphi\|_{L^2(\mathbb{T})} = 1$. Then $W_n(\varphi^{\otimes n})(v) = H_n(v(\varphi))$, where $W_n(...)(v)$ is a Wiener-Itô integral with respect to the white noise v, and for the stationary solution of the Ornstein-Uhlenbeck process we have

$$\begin{aligned}\mathrm{d}H_n(v_t(\varphi)) &= H_n'(v_t(\varphi))v_t(\Delta\varphi)\mathrm{d}t + H_n''(v_t(\varphi))\|(-\Delta)^{1/2}\varphi\|^2_{L^2(\mathbb{T})}\mathrm{d}t + \mathrm{d}M_t \\ &= nH_{n-1}(v_t(\varphi))H_1(v_t(\Delta\varphi))\mathrm{d}t - n(n-1)H_{n-2}(v_t(\varphi))\langle\varphi,\Delta\varphi\rangle_{L^2(\mathbb{T})}\mathrm{d}t + \mathrm{d}M_t \\ &= nW_{n-1}(\varphi^{\otimes n-1})W_1(\Delta\varphi)\mathrm{d}t - n(n-1)W_{n-2}(\varphi^{\otimes n-2})\langle\varphi,\Delta\varphi\rangle_{L^2(\mathbb{T})}\mathrm{d}t + \mathrm{d}M_t,\end{aligned}$$

where we used that $H_n' = nH_{n-1}$. Now we use the multiplication rule for the Wiener chaos, see [17], Proposition 1.1.2, and rewrite the first term on the right hand side as

$$\begin{aligned}nW_{n-1}(\varphi^{\otimes n-1})W_1(\Delta\varphi) &= nW_n(\varphi^{\otimes n-1}\otimes\Delta\varphi) + n(n-1)W_{n-2}(\varphi^{\otimes n-2})\langle\varphi,\Delta\varphi\rangle_{L^2(\mathbb{T})} \\ &= W_n(\Delta\varphi^{\otimes n}) + n(n-1)W_{n-2}(\varphi^{\otimes n-2})\langle\varphi,\Delta\varphi\rangle_{L^2(\mathbb{T})}.\end{aligned}$$

The second term on the right hand side cancels with $-n(n-1)W_{n-2}(\varphi^{\otimes n-2})\langle\varphi,\Delta\varphi\rangle_{L^2(\mathbb{T})}$, and therefore

$$\mathrm{d}W_n(\varphi^{\otimes n})(v_t) = W_n(\Delta\varphi^{\otimes n})(v_t)\mathrm{d}t + \mathrm{d}M_t.$$

□

Lemma 4.3. (Fock space representation of the Burgers generator) *The operator $\mathcal{G}$ corresponding to the dynamics $\partial_x u^2$ is formally given by*

$$\mathcal{G} = \mathcal{G}_+ + \mathcal{G}_-, \qquad \mathcal{G}_\pm: L^2_s(\mathbb{T}^n) \to L^2_s(\mathbb{T}^{n\pm1}),$$

with

$$\begin{aligned}\mathcal{G}_+\varphi_n(x_{1:n+1}) &= \Pi(-n\delta(x_1-x_2)\partial_1\varphi_n(x_1,x_{3:n+1})), \\ \mathcal{G}_-\varphi_n(x_{1:n-1}) &= \Pi(-2n(n-1)\partial_1\varphi_n(x_1,x_1,x_{2:n-1})),\end{aligned}$$

where $x_{i:i+k}=(x_i,x_{i+1},\ldots,x_{i+k})$*, and where we recall that* Π *is the symmetrization operator.*

Proof. (Sketch) We take again $\varphi\in L^2(\mathbb{T})$ with $\|\varphi\|_{L^2}=1$ and consider $H_n(u_t(\varphi))$. Then

$$\mathrm{d}H_n(u_t(\varphi)) \;=\; -H_n'(u_t(\varphi))u_t^2(\partial_x\varphi)\mathrm{d}t=-nW_{n-1}(\varphi^{\otimes n-1})u_t^2(\partial_x\varphi)\mathrm{d}t.$$

Now, using that $\int C\partial_x\varphi(x)\mathrm{d}x=0$ for any constant C, we can replace u^2 by the Hermite polynomial and

$$u^2(\partial_x\varphi)=\int_{\mathbb{T}}W_2(\delta_y^{\otimes 2})\partial_y\varphi(y)\mathrm{d}y=W_2\bigg(\int_{\mathbb{T}}\delta_y^{\otimes 2}\partial_y\varphi(y)\mathrm{d}y\bigg).$$

The general multiplication rule for Wiener chaos variables from [17], Proposition 1.1.3, yields

$$W_{n-1}(f)W_2(g) \;=\; W_{n+1}(f\otimes g)+2(n-1)W_{n-1}(f\otimes_1 g)+(n-1)(n-2)W_{n-3}(f\otimes_2 g),$$

where

$$f\otimes_r g(x_{1:n+1-2r})=\Pi\bigg(\int f(x_{1:n-1-r},y_{1:r})g(x_{n-1-r+1:n+1-2r},y_{1:r})\mathrm{d}y_{1:r}\bigg).$$

Now observe that

$$\bigg(\int_{\mathbb{T}}\delta_y^{\otimes 2}\partial_y\varphi(y)\mathrm{d}y\bigg)(x_1,x_2)=\delta(x_1-x_2)\partial_{x_1}\varphi(x_1),$$

from where we can directly read off $\mathcal{G}_+$ and $\mathcal{G}_-$. It looks like there is still a contribution $\mathcal{G}_{-3}$, but note that

$$\int_{y_1,y_2}\varphi(y_1)\varphi(y_2)\delta(y_1-y_2)\partial_{y_1}\varphi(y_1)=\int_{y_1}\varphi(y_1)^2\partial_{y_1}\varphi(y_1)=\frac{1}{3}\int_y\partial_y(\varphi(y)^3)=0$$

and therefore $\mathcal{G}_{-3}=0$. □

Remark 4.4. You may feel very uneasy at this point about all the formal manipulations that we made in the last "proof" and that seem impossible to make rigorous. Instead, we should truncate the Burgers nonlinearity by introducing a Fourier projection operator $\mathcal{P}_\varepsilon u=\mathcal{F}^{-1}(\mathbb{1}_{[-\varepsilon^{-1},\varepsilon^{-1}]}\mathcal{F}u)$ and consider

$$\mathcal{P}_\varepsilon(\partial_x(\mathcal{P}_\varepsilon u)^2).$$

Then all the Dirac deltas get replaced by approximations $\mathcal{P}_\varepsilon\delta$ and we can make the previous arguments rigorous.

Lemma 4.5. (Bounds for the Burgers generator) *Let* $\mathcal{G}$ *be as in the previous lemma. Then we have for each* $\beta\in\mathbb{R}$*:*

$$\|\mathcal{G}_\pm\varphi\|_{\mathcal{H}^{-1}_{\beta-1}} \;\lesssim\; \|\varphi\|_{\mathcal{H}^1_\beta}.$$

For $\varphi\in\mathcal{H}^1_\alpha$ *and* $\psi\in\mathcal{H}^1_\beta$ *with* $\alpha+\beta\geqslant 1$ *we have*

$$\langle\mathcal{G}_+\varphi,\psi\rangle=-\langle\varphi,\mathcal{G}_-\psi\rangle,$$

and therefore

$$\langle\mathcal{G}\varphi,\psi\rangle=-\langle\varphi,\mathcal{G}\psi\rangle.$$

Proof. (Sketch) We can rewrite the operators using Fourier series, which we should actually still symmetrize but we can omit that by using the bound $\|\Pi f\|_{H^\alpha}\leqslant\|f\|_{H^\alpha}$:

$$\begin{aligned}\mathcal{F}(\mathcal{G}_+\varphi_n)(k_{1:n+1}) &= -2\pi i n(k_1+k_2)\hat{\varphi}_n(k_1+k_2,k_{3:n+1}),\\ \mathcal{F}(\mathcal{G}_-\varphi_n)(k_{1:n-1}) &= -2\pi i n(n-1)\,k_1\sum_{p+q=k_1}\hat{\varphi}_n(q,p,k_{2:n-1}).\end{aligned}$$

Then the claimed bound follows by plugging in the definition of the $\mathcal{H}^1_{\beta-1}$ norm and by some explicit estimations of Fourier series. See [10] for details, but it can be a good exercise to do the computation yourself.[4.1]

Once we have the bound, it suffices to prove $\langle\mathcal{G}_+\varphi,\psi\rangle=-\langle\varphi,\mathcal{G}_-\psi\rangle$ for $\varphi,\psi\in\mathcal{H}^\infty_\infty=\bigcap_m\mathcal{H}^m_m$. This follows again by a direct computation. □

4.1. If you try the computation, maybe go for $\|\mathcal{G}_\pm\varphi\|_{\mathcal{H}^{-1}_{\beta-k}}$ for *some* $k>0$ (maybe $k=2$ or so). Getting $k=1$ is a bit subtle, although in the recent paper [3] it is shown that the estimate even holds with $k=1/2$.

These bounds can be improved, for example we actually only lose $3/2$ degrees of regularity in the upper variable and not 2. But there are two important limits of the estimates: To control $\mathcal{G}_-\varphi$ we need $\varphi \in \mathcal{H}_\beta^{1/2+\varepsilon}$ for some $\varepsilon > 0$, and no matter how smooth φ is, $\mathcal{G}_+\varphi$ is never better than $\mathcal{H}_\infty^{-1/2-\varepsilon}$.

4.3 Construction of energy solutions

We have now two operators $\mathcal{L}_0$ and $\mathcal{G}$, such that $\mathcal{H}_0^\alpha = (1-\mathcal{L}_0)^{\alpha/2}\Gamma L^2$, such that $\langle \mathcal{G}\varphi, \varphi\rangle = 0$ for sufficiently nice φ (i.e. $\varphi \in \mathcal{H}_1^1$), and such that $\mathcal{G}$ is in some sense controlled by $\mathcal{L}_0$, i.e. $\mathcal{G}$ is a bounded operator from $\mathcal{H}_\beta^1 \to \mathcal{H}_{\beta-1}^{-1}$. If we ignore the lower regularity index, then this looks very similar to the assumptions that we made in the periodic diffusion example.

Let us see how far we can push the arguments from that example to our current setting. To construct energy solutions we consider the approximation

$$\partial_t u^\varepsilon = \Delta u^\varepsilon + \mathcal{P}_\varepsilon \partial_x (\mathcal{P}_\varepsilon u^\varepsilon)^2 + \sqrt{2(-\Delta)}\,\xi.$$

This equation is well-posed because we can decompose it as

$$u^\varepsilon = u^{\varepsilon,<} + u^{\varepsilon,>} := \mathcal{P}_\varepsilon u^\varepsilon + (1-\mathcal{P}_\varepsilon) u^\varepsilon,$$

where $u^{\varepsilon,<}$ solves the finite-dimensional equation

$$\partial_t u^{\varepsilon,<} = \Delta u^{\varepsilon,<} + \mathcal{P}_\varepsilon \partial_x (u^{\varepsilon,<})^2 + \sqrt{2(-\Delta)}\mathcal{P}_\varepsilon \xi, \qquad u^{\varepsilon,<} = \mathcal{P}_\varepsilon u^\varepsilon,$$

and where $u^{\varepsilon,>}$ solves the infinite-dimensional linear equation

$$\partial_t u^{\varepsilon,>} = \Delta u^{\varepsilon,>} + \sqrt{2(-\Delta)}(1-\mathcal{P}_\varepsilon)\xi \qquad u^{\varepsilon,>} = (1-\mathcal{P}_\varepsilon) u^\varepsilon.$$

Based on this representation we can check that the white noise is also invariant for u^ε and that u^ε has the generator $\mathcal{G}^\varepsilon$ which has the same bounds and antisymmetry properties as $\mathcal{G}$, uniformly in ε.

Lemma 4.6. (Itô trick) *Assume that* law $u_0^\varepsilon = \nu_\varepsilon$ *and that with the white noise law* μ *we have* $\frac{\mathrm{d}\nu_\varepsilon}{\mathrm{d}\mu} \in L^2(\mu)$. *Then we have uniformly in* ε

$$\mathbb{E}\left[\sup_{t \leqslant T}\left|\int_0^t \mathcal{L}_0 \varphi(u_s^\varepsilon)\mathrm{d}s\right|^p\right] \lesssim \left\|\frac{\mathrm{d}\nu_\varepsilon}{\mathrm{d}\mu}\right\|_{L^2(\mu)} T^{\frac{p}{2}} \|\mathcal{E}(\varphi)\|_{L^{2p}(\mu)}^p,$$

where

$$\mathcal{E}(\varphi) = 2\int_{\mathbb{T}} |\partial_x D_x \varphi|^2 \mathrm{d}x,$$

for

$$D_x \varphi_n(x_{1:n-1}) = n \varphi_n(x, x_{1:n-1}).$$

Proof. This follows by exactly the same arguments as in the finite-dimensional setting. Note that $\mathcal{E}(\varphi)$ corresponds to $|\nabla f|^2$ in finite dimensions, and this term comes from the quadratic variation of the forward and the backward martingale. □

The $L^{2p}(\mu)$ norm of $\mathcal{E}(\varphi)$ may be difficult to compute. But if $\varphi_n = 0$ for all $n \geqslant N$, then we can use Gaussian hypercontractivity to replace it by the $L^1(\mu)$ norm which is

$$\mathbb{E}\left[2\int_{\mathbb{T}} |\partial_x D_x \varphi|^2 \mathrm{d}x\right] = 2\|(-\mathcal{L}_0)^{1/2}\varphi\|_{L^2(\mu)}^2.$$

This last identity follows from a direct computation, or by Dynkin's formula for the quadratic variation,

$$\mathcal{E}(\varphi) = (\mathcal{L}_0 + \mathcal{G}^\varepsilon)\varphi^2 - 2\varphi(\mathcal{L}_0 + \mathcal{G}^\varepsilon)\varphi = \mathcal{L}_0 \varphi^2 - 2\varphi \mathcal{L}_0 \varphi,$$

where we used that $\mathcal{G}^\varepsilon$ is a first order differential operator and it satisfies Leibniz's rule.

To prove tightness of (u^ε), we will use Mitoma's criterion, which says that $(u^\varepsilon)_\varepsilon$ is tight in $C(\mathbb{R}_+, \mathcal{S}')$ if and only if for all $f \in C^\infty$ the family of real-valued processes $(u^\varepsilon(f))_\varepsilon$ is tight in $C(\mathbb{R}_+, \mathbb{R})$.

We also write

$$\mathcal{C} = \{(\varphi_n) : \exists N \text{ s.t. } \mathcal{F}\varphi_n(k) = 0 \text{ if } n \geqslant N \text{ or } |k| \geqslant N\}$$

for the *cylinder functions*, i.e. the polynomials depending on finitely many Fourier modes.

Theorem 4.7. (Existence of energy solutions, Gonçalves-Jara [4], Gubinelli-Jara [7]) *Assume that* law $u_0^\varepsilon = \nu_\varepsilon$ *and that with the white noise law* μ *we have* $\frac{\mathrm{d}\nu_\varepsilon}{\mathrm{d}\mu} \in L^2(\mu)$ *such that* $\frac{\mathrm{d}\nu_\varepsilon}{\mathrm{d}\mu}$ *converges in* $L^2(\mu)$ *to some* $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$. *Then* (u^ε) *is tight in* $C(\mathbb{R}_+, \mathcal{S}')$ *and any limit point* u *satisfies:*

i. *u weakly solves the stochastic Burgers equation with initial distribution ν: $u_0 \sim \nu$ and for all $f \in C^\infty(\mathbb{T})$*

$$u_t(f) = u_0(f) + \int_0^t u_s(\Delta f)\mathrm{d}s - \lim_{\delta \to 0} \int_0^t (\mathcal{P}_\delta u_s)^2(\partial_x f)\mathrm{d}s + M_t(f),$$

where $M(f)$ is a martingale with quadratic variation $\langle M(f)\rangle_t = 2t\|(-\Delta)^{1/2} f\|_{L^2}^2$. Moreover, for all cylinder functions $\varphi \in \mathcal{C}$ the process

$$\varphi(u_t) - \varphi(u_0) - \lim_{\delta \to 0} \int_0^t \mathcal{L}^\delta \varphi(u_s)\mathrm{d}s$$

is a martingale.

ii. *u is* incompressible: *For all $\varphi \in L^2(\mu)$ and all $t \geqslant 0$ we have*

$$\mathbb{E}[|\varphi(u_t)|] \lesssim \|\varphi\|_{L^2(\mathbb{T})}.$$

iii. *u is* admissible / satisfies an energy estimate*:*

$$\mathbb{E}\left[\sup_{t \leqslant T}\left|\int_0^t \varphi(u_s)\mathrm{d}s\right|\right] \lesssim \|(-\mathcal{L}_0)^{-1/2}\varphi\|_{L^2(\mu)}.$$

We call such u an energy solution *of the stochastic Burgers equation. In that case we can extend the map $I: L^2(\mu) \to L^1(\Omega, C([0,T]))$,*

$$I(\varphi)_t = \int_0^t \varphi(X_s)\mathrm{d}s,$$

continuously to $\mathcal{H}_0^{-1}$, and we denote the extension with the same symbol I.

Remark 4.8. This construction works whenever $\mathcal{G}\varphi \in \mathcal{H}_0^{-1}$ for all $\varphi \in \mathcal{C}$, and if for $\varphi \in \mathcal{C}$ we can improve either the regularity or the integrability in the estimate

$$\mathbb{E}\left[\sup_{t \leqslant T}\left|\int_0^t \varphi(u_s)\mathrm{d}s\right|^2\right]^{1/2} \lesssim T^{1/2}\|\varphi\|_{\mathcal{H}_0^{-1}},$$

under stationary initial conditions. If we can take the $L^p(\Omega)$ norm on the left hand side, then we could apply Kolmogorov's continuity criterion. By hypercontractivity, this is always the case if the invariant measure is Gaussian and if $\mathcal{G}$ comes from a polynomial nonlinearity, and if $\varphi \in \mathcal{C}$. Similarly, if we had some regularity to spare we could use an interpolation argument with the trivial bound where we pull the absolute value inside the integral, see [9], and in this way we could improve the factor $T^{1/2}$ to $T^{(1+\kappa)/2}$ for some $\kappa > 0$ which again is enough for Kolmogorov's continuity criterion.

For example, we can use the same construction for:

i. Fractional, multi-component Burgers equation:

$$\partial_t u = -(-\Delta)^\theta u + \partial_x(u \cdot \Gamma u) + \sqrt{2}(-\Delta)^{\theta/2}\xi,$$

if Γ is fully symmetric in its three arguments, and if $\theta > \frac{1}{2}$. Note that $\theta = \frac{3}{4}$ is scaling critical, and therefore the construction works in the supercritical regime.

ii. Stochastic surface quasi-geostrophic equation:

$$\begin{aligned}\partial_t \theta + u \cdot \nabla \theta &= (-\Delta)^\gamma \theta + \sqrt{2}(-\Delta)^{\gamma/2}\xi,\\ u &= \nabla^\perp(-\Delta)^{-1/2}\theta,\end{aligned}$$

for a space-time white noise ξ, and for $\gamma > 0$. The equation is critical for $\gamma = 1$.

iii. Vorticity formulation of 2d Navier-Stokes under the enstrophy measure:

$$\begin{aligned}\partial_t \omega + u \cdot \nabla \omega &= (-\Delta)^\gamma \omega + \sqrt{2}(-\Delta)^{\gamma/2}\xi,\\ u &= \nabla^\perp(-\Delta)^{-1}\omega,\end{aligned}$$

for a space-time white noise ξ, and for $\gamma \geqslant 0$. The equation is critical for $\gamma = 1/2$.

The bound

$$\|\mathcal{G}_\varepsilon \varphi\|_{\mathcal{H}^{-1}_{\beta-1}} \lesssim \|\varphi\|_{\mathcal{H}^1_\beta}$$

is also uniformly satisfied for many other critical or supercritical models, for example Landau-Lifshitz stochastic Navier-Stokes equations or stochastic Burgers equations in the supercritical dimensions $d \geqslant 3$, or in the weak coupling regime in the critical dimension $d = 2$, where also the anisotropic KPZ equation satisfies the same bound. But in those cases the operators $(\mathcal{G}_\varepsilon)$ do not converge and describing the limit of (u^ε) is more difficult, see [3] and the references therein.

4.4 A first construction of semigroups, failed duality

Now that we constructed energy solutions, we would like to use duality with the semigroup/Kolmogorov backward equation to prove their weak uniqueness. For this purpose we follow the same arguments as in Theorem 2.5:

Proposition 4.9. (Construction of the semigroup / solution of the Kolmogorov backward equation) *Let $\mathcal{G}\colon \mathcal{H}^1_0 \to \mathcal{H}^{-\alpha}_{-\beta}$ as a bounded operator, for some $\alpha, \beta \in \mathbb{R}$. Then there exists a dense domain $\tilde{\mathcal{D}}(\mathcal{L})$ for $\mathcal{L} = \mathcal{L}_0 + \mathcal{G}$ such that the closure of $(\tilde{\mathcal{D}}(\mathcal{L}), \mathcal{L})$ generates a contraction semigroup $(S_t)_{t \geqslant 0}$ on $L^2(\mathbb{T}^d)$.*

Proof. This follows from exactly the same arguments as in the finite-dimensional Theorem 2.5 if we take the approximations $\mathcal{L}^\varepsilon = \mathcal{L}_0 + \mathbb{1}_{\mathcal{N} \leqslant \frac{1}{\varepsilon}} \mathbb{1}_{\mathcal{L}_0 \leqslant \frac{1}{\varepsilon}} \mathcal{G} \mathbb{1}_{\mathcal{N} \leqslant \frac{1}{\varepsilon}} \mathbb{1}_{\mathcal{L}_0 \leqslant \frac{1}{\varepsilon}}$. In fact, Steps 1.-3. of the proof did not use that $\mathcal{G}$ maps H^1 to $H^{-1+\kappa}$, but just that continuously $\mathcal{G}$ maps H^1 to some $H^{-\gamma}$. We only needed a "good image space" for $\mathcal{G}$ to show that

$$\langle \varphi, \mathcal{L}\varphi \rangle = -\|(-\mathcal{L}_0)^{1/2} \varphi\|^2,$$

from which we deduced that our domain $\tilde{\mathcal{D}}(\mathcal{L})$ (which we constructed quite arbitrarily) agrees with the maximal domain $\mathcal{D}(\mathcal{L}) = \mathcal{L}^{-1} \mathcal{H}^0_0$. □

These conditions are satisfied for all the examples mentioned above, in particular for the fractional stochastic Burgers equation with $\theta > \frac{1}{2}$, so also in the supercritical regime $\theta \in \left(\frac{1}{2}, \frac{3}{4}\right)$.

And in principle we do not care very much whether $\tilde{\mathcal{D}}(\mathcal{L}) = \mathcal{D}(\mathcal{L})$ or not or whether $(\tilde{\mathcal{D}}(\mathcal{L}), \mathcal{L})$ is closed. But without $\tilde{\mathcal{D}}(\mathcal{L}) = \mathcal{D}(\mathcal{L})$ there is already an issue with uniqueness: We construct the domain by selecting some subsequences. It is conceivable that other subsequences converge to different limits and that this gives different domains, which lead to different semigroups. At this point, we do not know if this is the case.

Ultimately, we are interested in the uniqueness of energy solutions (and the duality between energy solution and semigroup would then also give the uniqueness of the semigroup). Looking back at the proof of this uniqueness, we needed to show that for $\varphi \in \mathcal{C}$

$$\mathcal{G}^m \varphi \to \mathcal{G}\varphi \qquad \text{in } \mathcal{H}^{-1}_0,$$

which basically means that $\mathcal{G}$ maps $\mathcal{C}$ to $\mathcal{H}^{-1}_0$ and which is again satisfied in all the examples mentioned above. But then we need to find for all $\varphi \in \tilde{\mathcal{D}}(\mathcal{L})$ some $(\varphi^\varepsilon) \subset \mathcal{C}$ such that $\varphi^\varepsilon \to \varphi$ in $\mathcal{H}^0_0$ (easy to find) and such that $\mathcal{G}\varphi^\varepsilon \to \mathcal{G}\varphi$ in $\mathcal{H}^{-1}_0$. Since $\mathcal{G}$ only maps $\mathcal{H}^1_0$ to $\mathcal{H}^{-1}_{-1}$ and since we only know $\tilde{\mathcal{D}}(\mathcal{L}) \subset \mathcal{H}^1_0$, this last condition is out of reach with our current arguments.

4.5 Commutator with the number operator, improved generator

Our goal is nearly in reach, but somehow we still need to improve the $\mathcal{N}$-regularity of the domain. We will achieve this with the help of the following commutator estimate:

Lemma 4.10. (Commutator estimate) *Let $\mathcal{G}$ be the operator corresponding to the Burgers nonlinearity. Then the commutator $[\mathcal{N}, \mathcal{G}] := \mathcal{N}\mathcal{G} - \mathcal{G}\mathcal{N}$ satisfies the estimate*

$$\|[\mathcal{N}, \mathcal{G}]\varphi\|_{\mathcal{H}^{-1}_0} \lesssim \|\varphi\|_{\mathcal{H}^1_1}.$$

Proof. We have

$$\begin{aligned}
[\mathcal{N},\mathcal{G}_+]\varphi_n &= (n+1)\mathcal{G}_+\varphi_n - \mathcal{G}_+(n\varphi_n) = \mathcal{G}_+\varphi_n\\
[\mathcal{N},\mathcal{G}_-]\varphi_n &= (n-1)\mathcal{G}_-\varphi_n - \mathcal{G}_-(n\varphi_n) = -\mathcal{G}_-\varphi_n,
\end{aligned}$$

so the claim follows from our estimates for $\mathcal{G}$. □

Theorem 4.11. (Domain with improved regularity) *Let $\mathcal{G}:\mathcal{H}_0^1\to\mathcal{H}_{-1}^{-1}$ and $[\mathcal{N},\mathcal{G}]:\mathcal{H}_1^1\to\mathcal{H}_0^{-1}$, both as bounded operators. Let $\mathcal{D}(\mathcal{L})=\mathcal{L}^{-1}\mathcal{H}_0^0=\{\varphi\in\mathcal{H}_0^1:\mathcal{L}\varphi\in\mathcal{H}_0^0\}$. Then:*

i. $(1-\mathcal{L})$ is injective on $\mathcal{H}_0^1$;

ii. $\mathcal{D}(\mathcal{L})$ is dense and the closed operator $(\mathcal{D}(\mathcal{L}),\mathcal{L})$ generates a contraction semigroup $(S_t)_{t\geqslant 0}$ on $\mathcal{H}_0^0$.

iii. $(1-\mathcal{L})\mathcal{C}$ is dense in $\mathcal{H}_0^{-1}$.

iv. We have $\langle(1-\mathcal{L})\varphi,\varphi\rangle=\|\varphi\|_{\mathcal{H}_0^1}^2$ for all $\varphi\in\mathcal{D}(\mathcal{L})$.

Proof.

i. $(1-\mathcal{L})$ is injective on $\mathcal{H}_0^1$: Let $\varphi\in\mathcal{H}_0^1$ be such that $(1-\mathcal{L})\varphi=0$. We introduce the operator

$$\mathcal{L}^\lambda=\mathcal{L}_0+(\lambda+\mathcal{N})\mathcal{G}(\lambda+\mathcal{N})^{-1}.$$

Then

$$\begin{aligned}
0 &= \lambda^2\langle(\lambda+\mathcal{N})^{-2}(1-\mathcal{L})\varphi,\varphi\rangle\\
&= \lambda^2\langle(\lambda+\mathcal{N})^{-2}(1-\mathcal{L}^\lambda)\varphi,\varphi\rangle+\lambda^2\langle(\lambda+\mathcal{N})^{-2}(\mathcal{L}^\lambda-\mathcal{L})\varphi,\varphi\rangle\\
&= \langle\lambda^2(\lambda+\mathcal{N})^{-2}(1-\mathcal{L}_0)\varphi,\varphi\rangle+\lambda^2\langle(\lambda+\mathcal{N})^{-2}(\mathcal{L}^\lambda-\mathcal{L})\varphi,\varphi\rangle,
\end{aligned}$$

where we used that $(\lambda+\mathcal{N})^{-1}\varphi\in\mathcal{H}_1^1$ and therefore $\langle\mathcal{G}(\lambda+\mathcal{N})^{-1}\varphi,(\lambda+\mathcal{N})^{-1}\varphi\rangle=0$ which leads to

$$\langle(\lambda+\mathcal{N})^{-2}(1-\mathcal{L}^\lambda)\varphi,\varphi\rangle=\langle(\lambda+\mathcal{N})^{-2}(1-\mathcal{L}_0)\varphi,\varphi\rangle.$$

Now we want to send $\lambda\to\infty$. By the dominated convergence theorem we have

$$\langle\lambda^2(\lambda+\mathcal{N})^{-2}(1-\mathcal{L}_0)\varphi,\varphi\rangle\to\|\varphi\|_{\mathcal{H}_0^1}^2,$$

while

$$\begin{aligned}
|\lambda^2\langle(\lambda+\mathcal{N})^{-2}(\mathcal{L}^\lambda-\mathcal{L})\varphi,\varphi\rangle| &\leqslant |\langle(\mathcal{L}^\lambda-\mathcal{L})\varphi,\varphi\rangle|\\
&\leqslant \|(\mathcal{L}^\lambda-\mathcal{L})\varphi\|_{\mathcal{H}_0^{-1}}\|\varphi\|_{\mathcal{H}_0^1}\\
&= \|[(\lambda+\mathcal{N}),\mathcal{G}](\lambda+\mathcal{N})^{-1}\varphi\|_{\mathcal{H}_0^{-1}}\|\varphi\|_{\mathcal{H}_0^1}\\
&= \|[\mathcal{N},\mathcal{G}](\lambda+\mathcal{N})^{-1}\varphi\|_{\mathcal{H}_0^{-1}}\|\varphi\|_{\mathcal{H}_0^1}\\
&\lesssim \|(\lambda+\mathcal{N})^{-1}\varphi\|_{\mathcal{H}_1^1}\|\varphi\|_{\mathcal{H}_0^1}\\
&= \|(\lambda+\mathcal{N})^{-1}(1+\mathcal{N})\varphi\|_{\mathcal{H}_0^1}\|\varphi\|_{\mathcal{H}_0^1}
\end{aligned}$$

converges to 0, again by the dominated convergence theorem. Therefore, $\varphi=0$.

ii. $\tilde{\mathcal{D}}(\mathcal{L})=\mathcal{D}(\mathcal{L})$: Recall that we constructed a map $R_1:\mathcal{H}_0^0\to\mathcal{H}_0^1$ (actually also $R_1:\mathcal{H}_0^{-1}\to\mathcal{H}_0^1$) such that $(1-\mathcal{L})R_1\varphi=\varphi$ for all $\varphi\in\mathcal{H}_0^0$, and that $\tilde{\mathcal{D}}(\mathcal{L})=R_1\mathcal{H}_0^0$. For $\varphi\in\mathcal{L}^{-1}\mathcal{H}_0^0$, let $\psi=(1-\mathcal{L})\varphi$, so that $(1-\mathcal{L})R_1(1-\mathcal{L})\psi=(1-\mathcal{L})\psi$. By injectivity of $(1-\mathcal{L})$ on $\mathcal{H}_0^1$ we must have $R_1(1-\mathcal{L})\psi=\psi$ and therefore $\psi\in\tilde{\mathcal{D}}(\mathcal{L})$.

iii. $(1-\mathcal{L})\mathcal{C}$ is dense in $\mathcal{H}_0^{-1}$: If $\varphi\in\mathcal{H}_0^{-1}$ is such that $\langle(1-\mathcal{L}_0)^{-1/2}(1-\mathcal{L})\psi,(1-\mathcal{L}_0)^{-1/2}\varphi\rangle=0$ for all $\psi\in\mathcal{C}$, then

$$0=\langle(1-\mathcal{L}_0)^{-1}(1-\mathcal{L})\psi,\varphi\rangle=\langle\psi,(1-\mathcal{L}^*)(1-\mathcal{L}_0)^{-1}\varphi\rangle$$

for all $\psi\in\mathcal{C}$, where we used that $\psi\in\mathcal{H}_\infty^1$ and $(1-\mathcal{L}_0)^{-1}\varphi\in\mathcal{H}_0^1$ to justify the integration by parts. Therefore, $(1-\mathcal{L}^*)(1-\mathcal{L}_0)^{-1}\varphi=0$. But since $\mathcal{L}^*=\mathcal{L}_0-\mathcal{G}$ and since we never used the sign of $\mathcal{G}$ it follows by exactly the same arguments as for $\mathcal{L}$ that $(1-\mathcal{L}^*)$ is injective and therefore $(1-\mathcal{L}_0)^{-1}\varphi=\varphi$, thus $\varphi=0$.

iv. $\langle(1-\mathcal{L})\varphi,\varphi\rangle=\|\varphi\|^2_{\mathcal{H}_0^1}$ for $\varphi\in\mathcal{D}(\mathcal{L})$: Since now we know that $(1-\mathcal{L})\mathcal{C}$ is dense in $\mathcal{H}_0^{-1}$, we can find φ^ε such that $(1-\mathcal{L})\varphi^\varepsilon$ converges to $(1-\mathcal{L})\varphi$ in $\mathcal{H}_0^{-1}$. Since R_1 is a bounded operator from $\mathcal{H}_0^{-1}$ to $\mathcal{H}_0^1$, we get that $R_1(1-\mathcal{L})\varphi^\varepsilon$ converges to $R_1(1-\mathcal{L})\varphi$ in $\mathcal{H}_0^1$. And we saw in step ii. that $R_1(1-\mathcal{L})\psi=\psi$, so that φ^ε converges to φ in $\mathcal{H}_0^1$. Therefore,

$$\langle(1-\mathcal{L})\varphi,\varphi\rangle=\lim_{\varepsilon\to0}\langle(1-\mathcal{L})\varphi^\varepsilon,\varphi^\varepsilon\rangle=\lim_{\varepsilon\to0}\langle(1-\mathcal{L}_0)\varphi^\varepsilon,\varphi^\varepsilon\rangle=\lim_{\varepsilon\to0}\|\varphi^\varepsilon\|^2_{\mathcal{H}_0^1}=\|\varphi\|^2_{\mathcal{H}_0^1},$$

where we used that $\langle\mathcal{G}\varphi^\varepsilon,\varphi^\varepsilon\rangle=0$ on cylinder functions φ^ε. □

The difference of this result compared to the previous one is very subtle, but these small improvements allow us to build the connection to energy solutions. If we would assume slightly better estimates for $\mathcal{G}$, say $\|\mathcal{G}\varphi\|_{\mathcal{H}_\lambda^{-1+\kappa}}\lesssim\|\varphi\|_{\mathcal{H}_1^1}$ for κ and/or $\lambda>0$, together with the commutator estimate, then we would be able to prove that $\mathcal{H}_\beta^1\cap\mathcal{D}(\mathcal{L})$ is dense for any $\beta\geqslant0$; alternatively, we could assume that $\|\mathcal{G}\varphi\|_{\mathcal{H}_0^{-1}}\leqslant\delta\|\varphi\|_{\mathcal{H}_1^1}$ for some sufficiently small $\delta>0$. Both of this essentially corresponds to a subcritical regime, and to handle the critical case where we lose one order of regularity both in the upper and in the lower variable without smallness assumption, we only know how to get the weaker results from the theorem. Next, we see that those are nonetheless sufficient to prove well-posedness of energy solutions.

4.6 Duality of energy solutions and semigroup

Here we combine the previous results to prove the uniqueness in law and Markov property of energy solutions. Let us first connect energy solutions with the operator $\mathcal{L}$:

Lemma 4.12. *Let u be an energy solution as in Theorem 4.7, and let $(\mathcal{D}(\mathcal{L}),\mathcal{L})$ be the operator constructed in Theorem 4.11. Then u solves the martingale problem for $\mathcal{L}$: For all $\varphi\in\mathcal{D}(\mathcal{L})$ the process*

$$\varphi(u_t)-\varphi(u_0)-\int_0^t\mathcal{L}\varphi(u_s)\mathrm{d}s$$

is a martingale.

Proof. Let $\varphi\in\mathcal{D}(\mathcal{L})$. Since $(1-\mathcal{L})\mathcal{C}$ is dense in $\mathcal{H}_0^{-1}$, we can find φ^ε such that $(1-\mathcal{L})\varphi^\varepsilon$ converges to $(1-\mathcal{L})\varphi$ in $\mathcal{H}_0^{-1}$. Since R_1 (from the proof of Theorem 4.11) is a bounded operator from $\mathcal{H}_0^{-1}$ to $\mathcal{H}_0^1$, we get that $R_1(1-\mathcal{L})\varphi^\varepsilon$ converges to $R_1(1-\mathcal{L})\varphi$ in $\mathcal{H}_0^1$. But in the proof of Theorem 4.11 we showed that $R_1(1-\mathcal{L})\psi=\psi$ and therefore φ^ε converges to φ in $\mathcal{H}_0^1$, and then also $\mathcal{L}\varphi^\varepsilon$ converges to $\mathcal{L}\varphi$ in $\mathcal{H}_0^{-1}$. Since $\varphi^\varepsilon\in\mathcal{C}$ we can use the martingale problem for cylinder functions, and we get that

$$\varphi^\varepsilon(u_t)-\varphi^\varepsilon(u_0)-\lim_{\delta\to0}\int_0^t\mathcal{L}^\delta\varphi^\varepsilon(u_s)\mathrm{d}s$$

is a martingale. Since $\varphi^\varepsilon\in\mathcal{C}\subset\mathcal{H}_1^1$, we get that $\mathcal{L}^\delta\varphi^\varepsilon$ converges to $\mathcal{L}\varphi^\varepsilon$ in $\mathcal{H}_0^{-1}$ as $\delta\to0$, and therefore by the admissibility condition we get that

$$\lim_{\delta\to0}\int_0^t\mathcal{L}^\delta\varphi^\varepsilon(u_s)\mathrm{d}s=I(\mathcal{L}\varphi^\varepsilon)_t,$$

with convergence in L^1. Now we use the admissibility condition once more, together with the fact that $\mathcal{L}\varphi^\varepsilon$ converges to $\mathcal{L}\varphi$ in $\mathcal{H}_0^{-1}$, to get $I(\mathcal{L}\varphi^\varepsilon)_t\to I(\mathcal{L}\varphi)_t$ as $\varepsilon\to0$, again with L^1-convergence. The incompressibility now yields $\varphi^\varepsilon(u_t)-\varphi^\varepsilon(u_0)\to\varphi(u_t)-\varphi(u_0)$ in L^1, and since an L^1-limit of martingales is again a martingale, the proof is complete. □

The next two results are now shown line-by-line with the same arguments as for the singular diffusion:

Lemma 4.13. *If u is incompressible and it solves the martingale problem for $\mathcal{L}$, then for all $\varphi\in C(\mathbb{R}_+,\mathcal{D}(\mathcal{L}))\cap C^1(\mathbb{R}_+,L^2)$ the process*

$$\varphi(t,u_t)-\varphi(0,u_0)-\int_0^t(\partial_s+\mathcal{L})\varphi(s,u_s)\mathrm{d}s$$

is a martingale.

Lemma 4.14. *Let u solve the martingale problem for $\mathcal{L}$, and let $(S_t)_{t \geqslant 0}$ be the semigroup constructed in Theorem 4.9 or 4.11. Then the law of u is unique and the finite-dimensional distributions are given by*

$$\mathbb{E}[\varphi_1(u_{t_1}) \cdots \varphi_n(u_{t_n})] = \int_{\mathbb{T}^d} (S_{t_1}\varphi_1 S_{t_2 - t_1}\varphi_2 \ldots S_{t_n - t_{n-1}}\varphi_n)(u)\mu(\mathrm{d}u),$$

where (S_t) is the semigroup generated by $\mathcal{L}$. In particular, u is a Markov process.

Combining these results, we obtain:

Theorem 4.15. *Let u be an energy solution as in Theorem 4.7, and let $(\mathcal{D}(\mathcal{L}), \mathcal{L})$ be the operator constructed in Theorem 4.11, with semigroup (S_t). Then the law of u is unique and u is a Markov process with transition function (S_t).*

5 Lecture 4: Applications

5.1 Critical equations

Example 5.1. The conditions of Theorem 4.15 are satisfied for:

i. The fractional stochastic Burgers equation on $\mathbb{R}_+ \times \mathbb{T}$ or $\mathbb{R}_+ \times \mathbb{R}$,

$$\partial_t u = -(-\Delta)^\theta u + \partial_x(u \cdot \Gamma u) + \sqrt{2}(-\Delta)^{\theta/2}\xi,$$

if $\Gamma \in \mathbb{R}^{d \times d \times d}$ is fully symmetric and if $\theta \geqslant \frac{3}{4}$. $\theta = \frac{3}{4}$ is scaling invariant and therefore scaling critical. We can construct energy solutions and semigroups for $\gamma > \frac{1}{2}$, but for $\gamma \in \left(\frac{1}{2}, \frac{3}{4}\right)$ we are lacking the duality and therefore the uniqueness.

ii. The surface quasi-geostrophic equation

$$\partial_t \theta + u \cdot \nabla \theta = -(-\Delta)^\gamma \theta + \sqrt{2}(-\Delta)^{\gamma/2}\xi, \qquad u = \nabla^\perp(-\Delta)^{-1/2}\theta$$

for $\gamma \geqslant 1$. $\gamma = 1$ is scaling invariant and therefore scaling critical. We can construct energy solutions and semigroups for $\gamma > 0$, but for $\gamma \in (0, 1)$ we are lacking the duality and therefore the uniqueness.

iii. 2d Navier-Stokes under the enstrophy measure:

$$\begin{aligned} \partial_t \omega + u \cdot \nabla \omega &= (-\Delta)^\gamma \omega + \sqrt{2}(-\Delta)^{\gamma/2}\xi, \\ u &= \nabla^\perp(-\Delta)^{-1}\omega, \end{aligned}$$

for a space-time white noise ξ, and for $\gamma \geqslant \frac{1}{2}$. $\gamma = 1/2$ is scaling invariant/critical. We can construct energy solutions and semigroups for $\gamma > 0$, but for $\gamma \in \left(0, \frac{1}{2}\right)$ we are lacking the duality and therefore the uniqueness.

5.2 Hairer-Quastel universality

This is from a joint work with Massimiliano Gubinelli [8], although written somewhat differently.

Consider the nonlinear SPDE

$$\partial_t v_\varepsilon = \Delta v_\varepsilon + \varepsilon^{1/2}\mathcal{P}_{1/2}\partial_x F(\mathcal{P}_{1/2}v_\varepsilon) + \sqrt{2(-\Delta)}\xi$$

on $\mathbb{R}_+ \times \varepsilon^{-1}\mathbb{T}$, where ξ is a space-time white noise, F is a suitable nonlinearity, and where we recall that $\mathcal{P}_{1/2}$ is the projection onto the Fourier modes $|\cdot| \leqslant 1/2$. For now we assume that $F \in L^2(\nu)$ for the standard normal distribution ν, and that $\int_{\mathbb{R}} F(x)x\nu(\mathrm{d}x) = 0$ (which is for simplicity, and which is for example the case if F is even). Let

$$u_\varepsilon(t, x) = \varepsilon^{-1/2}v_\varepsilon(\varepsilon^{-2}t, \varepsilon^{-1}x),$$

so that u_ε solves

$$\partial_t u_\varepsilon = \Delta u_\varepsilon + \varepsilon^{-1}\mathcal{P}_{(2\varepsilon)^{-1}}\partial_x F(\varepsilon^{1/2}\mathcal{P}_{(2\varepsilon)^{-1}}u_\varepsilon) + \sqrt{2(-\Delta)}\xi$$

on $\mathbb{R}_+ \times \mathbb{T}$, where ξ is a new space-time white noise (now on $\mathbb{R}_+ \times \mathbb{T}$) which we denote by the same symbol for simplicity. Then u_ε is invariant under the law of the white noise. This can be shown similarly as for the stochastic Burgers equation. The generator is $\mathcal{L}_0 + \mathcal{G}^\varepsilon$, where $\mathcal{G}^\varepsilon$ corresponds to $\varepsilon^{-1}\mathcal{P}_{(2\varepsilon)^{-1}}\partial_x F(\varepsilon^{1/2}\mathcal{P}_{(2\varepsilon)^{-1}}u_\varepsilon)$.

To compute $\mathcal{G}^\varepsilon$, we would like to express $F(\varepsilon^{1/2}\mathcal{P}_{\varepsilon^{-1}}u_\varepsilon(x))$ as a series of Wiener integrals. For this purpose it is convenient to assume that the white noise u_ε has vanishing zero Fourier mode. Since the zero Fourier mode $\int_{\mathbb{T}} u_\varepsilon$ is conserved by the dynamics (every term on the right hand side is a derivative), this "average-zero" white noise is also invariant and we will work with this invariant measure. Then

Note that $\varepsilon^{1/2}\mathcal{P}_{\varepsilon^{-1}}u_\varepsilon(x)$ is centered Gaussian with

$$\mathbb{E}[(\varepsilon^{1/2}\mathcal{P}_{(2\varepsilon)^{-1}}u_\varepsilon(x))^2] = \varepsilon\mathbb{E}[u_\varepsilon(\delta_x^\varepsilon)^2] = \varepsilon\|\delta_x^\varepsilon - \hat{\delta_x^\varepsilon}(0)\|_{L^2}^2 = \varepsilon\|\delta^\varepsilon - \hat{\delta^\varepsilon}(0)\|_{L^2}^2 = \varepsilon\sum_{k\neq 0}|\hat{\delta^\varepsilon}(k)|^2 = \varepsilon\sum_{|k|\leqslant \frac{1}{2\varepsilon}, k\neq 0} 1,$$

where $\delta^\varepsilon = \mathcal{F}^{-1}\mathbb{1}_{[-(2\varepsilon)^{-1},(2\varepsilon)^{-1}]}$ and $\delta_x^\varepsilon = \delta^\varepsilon(x-\cdot)$. If $\frac{1}{2\varepsilon}$ is an integer, then the right hand side is 1. From now on we always make this assumption. Of course it is possible to adapt the analysis to treat also the white noise with random zero Fourier mode and to treat general ε, but the formulas will be easier if $\varepsilon^{1/2}\mathcal{P}_{(2\varepsilon)^{-1}}u_\varepsilon(x)$ is a standard normal variable.

Then we can decompose, using that $\left(\frac{1}{\sqrt{m!}}H_m\right)_{m\in\mathbb{N}_0}$ (Hermite polynomials) is an orthonormal basis in $L^2(\nu)$:

$$\begin{aligned} F(\varepsilon^{1/2}\mathcal{P}_{(2\varepsilon)^{-1}}u_\varepsilon(x)) &= \sum_{m=0}^\infty \mathbb{E}\left[\frac{1}{\sqrt{m!}}H_m(X)F(X)\right]\frac{1}{\sqrt{m!}}H_m(\varepsilon^{1/2}\mathcal{P}_{(2\varepsilon)^{-1}}u_\varepsilon(x)) \\ &= \sum_{m=0}^\infty \underbrace{\frac{1}{m!}\mathbb{E}[H_m(X)F(X)]}_{=:c_m(F)}H_m(\varepsilon^{1/2}\mathcal{P}_{(2\varepsilon)^{-1}}u_\varepsilon(x)), \end{aligned}$$

for $X\sim\nu$ standard normal. By assumption the term for $m=1$ vanishes, and the term for $m=0$ is killed by the derivative ∂_x – without the derivative it would give a diverging contribution which we would have to remove by a renormalization.

Recall that

$$H_m(\varepsilon^{1/2}\mathcal{P}_{(2\varepsilon)^{-1}}u_\varepsilon(x)) = H_m(u_\varepsilon(\varepsilon^{1/2}\delta_x^\varepsilon)) = W_m((\varepsilon^{1/2}\delta_x^\varepsilon)^{\otimes m})(u_\varepsilon) = \varepsilon^{m/2}W_m((\delta_x^\varepsilon)^{\otimes m})(u_\varepsilon),$$

and therefore

$$\mathcal{G}^\varepsilon\varphi = \sum_{m\geqslant 2}\mathcal{G}^{\varepsilon,m}\varphi := \sum_{m\geqslant 2} c_m(F)\varepsilon^{\frac{m}{2}-1}\int_{\mathbb{T}} W_m(\partial_x(\delta_x^\varepsilon)^{\otimes m})D_x\varphi\mathrm{d}x,$$

where $D_x W_n(\varphi_n) = n W_{n-1}(\varphi_n(x,\cdot))$ is the Malliavin derivative. I expect that with similar arguments as for the quadratic term, we get for $m\geqslant 3$

$$\left\|\int_{\mathbb{T}} W_m(\partial_x(\delta_x^\varepsilon)^{\otimes m})D_x\varphi\mathrm{d}x\right\|_{\mathcal{H}_0^{-1}} \lesssim \varepsilon^{-\frac{m-3}{2}}\|\varphi\|_{\mathcal{H}^1_{\frac{m}{2}}},$$

where for $m=3$ we interpret $\varepsilon^{-\frac{m-3}{2}} = \sqrt{\log\frac{1}{\varepsilon}}$. But this is very tedious, so let us take a lighter approach and restrict to test functions φ in the first chaos, which effectively means that in the definition of an energy solution, Theorem 4.7, we are only going to check the first part of i., i.e. that u is a weak solution of the stochastic Burgers equation with the nonlinearity defined by approximation, and we will check ii. and iii., but we will not show that u solves the martingale problem for cylinder functions. For the subcritical Burgers equation we can deduce the martingale problem for cylinder functions from the other properties, see [10], basically by giving the drift higher time regularity than $\frac{1}{2}$ and using Young integration techniques. Only for critical equations it is crucial to directly prove the martingale problem for cylinder functions, because then we do not know how to get this from the weak solution description.

With that out of the way, let $\varphi(u)=u(f)$ for $\varphi_1\in C^\infty$. Then

$$\begin{aligned}\int_{\mathbb{T}} W_m(\partial_x(\delta_x^\varepsilon)^{\otimes m}) D_x\varphi \mathrm{d}x &= \int_{\mathbb{T}} W_m(\partial_x(\delta_x^\varepsilon)^{\otimes m}) f(x)\mathrm{d}x\\ &= -W_m\Big(\int_{\mathbb{T}}(\delta_x^\varepsilon)^{\otimes m}\partial_x f(x)\mathrm{d}x\Big),\end{aligned}$$

and taking the Fourier transform:

$$\begin{aligned}\mathcal{F}\Big(\int_{\mathbb{T}}(\delta_x^\varepsilon)^{\otimes m}\partial_x f(x)\mathrm{d}x\Big)(k_{1:n}) &= \int e^{-2\pi i k\cdot z}\int_{\mathbb{T}}(\delta_x^\varepsilon)^{\otimes m}(z_{1:m})\partial_x f(x)\mathrm{d}x\\ &= \int\prod_i \hat{\delta_x^\varepsilon}(k_i)\partial_x f(x)\mathrm{d}x\\ &= \mathbb{1}_{|k|_\infty\leqslant(2\varepsilon)^{-1}}(2\pi i(k_1+\cdots+k_m))\hat{f}(k_1+\cdots+k_m),\end{aligned}$$

so that

$$\begin{aligned}&\Big\|(1-\mathcal{L}_0)^{-1/2}\Big(\int_{\mathbb{T}}(\delta_x^\varepsilon)^{\otimes m}\partial_x f(x)\mathrm{d}x\Big)\Big\|^2\\ =&m!\sum_{|k_{1:m}|_\infty\leqslant(2\varepsilon)^{-1}}(1+|2\pi k|_2^2)^{-1}|(2\pi i(k_1+\cdots+k_m))\hat{f}(k_1+\cdots+k_m)|^2\\ \simeq& m!\sum_\ell\sum_{|k_{1:m}|_\infty\leqslant(2\varepsilon)^{-1}}\mathbb{1}_{k_1+\cdots+k_m=\ell}(1+|k|_2^2)^{-1}|\ell\hat{f}(\ell)|^2,\end{aligned}$$

and estimating the sum by an integral we get

$$\sum_{|k_{1:m}|_\infty\leqslant(2\varepsilon)^{-1}}\mathbb{1}_{k_1+\cdots+k_m=\ell}(1+|k|_2^2)^{-1} \lesssim \varepsilon^{-(m-3)},$$

again with interpretation $\varepsilon^{-(m-3)}=\log\frac{1}{\varepsilon}$ for $m=3$.

Therefore, we get for $\varphi(u)=u(f)$:

$$\begin{aligned}\|\mathcal{G}^\varepsilon\varphi-\mathcal{G}^{\varepsilon,2}\varphi\|_{\mathcal{H}_0^{-1}}^2 &= \Big\|\sum_{m\geqslant3}c_m(F)\varepsilon^{\frac{m}{2}-1}W_m\Big(\int_{\mathbb{T}}(\delta_x^\varepsilon)^{\otimes m}\partial_x f(x)\mathrm{d}x\Big)\Big\|_{\mathcal{H}_0^{-1}}^2\\ &= \sum_{m\geqslant3}c_m(F)^2\varepsilon^{m-2}m!\Big\|(1-\mathcal{L}_0)^{-1/2}\Big(\int_{\mathbb{T}}(\delta_x^\varepsilon)^{\otimes m}\partial_x f(x)\mathrm{d}x\Big)\Big\|^2\\ &\lesssim \sum_{m\geqslant3}c_m(F)^2\varepsilon^{m-2}m!\varepsilon^{-(m-3)}\\ &\lesssim \varepsilon\Big(1+\log\frac{1}{\varepsilon}\Big)\sum_{m\geqslant3}c_m(F)^2m!\\ &\lesssim \varepsilon\log\frac{1}{\varepsilon}\sum_{m\geqslant3}\frac{1}{m!}\mathbb{E}[H_m(X)F(X)]\\ &\leqslant \varepsilon\log\frac{1}{\varepsilon}\mathbb{E}[F(X)^2],\end{aligned}$$

where in the last step we used that $\Big(\frac{1}{\sqrt{m!}}H_m\Big)$ is an orthonormal basis in $L^2(\nu)$ and Parseval's identity. Therefore, $\mathcal{G}^\varepsilon\varphi$ converges in $\mathcal{H}_0^{-1}$ to the Burgers generator $\mathcal{G}\varphi$ and from this we readily get that any weak limit of (u_ε) is an energy solution to the stochastic Burgers equation.

5.3 Gaussian fluctuations for periodic KPZ

This part is from a joint work in progress with Huanyu Yang. The results for KPZ are known and were previously shown by Gu-Komorowski to do using the Cole-Hopf transform. Using energy solutions and the generator $\mathcal{L}$, we could extend them to fractional, multi-component Burgers equations by the same arguments.

Definition 5.2. *A stochastic process h with values in $C(\mathbb{R}_+, C(\mathbb{T}))$ is called a stationary* energy solution *of the periodic KPZ equation*

i. For all $f \in C^\infty(\mathbb{T})$

$$h_t(f) = h_0(f) + \int_0^t h_s(\Delta f)\mathrm{d}s + \lim_{\delta \to 0} \int_0^t ((\partial_x \mathcal{P}_\delta h_s)^2 - \|\mathcal{P}_\delta\|_{L^2}^2)(f)\mathrm{d}s + M_t(f),$$

where $M(f)$ is a martingale with quadratic variation $\langle M(f)\rangle_t = 2t\|f\|_{L^2}^2$.

ii. The distributional derivative $u = \partial_x h$ is a stationary energy solution of the stochastic Burgers equation, i.e. u_0 is a white noise, and moreover $\hat{u}_t(0) = 0$ for all $t \geqslant 0$ (so strictly speaking u_t is not a white noise but a mean free white noise).

Such stationary energy solutions of KPZ exist and they can be constructed in the same way as energy solutions of the stochastic Burgers equation.

Our goal is to prove the following result:

Theorem 5.3. *Let h be a stationary energy solution of the periodic KPZ equation. Then there exists $\sigma^2 \in (0, \infty)$ such that for each $x \in \mathbb{T}$ we have that*

$$\frac{1}{\sqrt{t}} h(t, x) \xrightarrow{w} \mathcal{N}(0, \sigma^2)$$

weakly, as $t \to \infty$.

Probably the assumption of stationary initial conditions can be relaxed.

This result may be surprising at first, because famously the KPZ equation has non-Gaussian fluctuations under the scaling $t^{-1/3}h(t, t^{2/3}x)$. But since we are on the torus, we cannot rescale space and the Burgers equation decorrelates exponentially fast, and this gives forces the fluctuations to be Gaussian. To see this, we use the mild formulation,

$$\begin{aligned} h_t(x) &= p_t * h_0(x) + \lim_{\delta \to 0} \int_0^t ((\partial_x \mathcal{P}_\delta h_s)^2 - \|\mathcal{P}_\delta\|_{L^2}^2)(p_{t-s}(x - \cdot))\mathrm{d}s + M_t^{p_{t-\cdot}(x-\cdot)} \\ &= p_t * h_0(x) + \lim_{\delta \to 0} \int_0^t ((\mathcal{P}_\delta u_s)^2 - \|\mathcal{P}_\delta\|_{L^2}^2)(p_{t-s}(x - \cdot))\mathrm{d}s + M_t^{p_{t-\cdot}(x-\cdot)} \end{aligned}$$

where p is the heat kernel of the periodic Laplacian

$$\mathcal{F}p_t(k) = e^{-|2\pi k|^2 t},$$

and where $\left(M_s = M_s^{p_{t-\cdot}(x-\cdot)}\right)_{s \in [0,t]}$ is a martingale with quadratic variation

$$\langle M^{p_{t-\cdot}(x-\cdot)}\rangle_t = \int_0^t \|p_{t-s}(x - \cdot)\|_{L^2}^2 \mathrm{d}s.$$

This quadratic variation is deterministic, so M is Gaussian (of course it is, because it is built by integrating the space-time white noise against a deterministic function). We will see that with the constant test function $\mathbb{1}$ we can interpret, up to a small error

$$\lim_{\delta \to 0} \int_0^t ((\mathcal{P}_\delta u_s)^2 - \|\mathcal{P}_\delta\|_{L^2}^2)(p_{t-s}(x - \cdot))\mathrm{d}s \simeq \int_0^t W_2\left(\int_{\mathbb{T}} \delta_x^{\otimes 2} \mathbb{1}(x)\mathrm{d}x\right)(u_s)\mathrm{d}s.$$

The functional $W_2(\ldots)(u_s)$ is in $\mathcal{H}_0^{-1}$, and since u decorrelates exponentially quickly we can decompose for $t = m$

$$\int_0^t W_2\left(\int_{\mathbb{T}} \delta_x^{\otimes 2} \mathbb{1}(x)\mathrm{d}x\right)(u_s)\mathrm{d}s = \sum_{i=0}^{m-1} \int_i^{i+1} W_2\left(\int_{\mathbb{T}} \delta_x^{\otimes 2} \mathbb{1}(x)\mathrm{d}x\right)(u_s)\mathrm{d}s$$

with L^2 random variables that are nearly independent. Therefore, by the central limit theorem also this contribution should give rise to a Gaussian limit.

To make this intuition rigorous, we first replace the point evaluation $h_t(x)$ by testing against $\mathbb{1}$, which is more regular:

Lemma 5.4. *We have*

$$\lim_{t\to\infty}\left\|\frac{1}{\sqrt{t}}(h_t(x)-h_t(\mathbb{1}))\right\|_{L^2}=0.$$

Proof. This is based on the fact that by Parseval's identity

$$\|p_t(x-\cdot)-\mathbb{1}\|_{L^2}^2=\sum_{k\neq 0} e^{-|2\pi k|^2 t}\lesssim e^{-|2\pi|^2 t}$$

for large t. Moreover,

$$h_t(\mathbb{1})=h_0(\mathbb{1})+0+\lim_{\delta\to 0}\int_0^t((\mathcal{P}_\delta u_s)^2-\|\mathcal{P}_\delta\|_{L^2}^2)(\mathbb{1})\mathrm{d}s+M_t(\mathbb{1}).$$

We subtract the mild formulation and bound the differences of the three different terms separately. For the initial condition we have

$$|p_t*h_0(x)-h_0(\mathbb{1})|=|h_0(p_t(x-\cdot)-\mathbb{1})|\leqslant\|h_0\|_{L^2}\|p_t(x-\cdot)-1\|_{L^2}\lesssim\|h_0\|_{L^2}e^{-Ct},$$

so this vanishes even without dividing by $\sqrt{t}$. For the drift we have to adapt the energy estimate to allow time-dependent functions. This can be done by the same arguments, see [9]. Moreover, we can improve the energy estimate to an L^2 estimate because we are stationary and do not need to apply Cauchy-Schwarz to pass from non-stationary to stationary initial conditions. This yields

$$\begin{aligned}&\mathbb{E}\left[\int_0^t((\mathcal{P}_\delta u_s)^2-\|\mathcal{P}_\delta\|_{L^2}^2)(p_{t-s}(x-\cdot)-\mathbb{1})\mathrm{d}s\right]\\&\lesssim\int_0^t\|(-\mathcal{L}_0)^{-1/2}((\mathcal{P}_\delta u)^2-\|\mathcal{P}_\delta\|_{L^2}^2)(p_{t-s}(x-\cdot)-\mathbb{1})\|^2\mathrm{d}s\\&=\int_0^t\left\|(-\mathcal{L}_0)^{-1/2}W_2\left(\int(\delta_y^\delta)^{\otimes 2}(p_{t-s}(x-y)-\mathbb{1})\mathrm{d}y\right)\right\|^2\mathrm{d}s,\end{aligned}$$

where as usualy $\delta_y^\delta=\delta^\delta(y-\cdot)$ for $\delta^\delta=\mathcal{F}^{-1}\mathbb{1}_{[-\delta^{-1},\delta^{-1}]}$. The term

$$W_2\left(\int(\delta_y^\delta)^{\otimes 2}(p_{t-s}(x-y)-\mathbb{1})\mathrm{d}y\right)$$

is basically $\mathcal{G}\varphi_1$ for $\varphi_1=p_{t-s}(x-y)-\mathbb{1}$, except that we are missing a derivative. We could treat this by hand, but since φ_1 has no zero Fourier mode we can cheat and write

$$\varphi_1=\partial_x\partial_x^{-1}(p_{t-s}(x-y)-\mathbb{1}),$$

where

$$\mathcal{F}(\partial_x^{-1}f)(k)=(-2\pi i k)^{-1}\hat{f}(k),\qquad k\neq 0,$$

and ∂_x^{-1} is not defined if $\hat{f}(0)\neq 0$. Then we can apply our estimate for $\mathcal{G}$ and get

$$\begin{aligned}&\int_0^t\left\|(-\mathcal{L}_0)^{-1/2}W_2\left(\int(\delta_y^\delta)^{\otimes 2}(p_{t-s}(x-y)-\mathbb{1})\mathrm{d}y\right)\right\|^2\mathrm{d}s\\&=\int_0^t\|(-\mathcal{L}_0)^{-1/2}\mathcal{G}\partial_x^{-1}(p_{t-s}(x-y)-\mathbb{1})\|^2\mathrm{d}s\\&\lesssim\int_0^t\|(1+\mathcal{N})(-\mathcal{L}_0)^{1/2}\partial_x^{-1}(p_{t-s}(x-y)-\mathbb{1})\|^2\mathrm{d}s\\&\lesssim\int_0^t\|(-\Delta)^{1/2}\partial_x^{-1}(p_{t-s}(x-y)-\mathbb{1})\|_{L^2}^2\mathrm{d}s\\&\simeq\int_0^t\|p_{t-s}(x-y)-\mathbb{1}\|_{L^2}^2\mathrm{d}s\\&\lesssim\int_0^t e^{-C(t-s)}\mathrm{d}s\lesssim 1,\end{aligned}$$

and after dividing by $\sqrt{t}$ this contribution vanishes.

For the martingale part we have a similar estimate:

$$\mathbb{E}[|M_t^{p_{t-\cdot}(x-\cdot)} - M_t^{\mathbb{1}}|^2] = \int_0^t \|p_{t-s}(x-y) - \mathbb{1}\|_{L^2}^2 \mathrm{d}s \lesssim \int_0^t e^{-C(t-s)} \mathrm{d}s \lesssim 1,$$

so again division by $\sqrt{t}$ kills this term for $t \to \infty$. □

Therefore, it suffices to prove the claimed convergence to a Gaussian for $\frac{1}{\sqrt{t}} h_t(\mathbb{1})$. Of course, $\frac{1}{\sqrt{t}} h_0(\mathbb{1})$ vanishes for $t \to \infty$, so we only have to handle the drift and the martingale.

Lemma 5.5. *The stochastic Burgers equation on $\mathbb{T}$ has a spectral gap, i.e. for all $\varphi \in \mathcal{D}(\mathcal{L})$ with $\varphi_0 = 0$ we have*

$$\langle (-\mathcal{L})\varphi, \varphi \rangle \geqslant |2\pi|^2 \|\varphi\|^2,$$

and therefore the semigroup satisfies for all $\varphi \in L^2(\mu)$

$$\|S_t \varphi - \varphi_0\|_{L^2}^2 \leqslant e^{-2|2\pi|^2 t}.$$

Proof. Let $\varphi \in \mathcal{D}(\mathcal{L})$ with $\varphi_0 = 0$. We showed that

$$\langle (-\mathcal{L})\varphi, \varphi \rangle = \|(-\mathcal{L}_0)^{1/2} \varphi\|^2,$$

and we have by Parseval with $\mathbb{Z}_0 = \mathbb{Z} \setminus \{0\}$:

$$\|(-\mathcal{L}_0)^{1/2} \varphi\|^2 = \sum_{n=1}^{\infty} n! \sum_{k \in \mathbb{Z}_0^n} |2\pi k|^2 |\hat{\varphi}(k)|^2 \geqslant |2\pi|^2 \sum_{n=1}^{\infty} n! \sum_{k \in \mathbb{Z}_0^n} |\hat{\varphi}(k)|^2 = |2\pi|^2 \|\varphi\|^2.$$

For the convergence of the semigroup we simply differentiate $\|S_t \varphi\|_{L^2}^2$: Since $S_t \varphi_0 = \varphi_0 = 0$, we get

$$\begin{aligned} \partial_t \|S_t \varphi\|^2 &= 2\langle S_t \varphi, \mathcal{L} S_t \varphi \rangle \\ &= -2\|(-\mathcal{L}_0)^{1/2} S_t \varphi\|^2 \\ &\leqslant -2|2\pi|^2 \|S_t \varphi\|^2, \end{aligned}$$

so the claim follows from Gronwall's inequality. For general φ (not necessarily in $\mathcal{D}(\mathcal{L})$ or $\varphi_0 = 0$) we first replace φ by $\varphi - \varphi_0$ and then we use an approximation argument, because $\mathcal{D}(\mathcal{L})$ is dense. □

Corollary 5.6. *For each $\varphi \in \mathcal{H}_0^{-1}$ with $\varphi_0 = 0$ there exists a unique solution $\psi \in \mathcal{H}_0^1$ to the Poisson equation*

$$-\mathcal{L}\psi = \varphi, \qquad \psi_0 = 0.$$

Proof. As in our construction of $R_1 = (1 - \mathcal{L})^{-1}$, we can also construct $R_\lambda = (\lambda - \mathcal{L})^{-1}$ for each $\lambda > 0$. Moreover,

$$\begin{aligned} \|(-\mathcal{L}_0)^{-1/2} \varphi\| \cdot \|(-\mathcal{L}_0)^{1/2} R_\lambda \varphi\| &\geqslant \langle \varphi, R_\lambda \varphi \rangle \\ &= \langle (\lambda - \mathcal{L}) R_\lambda \varphi, R_\lambda \varphi \rangle \\ &= \lambda \|R_\lambda \varphi\|^2 + \|(-\mathcal{L}_0)^{1/2} R_\lambda \varphi\|^2 \\ &\geqslant \|(-\mathcal{L}_0)^{1/2} R_\lambda \varphi\|^2, \end{aligned}$$

so uniformly in λ:

$$\|(-\mathcal{L}_0)^{1/2} R_\lambda \varphi\|^2 \leqslant \|(-\mathcal{L}_0)^{-1/2} \varphi\|.$$

And since $(R_\lambda \varphi)_0 = \lambda^{-1} \varphi_0 = 0$, we can estimate

$$\|R_\lambda \varphi\|_{\mathcal{H}_0^1} \overset{\text{spectral gap}}{\lesssim} \|(-\mathcal{L}_0)^{1/2} R_\lambda \varphi\|^2 \lesssim \|(-\mathcal{L}_0)^{-1/2} \varphi\| \overset{\varphi_0 = 0}{\lesssim} \|(1 - \mathcal{L}_0)^{-1/2} \varphi\|_{\mathcal{H}_0^{-1}},$$

again uniformly in λ. Now we can use a similar limit procedure as in the proof of Theorem 2.5 to show that $R_\lambda \varphi$ converges subsequentially weakly in $\mathcal{H}_0^1$ to some limit $R_0 \varphi$, and that $-\mathcal{L} R_0 \varphi = \varphi$. Moreover,

$$\langle \mathcal{L}\psi, \psi \rangle = \|(-\mathcal{L}_0)^{1/2} \psi\|^2,$$

so if $\mathcal{L}\psi = 0$ then $\psi = \psi_0$ and this proves the uniqueness of the solution to the Poisson equation. □

We are now ready to prove our main result of this section:

Proof. (Proof of Theorem 5.3) Consider

$$\psi = \psi_2 = \int_{\mathbb{T}} \delta_x^{\otimes 2} \mathbb{1}(x) \mathrm{d}x.$$

Then $\psi \in \mathcal{H}_\infty^{-1}$:

$$\begin{aligned} \hat{\psi}(k_{1:2}) &= \int_{\mathbb{T}^2} e^{-2\pi i k\cdot z} \int_{\mathbb{T}} \delta_x(z_1)\delta_x(z_2)\mathbb{1}(x)\mathrm{d}x\mathrm{d}z \\ &= \int_{\mathbb{T}^2} e^{-2\pi i k\cdot z}\delta(z_1 - z_2)\mathrm{d}z \\ &= \int_{\mathbb{T}} e^{-2\pi i(k_1 z_1 + k_2 z_1)}\mathrm{d}z_1 \\ &= \mathbb{1}_{k_1+k_2=0}, \end{aligned}$$

and therefore for all $\beta \in \mathbb{R}$

$$\|(1+\mathcal{N})^\beta(1-\mathcal{L}_0)^{-1/2}\psi\|^2 = 3^\beta 2! \sum_{k_1,k_2} (1+|2\pi k|^2)^{-1}\mathbb{1}_{k_1+k_2=0} \simeq \sum_k (1+|k|^2)^{-1} < \infty.$$

By the previous result we can thus solve the Poisson equation $-\mathcal{L}\varphi = \psi$, and by the martingale problem and the extension of the I map to $\mathcal{H}_0^{-1}$ we get that

$$N_t := \varphi(u_t) - \varphi(u_0) - I(\mathcal{L}\varphi)_t = \varphi(u_t) - \varphi(u_0) + I(\psi)_t$$

is a martingale, i.e.

$$I(\psi)_t = N_t - \varphi(u_0) + \varphi(u_t),$$

and trivially

$$\frac{1}{\sqrt{t}}(-\varphi(u_0) + \varphi(u_t)) \to 0$$

in L^2. Thus, we have written

$$\frac{1}{\sqrt{t}} h_t(x) = \frac{1}{\sqrt{t}}(M_t^{\mathbb{1}} + N_t) + o(1),$$

where the $o(1)$ term converges strongly to 0 (in L^2, convergence in probability would also be sufficient).

To proceed, we have to compute the quadratic variation of the martingale $M^{\mathbb{1}} + N$. Note that for smooth $\chi(u) = F(u(f_1), \ldots, u(f_m)) \in \mathcal{C}$ we have from Itô's formula, using the equation for h and letting $\tilde{\chi}(u) = F(-u(\partial_x f_1), \ldots, -u(\partial_x f_m))$

$$\begin{aligned} \mathrm{d}\chi(u_t) &= \mathrm{d}\tilde{\chi}(h_t) \\ &= (\ldots)\mathrm{d}t + \sum_{i=1}^m \partial_i F(-h_t(\partial_x f_1), \ldots, -h_t(\partial_x f_m))\mathrm{d}M_t^{-\partial_x f^i} \\ &= (\ldots)\mathrm{d}t + \underbrace{\sum_{i=1}^m \partial_i F(u_t(f_1), \ldots, u_t(f_m))\mathrm{d}M_t^{-\partial_x f^i}}_{=:N_t^\chi}, \end{aligned}$$

and therefore

$$\begin{aligned} \mathrm{d}\langle N^\chi + M^g\rangle_t &= \sum_{i,j=1}^m \partial_i F(u_t(f_1), \ldots, u_t(f_m))\partial_j F(u_t(f_1), \ldots, u_t(f_m)) 2\langle \partial_x f^i, \partial_x f^j\rangle \mathrm{d}t \\ &\quad + 2\sum_{i=1}^m \partial_i F(u_t(f_1), \ldots, u_t(f_m)) 2\langle -\partial_x f^i, g\rangle \mathrm{d}t + 2\|g\|^2 \mathrm{d}t \\ &= \left(2\int (\partial_x D_x \chi(u_t))^2 \mathrm{d}x - 4\int_{\mathbb{T}} \partial_x D_x \chi(u_t) g(x)\mathrm{d}x + 2\|g\|^2\right)\mathrm{d}t \\ &= 2\|\partial_x D_x \chi(u_t) - g\|_{L^2}^2 \mathrm{d}t. \end{aligned}$$

By an approximation argument this identity remains true in our setting, and therefore

$$\mathrm{d}\langle M^{\mathbb{1}} + N\rangle_t = 2\|\partial_x D_x \varphi(u_t) - \mathbb{1}\|_{L^2}^2 \mathrm{d}t.$$

The expectation of this expression simplifies a lot:

$$\mathbb{E}[\langle M^{\mathbb{1}} + N\rangle_t] = 2t(\|(-\mathcal{L}_0)^{1/2}\varphi\|^2 + 1),$$

where we used that the mixed term vanishes because $\int \partial_x D_x \varphi \mathrm{d}x = 0$. By the ergodic theorem we get

$$\left\langle \frac{1}{\sqrt{n}}(M^{\mathbb{1}} + N) \right\rangle_{nt} = \frac{1}{n}\int_0^{nt} 2\|\partial_x D_x \varphi(u_s) - \mathbb{1}\|_{L^2}^2 \mathrm{d}s \longrightarrow 2t(\|(-\mathcal{L}_0)^{1/2}\varphi\|^2 + 1),$$

and therefore $\frac{1}{\sqrt{n}}(M_{n\cdot}^{\mathbb{1}} + N_{n\cdot})$ converges in distribution to σB, where B is a Brownian motion and where

$$\sigma^2 = 2(\|(-\mathcal{L}_0)^{1/2}\varphi\|^2 + 1) \in (0, \infty).$$

Now our claim follows by considering $t = 1$. □

Bibliography

[1] Hajer Bahouri, Jean-Yves Chemin, and Raphael Danchin. *Fourier analysis and nonlinear partial differential equations*. Springer, 2011.

[2] Lorenzo Bertini and Giambattista Giacomin. Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.*, 183(3):571–607, 1997.

[3] Giuseppe Cannizzaro, Massimiliano Gubinelli, and Fabio Toninelli. Gaussian fluctuations for the stochastic Burgers equation in dimension ≥2. *ArXiv preprint arXiv:2304.05730*, 2023.

[4] Patrícia Gonçalves and Milton Jara. Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Arch. Ration. Mech. Anal.*, 212(2):597–644, 2014.

[5] Patrícia Gonçalves and Milton Jara. Density fluctuations for exclusion processes with long jumps. *Probab. Theory Related Fields*, 170(1-2):311–362, 2018.

[6] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics, Pi*, 3(e6), 2015.

[7] Massimiliano Gubinelli and Milton Jara. Regularization by noise and stochastic Burgers equations. *Stochastic Partial Differential Equations: Analysis and Computations*, 1(2):325–350, 2013.

[8] Massimiliano Gubinelli and Nicolas Perkowski. The Hairer–Quastel universality result at stationarity. *RIMS Kôkyûroku Bessatsu*, B59, 2016.

[9] Massimiliano Gubinelli and Nicolas Perkowski. Energy solutions of KPZ are unique. *J. Amer. Math. Soc.*, 31(2):427–471, 2018.

[10] Massimiliano Gubinelli and Nicolas Perkowski. The infinitesimal generator of the stochastic Burgers equation. *Probab. Theory Related Fields*, 178(3-4):1067–1124, 2020.

[11] Martin Hairer. Rough stochastic PDEs. *Comm. Pure Appl. Math.*, 64(11):1547–1585, 2011.

[12] Martin Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[13] Martin Hairer and Hendrik Weber. Rough Burgers-like equations with multiplicative noise. *Probab. Theory Related Fields*, 155(1-2):71–126, 2013.

[14] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang. Dynamic scaling of growing interfaces. *Physical Review Letters*, 56(9):889–892, 1986.

[15] A. Kiselev, F. Nazarov, and A. Volberg. Global well-posedness for the critical 2D dissipative quasi-geostrophic equation. *Invent. Math.*, 167(3):445–453, 2007.

[16] Konstantin Matetski, Jeremy Quastel, and Daniel Remenik. The KPZ fixed point. *Acta Math.*, 227(1):115–203, 2021.

[17] David Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, Second edition, 2006.

[18] Jeremy Quastel and Sourav Sarkar. Convergence of exclusion processes and the KPZ equation to the KPZ fixed point. *J. Amer. Math. Soc.*, 36(1):251–289, 2023.

[19] Xicheng Zhang and Guohuan Zhao. Stochastic Lagrangian path for Leray's solutions of 3D Navier-Stokes equations. *Comm. Math. Phys.*, 381(2):491–525, 2021.