Energy solutions of singular SPDEs

BY NICOLAS PERKOWSKI

August 5, 2024

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 [GJ14] and later Gubinelli-Jara [GJ13] 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.

Most results are based on joint works with Lukas Gräfner.

Table of contents

Introduction

This is a work in progress and in particular there are missing references.

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 an influential physics paper from 1986, Kardar, Parisi and Zhang [KPZ86] 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: $h\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$,

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

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]. 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.

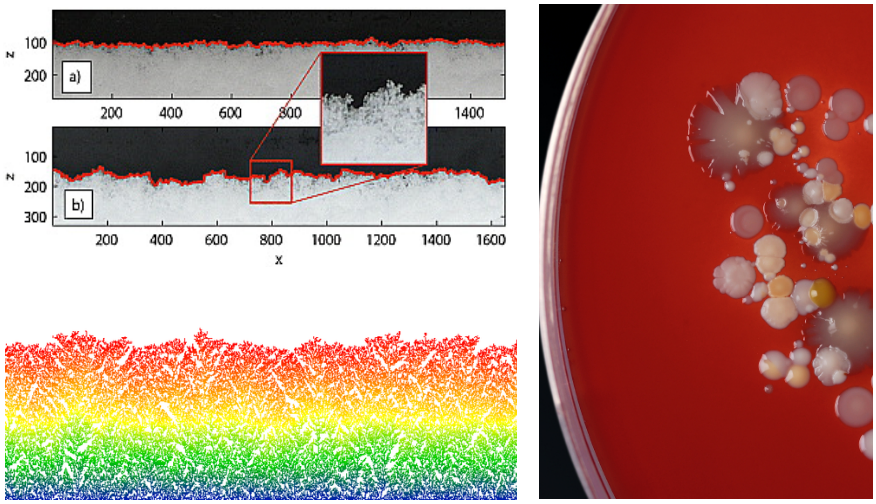

Figure 1. Growing interfaces. Image credit: Löwe et al., Geophys. Res. Letters, Vol. 34, L21507, 2007 (upper left), Nils Berglund (lower left), iStockphoto.com/rudigobbo (right)

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[2] breakthrough in singular SPDEs was a work by Hairer [Hai13] 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 [Hai14] and paracontrolled distributions by Gubinelli-Imkeller-Perkowski [GIP15] extend rough path integration to higher dimensions, which is necessary to treat equations with higher-dimensional space variables. 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 martingale techniques together with functional analytic considerations. Here we present this alternative approach. Let us give some examples of SPDEs where this is applicable:

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.

2. There was a previous work by Hairer [Hai11] 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 there is not as famous as the KPZ equation.

Example 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.

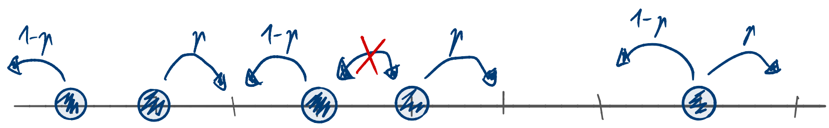

Figure 2. Simple exclusion process

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 $1-p$.

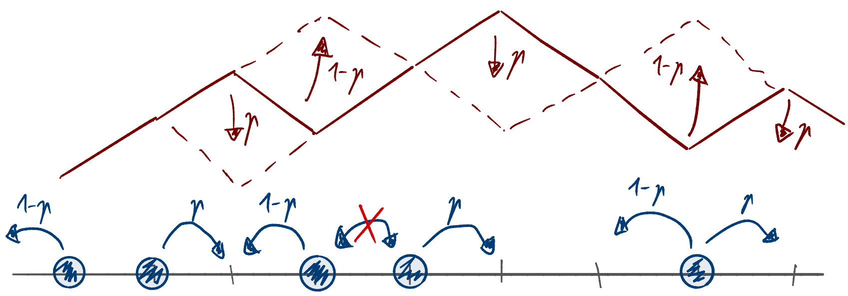

Figure 3. Simple exlusion as interface model

The oscillations of this random interface could for example be a toy model for the interfaces in Figure 1. 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 [MQR21, QS23].

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 [BG97].

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 [GJ18]. 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 [KNV07]:

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

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.

We will see that these examples can be interpreted as infinite-dimensional stochastic differential equations with a drift given by an infinite-dimensional distribution. The techniques developed for solving them also apply to finite-dimensional SDEs, and more precisely we can interpret the equations that are accessible with our approach as infinite-dimensional diffusions with distributional drift that is divergence free with respect to an "accessible" measure, and with additive noise representing a sort of Brownian motion under the invariant measure.

Example 2. Let $b \in \mathcal{S}'(\mathbb{T}^d, \mathbb{R}^d)$ be a divergence-free ($\nabla \cdot b \equiv 0$) vector field. The theory of renormalized solutions to transport equations and ordinary differential equations by Di Perna-Lions [DL89] gives well-posedness results to the SDE

$$\dot{x}(t) = b(x(t))$$

under low regularity assumptions on b, such as $b \in W^{1,1}$ (Sobolev space), based on commutator estimates. Adding a Brownian motion improves the situation and Le Bris-Lions [LL19] obtain weak well-posedness for

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

if $b \in L^2$, again based on commutator estimates. We will replace those commutator estimates with the *energy estimate* of energy solutions (although in the infinite-dimensional case we will also use commutator estimates in a similar way) to obtain weak well-posedness for much more singular b of the form $b = b_1 + b_2$ with $b_1 \in L^2$ and $b_2 \in B^{-1}_{\infty,1}$ (Besov space), which includes $b \in B^{-\gamma}_{\frac{2}{1-\gamma},1}$ for any $\gamma \in [0,1]$, and also for certain $b \in B^{-1}_{2+\varepsilon,2}$ for arbitrarily small $\varepsilon > 0$ under a structural assumption.

There is a closely related recent work by Hao and Zhang [HZ23] who show uniqueness of the limit of approximations to the above SDE for example if $\nabla \cdot b \equiv 0$ and $b \in (L^2 + B^{-1}_{\infty,2}) \cap H^{-1,p}$ for $p > d$, while the conditions we require here do not depend on the dimension.

Example 3. (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 [CHT22, CMOW22, ABK24, YY24] 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

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

Using the scale-invariance of the Gaussian free field and of the Brownian motion to translate the result to the small scale behavior, this suggets that taking the truncation (mollification with ρ) 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) $\xi \in B^{-1-\kappa}_{\infty,\infty}$ 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 the theory of regularity structures that this equation is scaling subcritical exactly if (locally) $b_{\varepsilon} \in B^{-1+\kappa}_{\infty,\infty}$ (then u_{ε} is a perturbation of the heat equation on small scales). This example shows that here 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 the problem slightly, say with a Fourier multiplier $\log(1+|x|^2)^{-\kappa}$ for some $\kappa > 1$? In that case the equation is (slightly) subcritical, but in the scale of Besov spaces the best regularity for b is $B^{-1}_{\infty,\infty}$, i.e. $b \notin \bigcup_{\kappa>0} B^{-1+\kappa}_{\infty,\infty}$. Thus we are unable to solve the equation using regularity structures, because the number of trees we would need to construct is infinite. Our approach will yield weak existence and uniqueness of energy solutions to this equation, leveraging the fact that $b \in B^{-1}_{\infty,1}$, which is (ever so slightly) better than $B^{-1}_{\infty,\infty}$.

We will also consider different integrability scales of Besov spaces, $B^{-\gamma}_{p,q}$ for $p < \infty$, and in that case we get well-posedness results for supercritical b.

Example 4. Once we proved well-posedness, we will study some properties of the SPDEs. Here we focus on weak universality result for the stochastic Burgers equation and a homogenization result for the KPZ equation:

i. Hairer-Quastel [HQ18] type weak universality: We consider the weakly asymmetric interface growth model

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

on $\mathbb{R}_+ \times \varepsilon^{-1}\mathbb{T}$, where ξ is a space-time white noise, F is a suitable nonlinearity, and $\mathcal{P}_{1/2}$ is the projection onto the Fourier modes $|\cdot| \leqslant 1/2$. Let

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

then we show that u_ε converges weakly to the unique energy solution u of

$$\partial_t u = \Delta u + c_2(F) \partial_x u^2 + \sqrt{2(-\Delta)}\,\xi,$$

where

$$c_2(F) = \frac{1}{2!} \mathbb{E}[H_2(X) F(X)],$$

for $X \sim \mathcal{N}(0,1)$ and the second Hermite polynomial $H_2(x) = x^2 - 1$. This is a natural and simple application of energy solutions, because they were introduced in [GJ14] exactly with the purpose of describing scaling limits.

ii. We consider an energy solution h to the the periodic KPZ equation

$$\partial_t h = \Delta h + (\partial_x h)^2 - \infty + \xi$$

and show that there exists $\sigma^2 \in (0,\infty)$ such that for all $x \in \mathbb{T}$

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

weakly, as $t \to \infty$, recovering a result from Gu-Komorowski [GK24] with a short proof that is robust enough to allow for fractional multi-component KPZ equations. This is another simple application of our methods, in combination with standard martingale methods for deriving such fluctuation results [KLO12].

1 Lecture 1: Construction of generator and semigroup

Here we develop general tools that let us construct semigroups for several singular SPDEs, including critical and supercritical ones, as well as for finite-dimensional singular stochastic differential equations. We will spend the rest of the lectures discussing applications, but at the beginning let us stay very general (and abstract) to increase the applicability.

Let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be a separable real Hilbert space (later we will take $\mathcal{H} = L^2(\mu)$, where μ is the candidate invariant measure of some stochastic dynamics). We are given a formal operator $\mathcal{L} = \mathcal{L}_0 + \mathcal{G}$, where $\mathcal{L}_0$ is "nice" (for example the generator of a linear SPDE) and $\mathcal{G}$ is "singular" (for example the generator of a singular nonlinearity), and our goal is to find a domain $\mathcal{D}(\mathcal{L})$ for $\mathcal{L}$ and to show that $(\mathcal{D}(\mathcal{L}), \mathcal{L})$ generates a contraction semigroup on $\mathcal{H}$.

We assume that $\mathcal{H}$ comes with a grading, i.e. it is given by the orthogonal decomposition

$$\mathcal{H} = \overline{\bigoplus_{n=0}^{\infty} \mathcal{H}_n},$$

where each $\mathcal{H}_n$ is a closed linear subspace of $\mathcal{H}$. Recall that the closure is necessary because the direct sum only allows finitely many non-zero entries. In our examples $(\mathcal{H}_n)_{n\in\mathbb{N}_0}$ will correspond to the chaos decomposition under a Gaussian measure, or the grading plays no role and $\mathcal{H}_1 = \mathcal{H}$ and $\mathcal{H}_n = \{0\}$ for $n \neq 1$. The chaos decomposition motivates the following convention of factoring out $\sqrt{n!}$: A general element of $\mathcal{H}$ is represented uniquely as

$$h = \sum_{n=0}^{\infty} \sqrt{n!} h_n, \qquad h_n \in \mathcal{H}_n,$$

and

$$\|h\|^2 = \sum_{n=0}^{\infty} n! \|h_n\|^2.$$

We also write

$$h = (h_n)_{n\in\mathbb{N}_0}.$$

We define the unbounded operator $\mathcal{N}$ which acts on $\mathcal{H}_n$ as multiplication with n. We consider also another (typically unbounded) self-adjoint and positive semi-definite operator

$$(\mathcal{D}(-\mathcal{L}_0), -\mathcal{L}_0),$$

such that $\mathcal{L}_0 \mathcal{H}_n \subset \mathcal{H}_n$ for all $n \in \mathbb{N}_0$. By the spectral theorem we get a functional calculus which allows us to define the powers $(1-\mathcal{L}_0)^{\alpha}$ for $\alpha \in \mathbb{R}$.

Definition 1.1. (Sobolev type spaces) *For $\alpha, \beta \in \mathbb{R}$ we define*

$$\mathcal{H}^{\alpha}_{\beta} := \{h : (1-\mathcal{L}_0)^{\alpha/2}(1+\mathcal{N})^{\beta} h \in \mathcal{H}\},$$

$$\begin{aligned} \|h\|^2_{\mathcal{H}^{\alpha}_{\beta}} &:= \|(1-\mathcal{L}_0)^{\alpha/2}(1+\mathcal{N})^{\beta} h\|^2 \\ &= \sum_{n=0}^{\infty} n!(1+n)^{2\beta} \|(1-\mathcal{L}_0)^{\alpha/2} h_n\|^2, \end{aligned}$$

where for $\alpha < 0$ or $\beta < 0$ we take the completion to obtain a Hilbert space.

The reason for taking $(1-\mathcal{L}_0)^{\alpha/2}$ in the definition of $\mathcal{H}^{\alpha}_{\beta}$ is that we think of $\mathcal{L}_0$ as an anlog of the Laplace operator, counting "two derivatives". In that sense α measures "spatial regularity", while β measures the decay of h_n in $h = (h_n)_{n\in\mathbb{N}_0}$ as $n \to \infty$. For $\alpha, \beta \geqslant 0$ this is a subspace of $\mathcal{H}$. But if $\alpha < 0$ or $\beta < 0$ the space $\mathcal{H}^{\alpha}_{\beta}$ contains "distributions", i.e. an element $h \in \mathcal{H}^{\alpha}_{\beta}$ can only be tested against "nice" $g \in \mathcal{H}$. In particular, if $\mathcal{H} = L^2(\mu)$, then $h \in \mathcal{H}^{\alpha}_{\beta}$ for $\alpha < 0$ or $\beta < 0$ is not a random variable and it cannot be evaluated for (almost all) $\omega \in \Omega$ and instead we can only make sense of the expectation $\mathbb{E}[hg]$ for "nice" random variables $g \in L^2(\mu)$.

The space of test functions is

$$\mathcal{C} := \bigcup_{m\in\mathbb{N}} \mathbb{1}_{\{-\mathcal{L}_0 \leqslant m\}} \mathbb{1}_{\{\mathcal{N} \leqslant m\}} \mathcal{H} \subset \bigcap_{\alpha,\beta\in\mathbb{R}} \mathcal{H}^{\alpha}_{\beta}.$$

We now consider a bounded linear operator $\mathcal{G}: \mathcal{H}^1_0 \to \mathcal{H}^{\alpha}_{\beta}$, we write $\mathcal{G} \in L(\mathcal{H}^1_0, \mathcal{H}^{\alpha}_{\beta})$, for some $\alpha, \beta \in \mathbb{R}$ (in the examples we will have $\alpha, \beta < 0$) such that

$$\langle \mathcal{G}h, g \rangle = -\langle h, \mathcal{G}g \rangle, \qquad g, h \in \mathcal{C}.$$

In particular,

$$\langle \mathcal{G}h, h \rangle = -\langle h, \mathcal{G}h \rangle = 0, \qquad h \in \mathcal{C}.$$

By density of $\mathcal{C}$ the relation $\langle \mathcal{G}h, g \rangle = -\langle h, \mathcal{G}g \rangle$ holds also for $h \in \mathcal{H}^1_0$ and $g \in \mathcal{C}$. Moreover, we get a dual regularity estimate: For all $h, g \in \mathcal{C}$

$$|\langle h, \mathcal{G}g \rangle| = |\langle \mathcal{G}h, g \rangle| \leqslant \|\mathcal{G}\|_{L(\mathcal{H}^1_0, \mathcal{H}^{\alpha}_{\beta})} \|h\|_{\mathcal{H}^1_0} \|g\|_{\mathcal{H}^{-\alpha}_{-\beta}},$$

so that again by density $\mathcal{G}\colon \mathcal{H}^{-\alpha}_{-\beta} \to \mathcal{H}^{-1}_0$.

In our applications we will not have $\mathcal{G}\mathcal{C} \subset \mathcal{H}$, so test functions are not in the domain of $\mathcal{G}$ seen as an operator on $\mathcal{H}$ and it is (very) difficult to identify a domain for $\mathcal{G}$, if it even exists. Therefore, instead of considering $\mathcal{G}$ alone, our goal is to find a domain $\mathcal{D}(\mathcal{L})$ for $\mathcal{L} = \mathcal{L}_0 + \mathcal{G}$, and to construct a semigroup generated by $(\mathcal{D}(\mathcal{L}), \mathcal{L})$.

Proposition 1.2. (Construction of a semigroup) *Under the above assumptions on $\mathcal{G}$ there exists a dense domain $\mathcal{D}(\mathcal{L}) \subset \mathcal{H}^1_0$ for $\mathcal{L} = \mathcal{L}_0 + \mathcal{G}$ such that the closure of $(\mathcal{D}(\mathcal{L}), \mathcal{L})$ generates a contraction semigroup $(T_t)_{t \geqslant 0}$ on $\mathcal{H}$.*

Proof.

i. *Approximate resolvent equation:*[1.1] Let

$$\mathcal{L}^{(m)} = \mathcal{L}_0 + \underbrace{\mathbb{1}_{\mathcal{N} \leqslant m} \mathbb{1}_{-\mathcal{L}_0 \leqslant m} \mathcal{G} \mathbb{1}_{\mathcal{N} \leqslant m} \mathbb{1}_{-\mathcal{L}_0 \leqslant m}}_{\mathcal{G}^{(m)}}.$$

Since $\mathcal{G}$ maps boundedly $\mathcal{H}^1_0 \to \mathcal{H}^\alpha_\beta$, the operator $\mathcal{G}^{(m)} \in L(\mathcal{H}, \mathcal{H})$ is bounded (with operator norm depending on m) and since $\|(\lambda - \mathcal{L}_0)^{-1}\|_{L(\mathcal{H}, \mathcal{H})} \leqslant \lambda^{-1}$ by dissipativity of $\mathcal{L}_0$ (we have $\langle \mathcal{L}_0 h, h \rangle \leqslant 0$ for all $h \in \mathcal{D}(\mathcal{L}_0)$ because $-\mathcal{L}_0$ is positive semi-definite), the map $\Phi\colon \mathcal{H} \to \mathcal{H}$,

$$\Phi(h) = (\lambda - \mathcal{L}_0)^{-1} \mathcal{G}^{(m)} h + (\lambda - \mathcal{L}_0)^{-1} g,$$

is a contraction for sufficiently large λ (depending on m), which means that it has a unique fixed point which solves the resolvent equation

$$(\lambda - \mathcal{L}_0) h = \mathcal{G}^{(m)} h + g \qquad \Leftrightarrow \qquad (\lambda - \mathcal{L}^{(m)}) h = g.$$

Since also $\mathcal{D}(\mathcal{L}^{(m)}) = \mathcal{D}(\mathcal{L}_0)$ by boundedness of $\mathcal{G}^{(m)}$, and since $\mathbb{1}_{\mathcal{N} \leqslant m} \mathbb{1}_{-\mathcal{L}_0 \leqslant m} h \in \mathcal{C}$, we get for all $h \in \mathcal{D}(\mathcal{L}^{(m)})$

$$\langle \mathcal{L}^{(m)} h, h \rangle = \langle \mathcal{L}_0 h, h \rangle + \left\langle \mathcal{G} \underbrace{\mathbb{1}_{\mathcal{N} \leqslant m} \mathbb{1}_{-\mathcal{L}_0 \leqslant m} h}_{\in \mathcal{C}}, \mathbb{1}_{\mathcal{N} \leqslant m} \mathbb{1}_{-\mathcal{L}_0 \leqslant m} h \right\rangle = \langle \mathcal{L}_0 h, h \rangle \leqslant 0,$$

and therefore $\mathcal{L}^{(m)}$ is dissipative. By the Lumer-Phillips theorem (see below) we get that $\mathcal{L}^{(m)}$ generates a contraction semigroup $(T^{(m)}_t)_{t \geqslant 0}$ on $\mathcal{H}$ and thus we can also solve (replacing λ by 1) the resolvent equation $(1 - \mathcal{L}^{(m)}) h^{(m)} = g$ via $h^{(m)} = \int_0^\infty e^{-t} T^{(m)}_t g \mathrm{d} t$. We write $h^{(m)} = (1 - \mathcal{L}^{(m)})^{-1} g$.

ii. *A priori estimate:* Testing the equation $(1 - \mathcal{L}^{(m)}) h^{(m)} = g$ against $h^{(m)} \in \mathcal{D}(\mathcal{L}^{(m)}) = \mathcal{D}(\mathcal{L}_0)$, we get

$$\|h^{(m)}\|^2_{\mathcal{H}^1_0} = \langle (1 - \mathcal{L}_0) h^{(m)}, h^{(m)} \rangle = \langle (1 - \mathcal{L}^{(m)}) h^{(m)}, h^{(m)} \rangle = \langle g, h^{(m)} \rangle \leqslant \|g\|_{\mathcal{H}^{-1}_0} \|h^{(m)}\|_{\mathcal{H}^1_0},$$

and thus

$$\|h^{(m)}\|_{\mathcal{H}^1_0} \leqslant \|g\|_{\mathcal{H}^{-1}_0}$$

and $(h^{(m)})_{m \in \mathbb{N}} \subset \mathcal{H}^1_0$ is uniformly bounded in m.

iii. *Solving the limiting resolvent equation*: We have shown that the operators $((1 - \mathcal{L}^{(m)})^{-1})_{m \in \mathbb{N}}$ are uniformly bounded in $L(\mathcal{H}^{-1}_0, \mathcal{H}^1_0)$. By a general result from functional analysis, this means that there exists a subsequence which converges in the weak operator topology to a limit $\mathcal{R}_1 \in L(\mathcal{H}^{-1}_0, \mathcal{H}^1_0)$, i.e. such that for all $h \in \mathcal{H}^1_0$ and all $g \in \mathcal{H}^{-1}_0$ we have

$$\lim_{k \to \infty} \langle (1 - \mathcal{L}^{(m_k)})^{-1} g, h \rangle_{\mathcal{H}^1_0} = \langle \mathcal{R}_1 g, h \rangle_{\mathcal{H}^1_0} = \langle \mathcal{R}_1 g, (1 - \mathcal{L}_0) h \rangle.$$

1.1. This step could also be achieved with the Lax-Milgram theorem, which would impose less strict assumptions on $\mathcal{L}_0$ and in particular would not require $\mathcal{L}_0$ to be self-adjoint. But without self-adjointness the approximation $\mathcal{G}^{(m)}$ and the description of weak continuity of $\mathcal{G}$ would require more care. Here we use the Lumer-Phillips theorem mainly because we need it later again, and so we avoid introducing two big tools from functional analysis.

Here is a direct argument based on a diagonal sequence: Let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$. By the a priori bound and a diagonal sequence argument we can extract a subsequence, denoted by abuse of notation again by (m), such that for all $k\in\mathbb{N}$ there exists a limit

$$\mathcal{R}_1 e_k := \lim_{m\to\infty} (1-\mathcal{L}^{(m)})^{-1} e_k,$$

weakly in $\mathcal{H}_0^1$. By linearity, this convergence extends to the span of the (e_k) and the resulting map $\sum_k a_k e_k \mapsto \sum_k a_k \mathcal{R}_1 e_k$ 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 uniform a priori bound $\|(1-\mathcal{L}^{(m)})^{-1}h\|_{\mathcal{H}_0^1} \leqslant \|h\|_{\mathcal{H}_0^{-1}}$. We write $\mathcal{R}_1 h = \lim_{m\to\infty}(1-\mathcal{L}^{(m)})^{-1}h$ for all $h\in\mathcal{H}$, again with weak convergence in $\mathcal{H}_0^1$.

Next, we want to show that $(1-\mathcal{L})\mathcal{R}_1 h = h$ for all $h\in\mathcal{H}$. We obtain for all $g\in\mathcal{C}\subset\mathcal{H}_{-\beta}^{-\alpha}\cap\mathcal{H}_0^1$

$$\begin{aligned}\langle h, g\rangle &= \lim_m \langle (1-\mathcal{L}^{(m)})(1-\mathcal{L}^{(m)})^{-1}h, g\rangle \\ &= \lim_m \langle (1-\mathcal{L}^{(m)})^{-1}h, (1-\mathcal{L}_0+\mathcal{G}^{(m)})g\rangle \\ &= \lim_m \langle (1-\mathcal{L}^{(m)})^{-1}h, (1-\mathcal{L}_0+\mathcal{G})g\rangle + \langle (1-\mathcal{L}^{(m)})^{-1}h, (\mathcal{G}^{(m)}-\mathcal{G})g\rangle.\end{aligned}$$

Since $g\in\mathcal{H}_{-\beta}^{-\alpha}\cap\mathcal{H}_0^1$ we have $(1-\mathcal{L}_0+\mathcal{G})g\in\mathcal{H}_0^{-1}$, and since $((1-\mathcal{L}^{(m)})^{-1}h)_m$ converges weakly in $\mathcal{H}_0^1$ to $\mathcal{R}_1 h$, the first term converges to

$$\langle \mathcal{R}_1 h, (1-\mathcal{L}_0+\mathcal{G})g\rangle \overset{g\in\mathcal{C}}{=} \langle (1-\mathcal{L})\mathcal{R}_1 h, g\rangle.$$

The remainder is bounded by

$$\begin{aligned}\langle (1-\mathcal{L}^{(m)})^{-1}h, (\mathcal{G}^{(m)}-\mathcal{G})g\rangle &\leqslant \|(1-\mathcal{L}^{(m)})^{-1}h\|_{\mathcal{H}_0^1}\|(\mathcal{G}^{(m)}-\mathcal{G})g\|_{\mathcal{H}_0^{-1}} \\ &\lesssim \|h\|_{\mathcal{H}_0^{-1}}\|\mathcal{G}^{(m)}-\mathcal{G}\|_{L(\mathcal{H}_{-\beta}^{-\alpha},\mathcal{H}_0^{-1})}\|g\|_{\mathcal{H}_{-\beta}^{-\alpha}},\end{aligned}$$

which converges to 0 because $\mathcal{G}\in L(\mathcal{H}_{-\beta}^{-\alpha},\mathcal{H}_0^{-1})$ and by the dominated convergence theorem applied to the spectral measure representation of $\mathcal{L}_0$ and to the summation in n in the definition of our norms. Therefore,

$$(1-\mathcal{L})\mathcal{R}_1 h = h,$$

and we have constructed a domain

$$\mathcal{D}(\mathcal{L}) = \{\mathcal{R}_1 h : h\in\mathcal{H}\},$$

such that every $\mathcal{R}_1 h\in\mathcal{D}(\mathcal{L})$ satisfies $(1-\mathcal{L})\mathcal{R}_1 h = h\in\mathcal{H}$.

iv. *Existence of the semigroup:* We apply the Lumer-Phillips theorem (see below): We showed that $1-\mathcal{L}$ is surjective, so it remains to show that $\mathcal{L}$ is dissipative. For $g=\mathcal{R}_1 h\in\mathcal{D}(\mathcal{L})$ and $g^{(m)} = (1-\mathcal{L}^{(m)})^{-1}h$ we have (along the subsequence from the previous point)

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

where we used weak lower semi-continuity of the norm and the weak convergence of $g^{(m)}$ to g. Now $h=(1-\mathcal{L}^{(m)})g^{(m)}$ and therefore

$$\langle g^{(m)}, h\rangle = \|g^{(m)}\|_{\mathcal{H}_0^1}^2 \geqslant \|g^{(m)}\|^2,$$

so finally

$$\langle g, \mathcal{L}g\rangle \leqslant \liminf_{n\to\infty} (\|g^{(m)}\|^2 - \|g^{(m)}\|_{\mathcal{H}_0^1}^2) \leqslant 0$$

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

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 $(T_t)_{t\geqslant 0}$ of linear operators $T_t: H \to H$ such that $t \mapsto T_t h \in H$ is (strongly) continuous for all $h \in H$, $T_0 = \mathrm{id}$, and $\|T_t h\| \leqslant \|h\|$ for all $h \in H$.

The semigroup in the Lumer-Phillips theorem is uniquely determined by $(D(A), A)$, and since $\mathcal{G}$ and $\mathcal{L}_0$ are well-defined operators on suitable subspaces of $\mathcal{H}$ that do not depend on our construction in the previous theorem, we might hope that the semigroup is unique. But this is not the case because the domain $\mathcal{D}(\mathcal{L})$ is not necessarily unique and we might obtain another domain by choosing another approximation or another subsequence. There are counterexamples to uniqueness even in the simple case $\mathcal{H}_n \simeq \mathbb{R}$ for all $n \in \mathbb{N}$ and $\mathcal{L}_0 = \mathrm{id}$.

Thus, we need more assumptions to guarantee uniqueness. A simple and classical solution would be to assume that $\mathcal{G} \in L(\mathcal{H}_0^1, \mathcal{H}_0^{-1})$. In the theory of Dirichlet forms, this is known as the *sector condition*, but it is not satisfied in the examples that interest us and we only have $L(\mathcal{H}_0^1, \mathcal{H}_{-1}^{-1})$. To recuperate the loss in $\mathcal{N}$-regularity, we use a commutator estimate.

Lemma 1.3. (Commutator estimate) *Let $\mathcal{G} \in L(\mathcal{H}_1^1, \mathcal{H}_0^{-1})$ be such that there exists $\ell \in \mathbb{N}$ with $\mathcal{G} = \sum_{k=-\ell}^{\ell} \mathcal{G}_k$, where*

$$\mathcal{G}_k: \mathcal{H}_n \to \mathcal{H}_{n+k}, \qquad n \in \mathbb{N}_0,$$

(with $\mathcal{H}_k = \{0\}$ for $k < 0$), and such that each $\mathcal{G}_k \in L(\mathcal{H}_1^1, \mathcal{H}_0^{-1})$. Then the commutator $[\mathcal{N}, \mathcal{G}] := \mathcal{N}\mathcal{G} - \mathcal{G}\mathcal{N}$ has the same bound as $\mathcal{G}$,

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

Proof. We have with $(n+k)^+ = \max\{n+k, 0\}$

$$[\mathcal{N}, \mathcal{G}_k]h = \sum_{n=0}^{\infty} ((n+k)^+ \mathcal{G}_k h_n - \mathcal{G}_k(n h_n)) = \sum_{n=0}^{\infty} k \mathcal{G}_k h_n = k \mathcal{G}_k h.$$ □

Theorem 1.4. (Uniqueness of the semigroup) *Let $\mathcal{G} \in L(\mathcal{H}_1^1, \mathcal{H}_0^{-1})$ and $[\mathcal{N}, \mathcal{G}] \in L(\mathcal{H}_1^1, \mathcal{H}_0^{-1})$ (and we assume always $\langle \mathcal{G}h, g\rangle = -\langle h, \mathcal{G}g\rangle$ for $g, h \in \mathcal{C}$, which as discussed above yields also the dual estimate $\mathcal{G} \in L(\mathcal{H}_0^1, \mathcal{H}_{-1}^{-1})$). Let $\mathcal{D}_{\max}(\mathcal{L}) = \mathcal{L}^{-1}\mathcal{H} = \{h \in \mathcal{H}_0^1: \mathcal{L}h \in \mathcal{H}\}$ be the maximal domain of $\mathcal{L}$. Then:*

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

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

iii. $\mathcal{D}_{\max}(\mathcal{L}) = \mathcal{D}(\mathcal{L})$, where $\mathcal{D}(\mathcal{L})$ is the domain that we constructed in Proposition 1.2;

iv. $\mathcal{D}_{\max}(\mathcal{L})$ is dense and the closed operator $(\mathcal{D}_{\max}(\mathcal{L}), \mathcal{L})$ generates a contraction semigroup $(T_t)_{t\geqslant 0}$ on $\mathcal{H}$.

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

Proof.

i. $\langle (1-\mathcal{L})h, h\rangle = \|h\|_{\mathcal{H}_0^1}^2$ for $h \in \mathcal{D}_{\max}(\mathcal{L})$: Let $h \in \mathcal{H}_0^1$ be such that $(1-\mathcal{L})h \in \mathcal{H}$. For $\lambda > 0$ we introduce the operator

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

The dominated convergence theorem yields

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

where in the third line we used that $(\lambda+\mathcal{N})^{-1}h\in\mathcal{H}_1^1$ and therefore $\langle\mathcal{G}(\lambda+\mathcal{N})^{-1}h,(\lambda+\mathcal{N})^{-1}h\rangle=0$, and in the fourth line we applied once more the dominated convergence theorem. Moreover,

$$
\begin{aligned}
\lim_{\lambda\to\infty}|\lambda^2\langle(\lambda+\mathcal{N})^{-2}(\mathcal{L}^\lambda-\mathcal{L})h,h\rangle| &\leqslant \lim_{\lambda\to\infty}|\langle(\mathcal{L}^\lambda-\mathcal{L})h,h\rangle|\\
&\leqslant \lim_{\lambda\to\infty}\|(\mathcal{L}^\lambda-\mathcal{L})h\|_{\mathcal{H}_0^{-1}}\|h\|_{\mathcal{H}_0^1}\\
&= \lim_{\lambda\to\infty}\|[(\lambda+\mathcal{N}),\mathcal{G}](\lambda+\mathcal{N})^{-1}h\|_{\mathcal{H}_0^{-1}}\|h\|_{\mathcal{H}_0^1}\\
&= \lim_{\lambda\to\infty}\|[\mathcal{N},\mathcal{G}](\lambda+\mathcal{N})^{-1}h\|_{\mathcal{H}_0^{-1}}\|h\|_{\mathcal{H}_0^1}\\
&\lesssim \lim_{\lambda\to\infty}\|(\lambda+\mathcal{N})^{-1}h\|_{\mathcal{H}_1^1}\|h\|_{\mathcal{H}_0^1}\\
&= \lim_{\lambda\to\infty}\|(\lambda+\mathcal{N})^{-1}(1+\mathcal{N})h\|_{\mathcal{H}_0^1}\|h\|_{\mathcal{H}_0^1}\\
&= 0,
\end{aligned}
$$

by yet another application of the dominated convergence theorem. Therefore, $\langle(1-\mathcal{L})h,h\rangle=\|h\|_{\mathcal{H}_0^1}^2$.

ii. $(1-\mathcal{L})$ is injective on $\mathcal{H}_0^1$: Let $h\in\mathcal{H}_0^1$ be such that $(1-\mathcal{L})h=0$. Then $h\in\mathcal{D}_{\max}(\mathcal{L})$ and

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

and thus $h=0$.

iii. $\mathcal{D}_{\max}(\mathcal{L})=\mathcal{D}(\mathcal{L})$: Clearly $\mathcal{D}(\mathcal{L})\subset\mathcal{D}_{\max}(\mathcal{L})$. To see the converse inclusion, recall that we constructed a map $\mathcal{R}_1:\mathcal{H}\to\mathcal{H}_0^1$ such that $(1-\mathcal{L})\mathcal{R}_1h=h$ for all $h\in\mathcal{H}$, and that $\mathcal{D}(\mathcal{L})=\mathcal{R}_1\mathcal{H}$. For $h\in\mathcal{L}^{-1}\mathcal{H}\subset\mathcal{H}_0^1$, let $g=(1-\mathcal{L})h\in\mathcal{H}$. Then $(1-\mathcal{L})\mathcal{R}_1g=g=(1-\mathcal{L})h$. By injectivity of $(1-\mathcal{L})$ on $\mathcal{H}_0^1$ we must have $\mathcal{R}_1g=h$ and therefore $h\in\mathcal{D}(\mathcal{L})$.

iv. $\mathcal{D}_{\max}(\mathcal{L})$ is dense and the closed operator $(\mathcal{D}_{\max}(\mathcal{L}),\mathcal{L})$ generates a contraction semigroup: This follows directly from iii. and Proposition 1.2.

v. $(1-\mathcal{L})\mathcal{C}$ is dense in $\mathcal{H}_0^{-1}$: Let $h\in\mathcal{H}_0^{-1}$ be such that for all $g\in\mathcal{C}$

$$
\begin{aligned}
0 &= \langle(1-\mathcal{L})g,h\rangle_{\mathcal{H}_0^{-1}}\\
&= \langle(1-\mathcal{L}_0)^{-1/2}(1-\mathcal{L})g,(1-\mathcal{L}_0)^{-1/2}h\rangle\\
&= \langle(1-\mathcal{L})g,(1-\mathcal{L}_0)^{-1}h\rangle\\
&= \langle g,(1-\mathcal{L}^*)(1-\mathcal{L}_0)^{-1}h\rangle,
\end{aligned}
$$

where $\mathcal{L}^*=\mathcal{L}_0-\mathcal{G}$ and we used that $g\in\mathcal{C}$ and $(1-\mathcal{L}_0)^{-1}h\in\mathcal{H}_0^1$ to justify the "integration by parts". Since this holds for all $g\in\mathcal{C}$, we must have $(1-\mathcal{L}^*)(1-\mathcal{L}_0)^{-1}h=0$. But since we never used the sign of $\mathcal{G}$ when showing that $1-\mathcal{L}$ is injective, it follows by exactly the same arguments that also $(1-\mathcal{L}^*)$ is injective and therefore $(1-\mathcal{L}_0)^{-1}h=0$, thus $h=0$. □

The difference of this result compared to the previous one is subtle, but the small improvement in *i.* gives the uniqueness of the semigroup, and the density statement in *v.* will also allow us to prove the uniqueness of energy solutions.

If we would assume slightly better estimates for $\mathcal{G}$, say $\|\mathcal{G}h\|_{\mathcal{H}_\lambda^{-1+\kappa}} \lesssim \|h\|_{\mathcal{H}_1^1}$ for $\kappa, \lambda \geqslant 0$ and at least one of them strictly positive, 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 $\beta \geqslant 0$; alternatively, we could assume that $\|\mathcal{G}h\|_{\mathcal{H}_0^{-1}} \leqslant \delta \|h\|_{\mathcal{H}_1^1}$ for some sufficiently small $\delta > 0$ to obtain the density of $\mathcal{H}_\beta^1 \cap \mathcal{D}(\mathcal{L})$. 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.

Remark 1.5. Note that steps iii.-v. all follow from the injectivity shown in step ii. (and the injectivity of $1 - \mathcal{L}^*$). Thus, the same conclusions hold if we can show the injectivity in some other way.

2 Lecture 2: Energy solutions to the stochastic Burgers equation

Here we will set up function spaces for singular SPDEs on which we can implement the construction of the semigroup from the first lecture, and we will give construct weak solutions to those singular SPDEs and also introduce the crucial energy estimate, which yields their uniqueness. We focus on the concrete example of the stochastic Burgers equation on the one-dimensional torus, where the function spaces will be centered around the white noise and its chaos representation isometry to the Fock space. At the end we give an abstract formulation which encompasses other singular SPDEs with bilinear nonlinearity and Gaussian invariant measure.

2.1 Why white noise?

To analyze the 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[2.1]. 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 consider first a finite-dimensional example:

Example 2.1. Let Σ be an invertible covariance matrix on $\mathbb{R}^d$ and consider an ODE with smooth vector field f (say of polynomial growth),

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

such that

$$\nabla \cdot \left(f(x) e^{-\frac{1}{2}\langle x, \Sigma^{-1} x\rangle} \right) = 0, \qquad x \in \mathbb{R}^d.$$

Let $\mu = \mathcal{N}(0, \Sigma)$, then we have for all sufficiently nice test functions φ (say C^1 with the partial derivatives of polynomial growth)

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

where $Z > 0$ 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,

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

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

The second condition is equivalent to the energy $E(x) := \frac{1}{2}\langle x, \Sigma^{-1}x\rangle$ 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 2.2. 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 $E(u) = \frac{1}{2}\int u^2 = \langle u, u\rangle_{L^2(\mathbb{T})}$ 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 variational 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

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

is formally preserved. How should we interpret this measure? On $\mathbb{R}^d$ a random variable $X \sim \frac{1}{Z} e^{-\langle x, x\rangle_{\mathbb{R}^d}}\mathrm{d}x$ is $\mathbb{R}^d$-valued, 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 $\eta \sim \frac{1}{Z} e^{-\frac{1}{2}\langle u, u\rangle_{L^2(\mathbb{T})}}\mathrm{d}u$ is an $L^2(\mathbb{T})$-valued, centered Gaussian with

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

This is nearly correct, except that η is almost surely not $L^2(\mathbb{T})$-valued but instead it takes values in a larger space of distributions (for example in $H^{-1/2-\kappa}(\mathbb{T})$ for any $\kappa > 0$). 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 2.3. *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 known 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[2.2]

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

2.2. 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 2.4. Let A be a symmetric, negative definite Fourier multiplier, i.e. $\mathcal{F}(Au)(k) = a(k)\mathcal{F}u(k)$ for some symmetric function $a: \mathbb{Z} \to [0, \infty)$. Then the equation

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

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

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

where $(\hat{\xi}(k))_{k \in \mathbb{Z}}$ are independent complex-valued white noises, up to the constraint $\hat{\xi}_t(-k) = \overline{\hat{\xi}_t(k)}$. 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 exactly the distribution of the white noise.

For $a(k) = (2\pi k)^2$ we get $A = -\Delta$.

For the σ-algebra $\mathcal{F} = \sigma(\eta(\varphi): \varphi \in L^2(\mathbb{T}))$ on Ω, we have

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

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. We consider the operators

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

Recall that

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

with

$$\begin{aligned} \|\varphi\|_{\mathcal{H}_\beta^\alpha}^2 &= \|(1 - \mathcal{L}_0)^{\alpha/2}(1 + \mathcal{N})^\beta \varphi\|^2 \\ &= \sum_{n=0}^\infty n!(1+n)^{2\beta} \|(1 - \Delta)^{\alpha/2}\varphi_n\|_{L_s^2(\mathbb{T}^n)}^2 \\ &=: \sum_{n=0}^\infty n!(1+n)^{2\beta} \|\varphi_n\|_{H_s^\alpha(\mathbb{T}^n)}^2, \end{aligned}$$

and that we defined the *cylinder functions* as

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

i.e. they are polynomials depending on finitely many Fourier modes.

2.2 Fock space representation of the generator, bounds

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

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

which we interpret in the weak sense: $v \in C(\mathbb{R}_+, \mathcal{S}')$ such that for all $f \in C^2(\mathbb{T})$

$$v_t(f) = v_0(f) + \int_0^t v_s(\Delta f)\mathrm{d}s + M_t^f,$$

where M^f is a continuous martingale with quadratic variation

$$\langle M^f \rangle_t = 2\|(-\Delta)^{1/2} f\|_{L^2(\mathbb{T})}^2 t.$$

We consider a stationary initial condition for now, i.e. v_0 is a white noise.

Lemma 2.5. (Fock space representation of the OU generator) *For* $\varphi \in \mathcal{H}_0^2 \subset \mathcal{D}(\mathcal{L}_0)$ *we have*

$$(\mathcal{L}_0 \varphi)_n = \Delta \varphi_n := (\partial_{x_1}^2 + \cdots + \partial_{x_n}^2)\varphi.$$

Proof. It suffices to prove this identity for $W_n(\varphi_n)$ with $\varphi_n \in H_s^2(\mathbb{T}^n)$. 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

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

where we used that $H_n' = n H_{n-1}$. Now we use the multiplication rule for the Wiener chaos, see [Nua06], Proposition 1.1.2, and rewrite the first term on the right hand side as

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

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

$$\mathcal{G} = \mathcal{G}_+ + \mathcal{G}_-, \qquad \mathcal{G}_\pm : L_s^2(\mathbb{T}^n) \to L_s^2(\mathbb{T}^{n\pm 1}),$$

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 $x_{i:i+k} = (x_i, x_{i+1}, ..., x_{i+k})$, *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

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

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

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

The general multiplication rule for Wiener chaos variables from [Nua06], 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

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

Now observe that

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

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 2.7. You may feel uneasy at this point about the formal manipulations in the last "proof". To make them rigorous, 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 by considering

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

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

Lemma 2.8. (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 $\varphi\in\mathcal{H}^1_\beta$ and $\psi\in\mathcal{H}^1_{\beta'}$ with $\beta+\beta'\geqslant 1$ 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 [GP20] for details, but it can be a good exercise to do the computation yourself.[2.3]

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. □

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 $\varphi\in\mathcal{H}^{1/2+\varepsilon}_\beta$ for some $\varepsilon>0$, and no matter how smooth φ is, $\mathcal{G}_+\varphi$ is never better than $\mathcal{H}^{-1/2-\varepsilon}_\infty$.

2.3 Construction of the semigroup

Theorem 2.9. *In the setting of the stochastic Burgers equation, there exists a unique semigroup $(T_t)_{t\geqslant 0}$ generated by $\mathcal{L}=\mathcal{L}_0+\mathcal{G}$ with domain $\mathcal{D}_{\max}(\mathcal{L})=\{\varphi\in\mathcal{H}^1_0:\mathcal{L}\varphi\in\mathcal{H}=\mathcal{H}^0_0\}$. Moreover, $(1-\mathcal{L})\mathcal{C}$ is dense in $\mathcal{H}^{-1}_0$.*

Proof. This follows directly from Theorem 1.4 because $\mathcal{G}\in L(\mathcal{H}^1_\beta,\mathcal{H}^{-1}_{\beta-1})$ for any $\beta\in\mathbb{R}$ and $\mathcal{G}=\mathcal{G}_++\mathcal{G}_-$ with $\mathcal{G}_\pm:\mathcal{H}_n\to\mathcal{H}_{n\pm1}$ and $\mathcal{G}_\pm\in L(\mathcal{H}^1_\beta,\mathcal{H}^{-1}_{\beta-1})$, so that we get the commutator estimate $[\mathcal{N},\mathcal{G}]\in L(\mathcal{H}^1_\beta,\mathcal{H}^{-1}_{\beta-1})$. □

2.4 Construction of energy solutions

We have now two operators $\mathcal{L}_0$ and $\mathcal{G}$, such that $\mathcal{H}^\alpha_0=(1-\mathcal{L}_0)^{\alpha/2}\Gamma L^2$, 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 $\mathcal{H}^1_\beta\to\mathcal{H}^{-1}_{\beta-1}$. To construct energy solutions we consider the approximation

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

2.3. 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 [CGT23] it is shown that the estimate even holds with $k=1/2$.

where $\mathcal{P}_\varepsilon u = \mathcal{F}^{-1}(\mathbb{1}_{[-\varepsilon^{-1},\varepsilon^{-1}]}\mathcal{F}u)$. 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 $\mathcal{L}_0 + \mathcal{G}^\varepsilon$, where $\mathcal{G}^\varepsilon$ has the same bounds and antisymmetry properties as $\mathcal{G}$, uniformly in ε.

Lemma 2.10. (Itô trick) [2.4] *Assume that* law $u_0^\varepsilon = \nu$ *and that with the white noise law* μ *we have* $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \in L^2(\mu)$ *and let* $p \geqslant 2$. *Then we have uniformly in* ε

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

where

$$\mathcal{E}(\varphi) = 2\int_{\mathbb{T}} \left|(-\Delta)_x^{1/2} D_x\varphi\right|^2 \mathrm{d}x,$$

for

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

Proof. Let $u_0^\varepsilon \sim \mu$ and let $\hat{u}_t^\varepsilon = u_{T-t}^\varepsilon$. Since u^ε is a stationary Markov process, also the time-reversed process $\hat{u}^\varepsilon$ is Markov, with generator $(\mathcal{L}_0 + \mathcal{G}^\varepsilon)^* = \mathcal{L}_0 - \mathcal{G}^\varepsilon$. Thus, we get from the martingale problem

$$\begin{aligned}
\varphi(u_t^\varepsilon) &= \varphi(u_0^\varepsilon) + \int_0^t (\mathcal{L}_0 + \mathcal{G}^\varepsilon)\varphi(u_s^\varepsilon)\mathrm{d}s + M_t^{\varphi,\varepsilon},\\
\varphi(\hat{u}_T^\varepsilon) &= f(\hat{u}_{T-t}^\varepsilon) + \int_{T-t}^T (\mathcal{L}_0 - \mathcal{G}^\varepsilon)\varphi(\hat{u}_s^\varepsilon)\mathrm{d}s + \hat{M}_T^{\varphi,\varepsilon} - \hat{M}_{T-t}^{\varphi,\varepsilon},
\end{aligned}$$

where $M^{\varphi,\varepsilon}$ is a martingale and $\hat{M}^{\varphi,\varepsilon}$ is a martingale in the backward filtration (the filtration generated by $\hat{u}^\varepsilon$). We can rewrite the equation for $\varphi(\hat{u}_T^\varepsilon)$ as

$$\begin{aligned}
\varphi(u_0^\varepsilon) &= \varphi(u_t^\varepsilon) + \int_{T-t}^T (\mathcal{L}_0 - \mathcal{G}^\varepsilon)\varphi(u_{T-s}^\varepsilon)\mathrm{d}s + \hat{M}_T^{\varphi,\varepsilon} - \hat{M}_{T-t}^{\varphi,\varepsilon}\\
&= \varphi(u_t^\varepsilon) + \int_0^t (\mathcal{L}_0 - \mathcal{G}^\varepsilon)\varphi(u_s^\varepsilon)\mathrm{d}s + \hat{M}_T^{\varphi,\varepsilon} - \hat{M}_{T-t}^{\varphi,\varepsilon}.
\end{aligned}$$

Adding this to the equation for $\varphi(u_t^\varepsilon)$ from above, we get

$$\int_0^t 2\mathcal{L}_0\varphi(u_s^\varepsilon)\mathrm{d}s = M_t^{\varphi,\varepsilon} + \hat{M}_T^{\varphi,\varepsilon} - \hat{M}_{T-t}^{\varphi,\varepsilon}.$$

Moreover, for cylinder functions $\varphi \in \mathcal{C}$ we have

$$M_t^{\varphi,\varepsilon} = \int_{[0,t]\times\mathbb{T}} \sqrt{2}(-\Delta)_x^{1/2} D_x\varphi(u_s^\varepsilon)\xi(\mathrm{d}s,\mathrm{d}x)$$

by Itô's formula, and thus

$$\langle M^{\varphi,\varepsilon}\rangle_t = \int_0^t \mathcal{E}(\varphi)(u_s^\varepsilon)\mathrm{d}s, \qquad \langle \hat{M}^{\varphi,\varepsilon}\rangle_t = \int_0^t \mathcal{E}(\varphi)(\hat{u}_s^\varepsilon)\mathrm{d}s,$$

2.4. It is possible to extend this to initial densities in $L^1(\mu)$, see [GP24], but then the formulation of the energy estimate involves probabilities rather than expectations and it is slightly more involved. For simplicity, we restrict to L^2-densities.

which by approximation extends to sufficiently regular φ. Thus, the Burkholder-Davis Gundy inequality, Minkowski's inequality and stationarity yield

$$\begin{aligned}\mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t \mathcal{L}_0\varphi(u_s^\varepsilon)\mathrm{d}s\right|^p\right] &\lesssim \mathbb{E}\left[\left(\int_0^T \mathcal{E}\varphi(u_s^\varepsilon)\mathrm{d}s\right)^{p/2}\right]\\ &\lesssim \left(\int_0^T \mathbb{E}[|\mathcal{E}\varphi(u_s^\varepsilon)|^{p/2}]^{2/p}\mathrm{d}s\right)^{p/2}\\ &= T^{\frac{p}{2}}\|\mathcal{E}\varphi\|_{L^{p/2}}^{p/2}.\end{aligned}$$

If $u_0^\varepsilon \sim \nu$, then we apply the Cauchy-Schwarz inequality to obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t \mathcal{L}_0\varphi(u_s^\varepsilon)\mathrm{d}s\right|^p\right] &= \int \mathbb{E}_u\left[\sup_{t\leqslant T}\left|\int_0^t \mathcal{L}_0\varphi(u_s^\varepsilon)\mathrm{d}s\right|^p\right]\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(u)\mu(\mathrm{d}u)\\ &\leqslant \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}\left(\int \mathbb{E}_u\left[\sup_{t\leqslant T}\left|\int_0^t \mathcal{L}_0\varphi(u_s^\varepsilon)\mathrm{d}s\right|^p\right]^2\mu(\mathrm{d}u)\right)^{1/2}\\ &\leqslant \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}\left(\int \mathbb{E}_u\left[\sup_{t\leqslant T}\left|\int_0^t \mathcal{L}_0\varphi(u_s^\varepsilon)\mathrm{d}s\right|^{2p}\right]\mu(\mathrm{d}u)\right)^{1/2}\\ &= \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}\mathbb{E}_\mu\left[\sup_{t\leqslant T}\left|\int_0^t \mathcal{L}_0\varphi(u_s^\varepsilon)\mathrm{d}s\right|^{2p}\right]^{1/2}\\ &\lesssim \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)} T^{\frac{p}{2}}\|\mathcal{E}\varphi\|_{L^p}^{p/2}.\end{aligned}$$

□

The $L^p(\mu)$ norm of $\mathcal{E}(\varphi)$ may be difficult to compute. But for $p=1$ the $L^1(\mu)$ norm is simply

$$\int\left(2\int_{\mathbb{T}}|(-\Delta)_x^{1/2}D_x\varphi|^2\mathrm{d}x\right)\mathrm{d}\mu = 2\|(-\mathcal{L}_0)^{1/2}\varphi\|_{L^2(\mu)}^2,$$

which follows from a direct computation, or by Dynkin's formula for the quadratic variation of $\varphi(u^\varepsilon)$,

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

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

Corollary 2.11. (Energy estimate) *Assume that* law $u_0^\varepsilon = \nu$ *and that with the white noise law* μ *we have* $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \in L^2(\mu)$ *and let* $p \geqslant 2$. *Then we have uniformly in* ε *and for all* $0 \leqslant s \leqslant t$

$$\mathbb{E}\left[\sup_{r\in[s,t]}\left|\int_s^r \varphi(u_\tau^\varepsilon)\mathrm{d}\tau\right|\right] \lesssim \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}(|t-s|+|t-s|^{1/2})\|\varphi\|_{\mathcal{H}_0^{-1}}.$$

If $\varphi \in \bigoplus_{n=0}^N \mathcal{H}_n$ *for some* $n \in \mathbb{N}$, *then we even have for all* $p \geqslant 1$

$$\mathbb{E}\left[\sup_{r\in[s,t]}\left|\int_s^r \varphi(u_\tau^\varepsilon)\mathrm{d}\tau\right|^p\right] \lesssim \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}(|t-s|^p+|t-s|^{p/2})\|\varphi\|_{\mathcal{H}_0^{-1}}^p.$$

Proof. Let $\psi = (1-\mathcal{L}_0)^{-1}\varphi$, so that $\varphi = \psi - \mathcal{L}_0\psi$ and thus by triangle inequality and Itô trick

$$\begin{aligned}\mathbb{E}\left[\sup_{r\in[s,t]}\left|\int_s^r \varphi(u_s^\varepsilon)\mathrm{d}s\right|\right] &\leqslant \mathbb{E}\left[\sup_{r\in[s,t]}\left|\int_s^r \psi(u_s^\varepsilon)\mathrm{d}s\right|\right] + \mathbb{E}\left[\sup_{r\in[s,t]}\left|\int_s^r \mathcal{L}_0\psi(u_s^\varepsilon)\mathrm{d}s\right|\right]\\ &\lesssim |t-s|\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}\|\psi\|_{L^2(\mu)} + \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}|t-s|^{1/2}\|\mathcal{E}(\psi)\|_{L^1(\mu)}^{1/2}\\ &= |t-s|\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}\|\psi\|_{L^2(\mu)} + \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}|t-s|^{1/2}\|(-\mathcal{L}_0)^{1/2}\psi\|_{L^2(\mu)},\end{aligned}$$

where we used that by stationarity under μ the Itô trick also holds on the interval $[s,t]$ and not only on $[0,T]$. Now it suffices to note that

$$\|\psi\|_{L^2(\mu)} = \|(1-\mathcal{L}_0)^{-1}\varphi\|_{L^2(\mu)} \leqslant \|(1-\mathcal{L}_0)^{-1/2}\varphi\|_{L^2(\mu)} = \|\varphi\|_{\mathcal{H}_0^{-1}}$$

and

$$\|(-\mathcal{L}_0)^{1/2}\psi\|_{L^2(\mu)} = \|(-\mathcal{L}_0)^{1/2}(1-\mathcal{L}_0)^{-1}\varphi\|_{L^2(\mu)} \leqslant \|(1-\mathcal{L}_0)^{-1/2}\varphi\|_{L^2(\mu)} = \|\varphi\|_{\mathcal{H}_0^{-1}}$$

to conclude the proof of the first claim.

If $\varphi \in \bigoplus_{n=0}^{N} \mathcal{H}_n$ we have also $\psi \in \bigoplus_{n=0}^{N} \mathcal{H}_n$ and $\mathcal{E}(\psi) \in \bigoplus_{n=0}^{2N-2} \mathcal{H}_n$ and thus we can apply Gaussian hypercontractivity to bound

$$\begin{aligned}\mathbb{E}\left[\sup_{r\in[s,t]}\left|\int_s^r \varphi(u_s^{\varepsilon})\mathrm{d}s\right|^p\right] &\lesssim |t-s|^p\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}\|\psi\|_{L^p(\mu)}^p + \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}|t-s|^{p/2}\|\mathcal{E}(\psi)\|_{L^p(\mu)}^{p/2}\\ &\lesssim |t-s|^p\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}\|\psi\|_{L^2(\mu)}^p + \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}|t-s|^{p/2}\|\mathcal{E}(\psi)\|_{L^1(\mu)}^{p/2},\end{aligned}$$

and then the second claim follows as above. □

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 $C(\mathbb{R}_+, \mathbb{R})$.

Theorem 2.12. (Existence of energy solutions, Gonçalves-Jara [GJ14], Gubinelli-Jara [GJ13]) *Assume that* law $u_0^{\varepsilon} = \nu$ *with* $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \in L^2(\mu)$*. Then* $(u^{\varepsilon})_{\varepsilon>0}$ *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 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, where $\mathcal{L}^{\delta} = \mathcal{L}_0 + \mathcal{G}^{\delta}$.

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

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

iii. *u* satisfies an energy estimate*:*

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

In particular, for each $T>0$ there is a unique continuous extension of the map

$$I: \mathcal{C} \to L^1(\Omega, C([0,T])), \qquad I(\varphi)_t = \int_0^t \varphi(X_s)\mathrm{d}s,$$

to $\mathcal{H}_0^{-1}$ (and the extensions for different T are consistent), and we denote the extension with the same symbol I.

We call such u an energy solution *of the stochastic Burgers equation.*

Proof.

- *Tightness:* Let $f \in C^{\infty}(\mathbb{T})$. Then $\varphi(u) = u(f)$ satisfies $\varphi_1 = f$ and $\varphi_n = 0$ otherwise and thus $\varphi \in \mathcal{H}_1$ and we can apply the L^p-version of the energy estimate:

$$\mathbb{E}\left[\left|\int_s^t u_r^{\varepsilon}(f)\mathrm{d}r\right|^p\right] \lesssim \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^2(\mu)}(|t-s|^p + |t-s|^{p/2})\|\varphi\|_{\mathcal{H}_0^{-1}}^p.$$

Since $u_0^{\varepsilon}(f) \sim \mathcal{N}(0, \|f\|_{L^2}^2)$ is tight by assumption, we get the uniform tightness of $(u^{\varepsilon}(f))_{\varepsilon>0}$ by Kolmogorov's continuity criterion.

- *Any limit point is incompressible and satisfies an energy estimate (ii. and iii.):* This follows from Fatou's lemma for weak convergence. Note that u^ε is incompressible uniformly in ε:

$$\mathbb{E}[|\varphi(u_t^\varepsilon)|] \leqslant \left\| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right\|_{L^2(\mu)} \mathbb{E}_\mu[|\varphi(u_t^\varepsilon)|^2]^{1/2} = \left\| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right\|_{L^2(\mu)} \|\varphi\|_{L^2(\mu)}.$$

- *Any limit point solves the stochastic Burgers equation:* Let $\varphi \in \mathcal{C} \subset \mathcal{H}_\infty^\infty := \bigcap_m \mathcal{H}_m^m$. Then $\mathcal{G}\varphi \in \mathcal{H}_0^{-1}$ and by the dominated convergence theorem

$$\int_0^\cdot \mathcal{G}^\delta \varphi(u_s^\varepsilon) \mathrm{d}s \to I(\mathcal{G}\varphi)$$

in $L^1(\Omega, C([0,T]))$. Let now $0 \leqslant s < t \leqslant T$ and let $G: C([0,s], \mathcal{S}') \to \mathbb{R}$ be continuous and bounded. By a monotone class argument, it suffices to show that

$$\mathbb{E}\left[\left(\varphi(u_t) - \varphi(u_s) - \int_s^t \mathcal{L}_0 \varphi(u_r) \mathrm{d}r - (I(\mathcal{G}\varphi)_t - I(\mathcal{G}\varphi)_s)\right) G((u_r)_{r \in [0,s]})\right] = 0.$$

The energy estimate yields

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

which converges to zero by the dominated convergence theorem. Moreover, since $\mathcal{G}^\delta \varphi(u)$ depends continuously[2.5] on $u \in \mathcal{S}'$ and is of polynomial growth, and since all moments of $u(f)$ are bounded by Gaussianity under the stationary measure, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\left(\varphi(u_t) - \varphi(u_s) - \int_s^t \mathcal{L}_0 \varphi(u_r) \mathrm{d}r - \int_s^t \mathcal{G}^\delta \varphi(u_r) \mathrm{d}r\right) G((u_r)_{r \in [0,s]})\right] \\
&= \lim_{\varepsilon \to 0} \mathbb{E}\left[\left(\varphi(u_t^\varepsilon) - \varphi(u_s^\varepsilon) - \int_s^t \mathcal{L}_0 \varphi(u_r^\varepsilon) \mathrm{d}r - \int_s^t \mathcal{G}^\delta \varphi(u_r^\varepsilon) \mathrm{d}r\right) G((u_r^\varepsilon)_{r \in [0,s]})\right] \\
&\leqslant \lim_{\varepsilon \to 0} \mathbb{E}\left[\left(\varphi(u_t^\varepsilon) - \varphi(u_s^\varepsilon) - \int_s^t \mathcal{L}_0 \varphi(u_r^\varepsilon) \mathrm{d}r - \int_s^t \mathcal{G}^\varepsilon \varphi(u_r^\varepsilon) \mathrm{d}r\right) G((u_r^\varepsilon)_{r \in [0,s]})\right] \\
&\quad + \lim_{\varepsilon \to 0} \mathbb{E}\left[\left(\int_s^t (\mathcal{G}^\delta \varphi - \mathcal{G}^\varepsilon \varphi)(u_r^\varepsilon) \mathrm{d}r\right) G((u_r^\varepsilon)_{r \in [0,s]})\right] \\
&\lesssim 0 + \lim_{\varepsilon \to 0} T^{1/2} \|\mathcal{G}^\delta \varphi - \mathcal{G}^\varepsilon \varphi\|_{\mathcal{H}_0^{-1}} \|G\|_\infty \\
&\simeq \|\mathcal{G}^\delta \varphi - \mathcal{G}\varphi\|_{\mathcal{H}_0^{-1}}.
\end{aligned}$$

Overall, we obtain

$$\left| \mathbb{E}\left[\left(\varphi(u_t) - \varphi(u_s) - \int_s^t \mathcal{L}_0 \varphi(u_r) \mathrm{d}r - (I(\mathcal{G}\varphi)_t - I(\mathcal{G}\varphi)_s)\right) G((u_r)_{r \in [0,s]})\right] \right| \lesssim \|\mathcal{G}^\delta \varphi - \mathcal{G}\varphi\|_{\mathcal{H}_0^{-1}},$$

for any $\delta > 0$. The claim follows by letting $\delta \to 0$. □

Remark 2.13.

i. We could replace *i.* by the following condition *i'.*: For all $f \in C^\infty(\mathbb{T})$ the process

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

is a continuous martingale with quadratic variation

$$\langle M^f \rangle_t = 2t \|(-\Delta)^{1/2} f\|_{L^2}^2.$$

Then *i.,ii.,iii.* are equivalent to *i'., ii., iii.* This is sometimes easier to verify because *i'.* allows us to work only with linear test functions instead of cylinder functions.

2.5. Recall that we are on the torus and hence the Fourier multiplier $\mathcal{P}_\delta$ is continuous from $\mathcal{S}'$ to $\mathcal{S}$. On $\mathbb{R}$ we would need a small additional argument.

ii. The same construction works whenever $\mathcal{G}\mathcal{C} \subset \mathcal{H}_0^{-1}$, i.e. without assuming $\mathcal{G}\mathcal{H}_\beta^1 \subset \mathcal{H}_{\beta-1}^{-1}$. This last condition is only needed to apply Theorem 1.4 which gives the density of $(1-\mathcal{L})\mathcal{C}$ in $\mathcal{H}_0^{-1}$, which we will need later to prove that energy solutions are solutions to the martingale problem for $(\mathcal{D}(\mathcal{L}), \mathcal{L})$ and thus to show their weak uniqueness.

2.5 Uniqueness of energy solutions via duality with the 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 2.14. *Let u be an energy solution of the stochastic Burgers equation with $u_0 \sim \nu$ and $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \in L^2(\mu)$, and let $(\mathcal{D}(\mathcal{L}), \mathcal{L})$ be the operator constructed in Theorem 2.9. 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 $(\varphi^{(m)})_m \subset \mathcal{C}$ such that $((1-\mathcal{L})\varphi^{(m)})_m$ converges to $(1-\mathcal{L})\varphi$ in $\mathcal{H}_0^{-1}$. Since $\mathcal{R}_1$ (from the proof of Theorem 1.4) is a bounded operator from $\mathcal{H}_0^{-1}$ to $\mathcal{H}_0^1$, we get that $(\mathcal{R}_1(1-\mathcal{L})\varphi^{(m)})_m$ converges to $\mathcal{R}_1(1-\mathcal{L})\varphi$ in $\mathcal{H}_0^1$. But in the proof of Theorem 1.4 we showed that $\mathcal{R}_1(1-\mathcal{L})\psi = \psi$ for all $\psi \in \mathcal{H}_0^0$ and therefore $(\varphi^{(m)})_m$ converges to φ in $\mathcal{H}_0^1$, and then also $(\mathcal{L}\varphi^{(m)})_m$ converges to $\mathcal{L}\varphi$ in $\mathcal{H}_0^{-1}$. Since $\varphi^{(m)} \in \mathcal{C}$, the martingale problem for cylinder functions yields that

$$\varphi^{(m)}(u_t) - \varphi^{(m)}(u_0) - I(\mathcal{L}\varphi^{(m)})_t, \qquad t \geqslant 0,$$

is a martingale. By the energy estimate together with the convergence $\mathcal{L}\varphi^{(m)} \to \mathcal{L}\varphi$ in $\mathcal{H}_0^{-1}$, we get $I(\mathcal{L}\varphi^{(m)})_t \to I(\mathcal{L}\varphi)_t$ in $L^1(\Omega)$. By incompressibility $\varphi^{(m)}(u_t) - \varphi^{(m)}(u_0) \to \varphi(u_t) - \varphi(u_0)$ in L^1, and since an L^1-limit of martingales is again a martingale, the proof is complete. □

Lemma 2.15. *If u is incompressible and it solves the martingale problem for $\mathcal{L}$, then for all $f \in C(\mathbb{R}_+, \mathcal{D}(\mathcal{L})) \cap C^1(\mathbb{R}_+, L^2(\mu))$ the process*

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

is a martingale.

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

$$\begin{aligned}\varphi(t, u_t) - \varphi(0, u_0) &= \sum_{k=0}^{n-1} (\varphi(t_{k+1}^n, u_{t_{k+1}^n}) - \varphi(t_k^n, u_{t_k^n})) \\ &= \sum_{k=0}^{n-1} (\varphi(t_{k+1}^n, u_{t_k^n}) - \varphi(t_k^n, u_{t_k^n})) + \sum_{k=0}^{n-1} (\varphi(t_{k+1}^n, u_{t_{k+1}^n}) - \varphi(t_{k+1}^n, u_{t_k^n})),\end{aligned}$$

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

We now are ready to prove the uniqueness of energy solutions:

Theorem 2.16. *Let $(\mathcal{D}(\mathcal{L}), \mathcal{L})$ be the operator constructed in Theorem 2.9 and let u be an incompressible solution to the martingale problem for $\mathcal{L}$. Then the law of u is uniquely determined by its initial distribution ν and the finite-dimensional distributions are given by*

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

where $(T_t)_{t \geqslant 0}$ is the semigroup from Theorem 2.9. In particular, u is a Markov process.

By Lemma 2.14 this holds in particular if u is an energy solution to the stochastic Burgers equation.

Proof. Markov property and uniqueness in law follow from the claimed identity. And the identity can be shown inductively in n, with the induction base and induction step using the same argument. So we only treat the case $n=1$. For $\varphi_1 \in \mathcal{D}(\mathcal{L})$ consider $\varphi(t,x)=T_{t_1-t}\varphi_1(x)$. Then $\varphi \in C(\mathbb{R}_+, \mathcal{D}(\mathcal{L}))\cap C^1(\mathbb{R}_+, L^2(\mu))$ and therefore the martingale problem with time-dependent functions gives

$$0=\mathbb{E}\left[\varphi(t_1,u_{t_1})-\varphi(0,u_0)-\int_0^{t_1}(\partial_s+\mathcal{L})\varphi(s,u_s)\mathrm{d}s\right]=\mathbb{E}[\varphi_1(u_{t_1})-T_{t_1}\varphi_1(u_0)]. \qquad \square$$

Remark 2.17. The construction and uniqueness proof for energy solutions did not rely strongly on the specific equation we considered. Essentially the same arguments work in the following setting[2.6]: Let H be a separable real Hilbert space and let $V\subset H\subset V^*$ be Hilbert spaces. Let A: $V\to V^*$ be a bounded symmetric linear operator with

$$\langle Av,w\rangle \leqslant C_A\|v\|_V\|w\|_V, \qquad \langle Av,v\rangle \geqslant c_A\|v\|_V^2,$$

for constants $c_A, C_A>0$. Let $K\colon V\times V\to V^*$ be a symmetric and bounded bilinear map such that

$$\|K(\varphi)\|_{V^*}\leqslant C_K\|\varphi\|_{H^{\otimes 1}(V)}:=C_K\left(\sum_k\left\|\sum_\ell \langle\varphi, e_k\otimes e_\ell\rangle_{H^{\otimes 2}}\right\|_V^2\right)^{1/2}$$

for all

$$\varphi\in\mathcal{S}_2:=\operatorname{span}\{\Pi(h_1\otimes h_2)\colon h_i\in V\}\subset H_s^{\otimes 2},$$

where $K(\Pi(h_1\otimes h_2)):=K(h_1,h_2)$ and extended linearly, and where $(e_k)_k$ is an orthonormal basis of H. Then there is a unique energy solution to the SPDE

$$\partial_t u=-Au+{:}K(u,u){:}+\sqrt{2}A^{1/2}\xi,$$

where ξ is a space-time white noise on H, i.e. such that $(\xi(f))_{f\in L^2(\mathbb{R}_+,H)}$ is centered Gaussian with $\mathbb{E}[\xi(f)\xi(g)]=\int_0^\infty\langle f(t),g(t)\rangle_H\mathrm{d}t$, and where formally the renormalized nonlinearity is defined as

$${:}K(u,u){:}=W_2(K)(u)=\sum_{k,\ell}(u(e_k)u(e_\ell)-\delta_{k,\ell})K(e_k,e_\ell)$$

for an orthonormal basis $(e_k)_k\subset V$ of H.

Example 2.18. For example, we can construct energy solutions of the following equations:

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 existence holds in the supercritical regime. For $\theta\geqslant\frac{3}{4}$ our arguments yield weak uniqueness.

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 scaling critical for $\gamma=1$. For $\gamma\geqslant 1$ our arguments yield weak uniqueness. For $\gamma\in[0,1)$ the recent work [HLZZ24] proves the non-uniqueness of weak solutions, but not of energy solutions. It remains unclear if the energy estimate can serve as a selection principle to select a unique weak solution in the supercritical case. See also the discussion below for singular diffusions.

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 scaling critical for $\gamma=1/2$. For $\gamma\geqslant 1/2$ our arguments yield weak uniqueness.

2.6. Work in progress with Lukas Gräfner and Shyam Popat.

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 [CGT23] and the references therein.

3 Lecture 3: Diffusions with singular drift

Here 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 $\operatorname{div} b = 0$, so by the discussion in Section 2.1 the ODE $\dot{x}(t) = b(x(t))$ preserves the Lebesgue measure. Since Lebesgue measure is also invariant for the Brownian motion, the SDE formally has the Lebesgue measure as stationary measure. To simplify the presentation we restrict our attention to b that are periodic with respect to integer shifts, and we interpret X as a $\mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$-valued process, so that the stationary Lebesgue measure is a probability measure. It is possible to treat also non-periodic b or to see X as an $\mathbb{R}^d$-valued process, see [GP24], but it would complicate the presentation.

We are interested in the case where b is only a distribution. We will measure regularity in Besov spaces, see [BCD11] for a good introduction. If you are not familiar with Besov spaces, you may also interpret $B^\alpha_{p,q}$ as the Bessel potential space $W^{\alpha,p}$ or as the Sobolev–Slobodeckij space $H^{\alpha,p}$; this suggests correctly that the upper index $\alpha \in \mathbb{R}$ measures the regularity of a function, while $p \in [1, \infty]$ describes the integrability scale in which the regularity is measured, and $q \in [1, \infty]$ is a fine-tuning index that is not very important for us. For $b \in B^{-\gamma}_{p,q}$ with $p > \frac{d}{1-\gamma}$ the SDE is subcritical, in the sense that under the rescaling $X^\varepsilon_t = \varepsilon X_{\varepsilon^{-2} t}$ we formally obtain an equation

$$\mathrm{d}X^\varepsilon_t = b_\varepsilon(X^\varepsilon_t)\mathrm{d}t + \mathrm{d}B_t$$

with a “better behaved” b_ε (smooth plus a non-smooth part which is “small” in $B^{-\gamma}_{p,q}$ norm[3.1]). We see that lower integrability requires higher regularity.

For $p > \frac{d}{1-\gamma}$ <u>and</u> $\gamma < \frac{1}{2}$ we can weakly solve the equation without relying on energy solutions, see [FIR17]. But for $\gamma > \frac{1}{2}$ there are counterexamples to weak uniqueness even for the “best case integrability”, $p = \infty$, see [KP23b]. These counterexamples are for non-divergence-free b, for which, inspired by [DD16], we need to include higher order “rough path type” information to guarantee weak uniqueness. We will see that for divergence-free b and by assuming additionally an energy estimate on top of the weak solution property, we can guarantee weak uniqueness in the supercritical regime $\gamma \in [0, 1)$ and

$$p > \frac{2}{1-\gamma},$$

or, in certain cases, even for $b \in B^{-1}_{2+\varepsilon,2}$.

3.1 Construction of energy solutions

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 3.1. *Let X solve*

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

3.1. The norm $\|b_\varepsilon\|_{B^{-\gamma}_{p,q}}$ is not necessarily small, because we are considering non-homogeneous spaces and for constant b we get $b_\varepsilon = \varepsilon^{-1} b$. In the corresponding homogeneous Besov space $\dot{B}^{-\gamma}_{p,q}$ the norm of b_ε would indeed converge to 0 as $\varepsilon \to 0$, provided that $b \in \dot{B}^{-\gamma}_{p,q}$.

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 (we call Y a periodic diffusion*) 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 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 short PDE-argument based on a spectral gap estimate: 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 $\langle \cdot, \cdot \rangle$

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

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 Itô trick now takes the following form:

Lemma 3.2. (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 $Y_0 \sim \nu$ with $\nu \ll \lambda$ for the Lebesgue measure λ on $\mathbb{T}^d$. Then we have for all $T > 0$, for all $f \in C^2(\mathbb{T}^d)$ and for all $p \in [2, \infty)$*

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

where the implicit constant on the right hand side is independent of b*.*

Proof. This follows from the same proof as for the approximate stochastic Burgers equation, see Lemma 2.10. □

Corollary 3.3. (Energy estimate) *Let* Y, ν, f *and* p *be as in the previous lemma. Then*

$$\mathbb{E}\left[\sup_{r\in[s,t]}\left|\int_s^r f(Y_\tau)\mathrm{d}\tau\right|^p\right]\lesssim\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\lambda}\right\|_{L^2(\mathbb{T}^d)}\left(|t-s|^p+|t-s|^{\frac{p}{2}}\right)\left\|\left(1-\frac{1}{2}\Delta\right)^{-1/2}f\right\|_{L^{2p}}^p,$$

where the implicit constant on the right hand side is independent of b*.*

Proof. The proof is again the same as for the stochastic Burgers equation, see Lemma 2.11. The only difference is that now we can work in L^p spaces instead of restricting to $p=1$, because $\nabla\left(1-\frac{1}{2}\Delta\right)^{-1/2}$ is a bounded operator on L^{2p}, which follows from Calderon-Zygmund theory. □

Theorem 3.4. (Existence of energy solutions) *Let* $(b_n)_{n\in\mathbb{N}}\subset C^\infty(\mathbb{T}^d)$ *be divergence-free and such that* $b_n\to b$ *in* $B^{-1}_{2+\varepsilon,2}$*. Let* $X^n_0\sim\nu$ *with* $\frac{\mathrm{d}\nu}{\mathrm{d}\lambda}\in L^2(\mathbb{T}^d)$ *and*

$$\mathrm{d}X^n_t=b_n(X^n_t)\mathrm{d}t+\mathrm{d}B_t,$$

where the Brownian motion B *is independent of* ν*, and let* $Y^n_t=X^n_t \bmod \mathbb{Z}^d$*. Then* $(Y^n)_{n\in\mathbb{N}}$ *is tight in* $C(\mathbb{R}_+,\mathbb{T}^d)$ *and any limit point* Y *satisfies:*

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

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

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

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

In particular, for each $T>0$ *there is a unique continuous extension of the map*

$$I\colon C(\mathbb{T}^d)\to L^1(\Omega,C([0,T])),\qquad I(f)_t=\int_0^t f(X_s)\mathrm{d}s,$$

to the Sobolev space $H^{-1}=B^{-1}_{2,2}$ *(and the extensions for different* T *are consistent), and we denote the extension with the same symbol* I*.*

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).*

Proof. We only address the tightness proof, as the rest is shown in exactly the same way as for the stochastic Burgers equation. To prove tightness, we apply the energy estimate to b_n and obtain under the stationary initial condition λ:

$$\begin{aligned}\mathbb{E}_\lambda\left[\left|\int_s^t b_n(Y^n_r)\mathrm{d}r\right|^{2+\varepsilon}\right] &\lesssim \left(|t-s|^{2+\varepsilon}+|t-s|^{\frac{2+\varepsilon}{2}}\right)\left\|\left(1-\frac{1}{2}\Delta\right)^{-1/2}b_n\right\|_{L^{2+\varepsilon}}^{2+\varepsilon}\\ &\lesssim \left(|t-s|^{2+\varepsilon}+|t-s|^{\frac{2+\varepsilon}{2}}\right)\left\|\left(1-\frac{1}{2}\Delta\right)^{-1/2}b_n\right\|_{B^0_{2+\varepsilon,2}}^{2+\varepsilon},\\ &\lesssim \left(|t-s|^{2+\varepsilon}+|t-s|^{\frac{2+\varepsilon}{2}}\right)\|b_n\|_{B^{-1}_{2+\varepsilon,2}}^{2+\varepsilon},\end{aligned}$$

where the second step follows from Theorem 2.40 in [BCD11] and the third step from Lemma 2.2 in [BCD11]. By assumption, the right hand side is uniformly bounded in n, and since the initial distribution of Y^n does not depend on n we obtain tightness under the stationary initial distribution. This tightness is preserved under an absolutely continuous change of measure, see [GP24]. $\square$

Note that to make sense of the martingale problem we need at least $b\cdot\nabla f\in B^{-1}_{2,2}$ for all $f\in C^\infty(\mathbb{T}^d)$, which only holds if $b\in B^{-1}_{2,2}$. In that sense our condition $b\in B^{-1}_{2+\varepsilon,2}$ for the existence of energy solutions is nearly equivalent to the condition that is required to even define energy solutions. It is also far in the supercritical regime, as $2+\varepsilon<\frac{2}{1-\gamma}=\infty$ because $\gamma=1$.

Under such general conditions we are unable to prove weak uniqueness. Before introducing additional conditions which guarantee the uniqueness of energy solutions, we construct a/the semigroup, under different conditions.

3.2 Construction of the semigroup and uniqueness of energy solutions

Let $\mathcal{H}=L^2(\mathbb{T}^d)$ and $\mathcal{L}_0=\frac{1}{2}\Delta$ and

$$\mathcal{H}_n=\begin{cases}\mathcal{H}, & n=1,\\ \{0\}, & n\neq 1.\end{cases}$$

Then

$$\mathcal{H}^\alpha_\beta=H^\alpha(\mathbb{T}^d),\qquad \alpha,\beta\in\mathbb{R},$$

and $\mathcal{C}$ consists of the functions in $C^\infty(\mathbb{T}^d)$ with compactly supported Fourier transform.

Let $b\in H^{-1}$ be divergence free and

$$\mathcal{G}f:=b\cdot\nabla f,$$

so that formally $\mathcal{G}$ is antisymmetric in $\mathcal{H}$, see the discussion of energy solutions above. One can show with paraproduct estimates and Besov embedding(Theorems 2.82 and 2.85 and Proposition 2.71 in [BCD11]) that for $f\in H^1=\mathcal{H}^1_0$

$$\mathcal{G}f=b\cdot\varolessthan\nabla f+b\cdot\varogreaterthan\nabla f+\nabla\cdot(b\odot f)\in B^{-1}_{1,1}+B^{-1}_{1,2}+B^{-1}_{1,\infty}\subset B^{-1}_{1,\infty}\subset B^{-1-\frac{d}{2}-\kappa}_{2,2}=H^{-1-\frac{d}{2}-\kappa}=\mathcal{H}_0^{-1-\frac{d}{2}-\kappa}$$

for any $\kappa>0$. By Proposition 1.2 there exists a domain $\mathcal{D}(\mathcal{L})=\mathcal{R}_1L^2$ for $\mathcal{L}=\mathcal{L}_0+\mathcal{G}$ and $(\mathcal{D}(\mathcal{L}),\mathcal{L})$ generates a possibly non-unique semigroup. We can slightly adapt the approximation to guarantee positivity of the resolvents along the approximation and thus also in the limit, see [GP24]: For all $f\geqslant 0$ we have $\mathcal{R}_1 f\geqslant 0$. Then $\mathcal{R}_1$ is a contraction in L^∞, which will be important below.

As discussed above, if we can show that $(1-\mathcal{L})\mathcal{C}$ is dense in $\mathcal{H}_0^{-1}$, then energy solutions are solutions to the martingale problem for $\mathcal{L}$, and thus unique in law. A sufficient condition for this is that $(1-\mathcal{L})$ and $(1-\mathcal{L}^*)$ are injective on $\mathcal{H}^1_0=H^1$.

Here is a sufficient condition to prove the injectivity of $(1-\mathcal{L})$ on $\mathcal{H}^1_0$. Note that we can always write a divergence-free $b\in\mathcal{S}'(\mathbb{T}^d)$ as divergence of an antisymmetric matrix, plus a constant, through the Helmholtz decomposition

$$A_{ij}(b)=\Delta^{-1}(\partial_i b^j-\partial_j b^i)\qquad\Rightarrow\qquad\sum_i\partial_i A_{ij}(b)=\Delta^{-1}\Delta b^j-\Delta^{-1}\partial_j\nabla\cdot b=b^j-\mathcal{F}b^j(0)-0,$$

and that A "gains one derivative" compared to b.

Theorem 3.5. *Let $\mu\ll\mathrm{Leb}$ be a probability measure on M. Let $b\in\mathcal{S}'(\mathbb{T}^d,\mathbb{R}^d)$ fulfill the following properties: $A(b)\in L^p$ for some $p\in(2,\infty]$, and there exists a sequence $(g_n)\subset L^\infty$ such that $\nabla g_n\in L^2\cap L^{\frac{2p}{p-2}}$ and*

a) $\nabla g_n\to 0$ weakly in L^2 and $\int g_n h\to\int h$ for all $h\in L^1$.

b) $g_n A\in L^\infty$, $n\in\mathbb{N}$.

c) $A\cdot\nabla g_n\to 0$ weakly in L^2.

Then $(1-\mathcal{L})$ and $(1-\mathcal{L}^*)$ are injective and thus energy solutions to the SDE with drift b are unique.

Proof. (Idea of the proof) The idea is to test the equation against u and to integrate by parts to see that $\langle(1-\mathcal{L})u, u\rangle = \|u\|^2_{\mathcal{H}^1_0}$. Due to the low regularity of u and A, this requires several approximations, which is where the sequence $(g_n)_{n\in\mathbb{N}}$ comes in. See [GP24] for details. □

In the following lemma we give a more explicit sufficient condition:

Lemma 3.6. *Let $A\in L^p(\mathbb{T}^d;\mathbb{R}^{d\times d})$ for $p\in(2,\infty]$ be antisymmetric* Leb − *a.e. and suppose there exists a compact set K such that for $\varepsilon>0$ and $B^\varepsilon=\{x: d(x,K)\leqslant\varepsilon\}$*

$$\sup_{\varepsilon>0}\varepsilon^{-2}\mathrm{Leb}(B^\varepsilon), \qquad \sup_{\varepsilon>0}\varepsilon^{-2}\int_{B^\varepsilon}|A|^2<\infty, \tag{3.1}$$

and for all $\varepsilon>0$,

$$A\mathbb{1}_{B^c_\varepsilon}\in L^\infty. \tag{3.2}$$

Then, there exists a sequence g_n satisfying the assumptions from Theorem 3.5.

We can construct examples of A where this condition is satisfies but such that A is not in L^p, where $p>2$ is arbitrarily close to 2; but the closer we get to 2, the higher-dimensional our example becomes. See again [GP24].

Here is a different type of condition to gurantee the well-posedness of the martingale problem.

Lemma 3.7. *Let $b\in H^{-1}$ be such that for some mollifier ρ with compactly supported Fourier transform and $\rho_n=n^d\rho(n\cdot)$ we have*

$$b\cdot\nabla(\rho_n * f)\to b\cdot\nabla f \qquad \text{in } H^{-1}$$

for all $f\in H^1\cap L^\infty$. Then any energy solution to the singular SDE with drift b is also a solution to the martingale problem for $(\mathcal{D}(\mathcal{L}),\mathcal{L})$, where $\mathcal{D}(\mathcal{L})=\mathcal{R}_1L^2$ is the domain constructed in Proposition 1.2, and in particular it is unique in law and a Markov process.

Proof. Let $f\in\mathcal{D}^\infty(\mathcal{L}):=\mathcal{R}_1L^\infty$. Since $\mathcal{R}_1$ is a contraction in L^∞ and $\mathcal{R}_1L^\infty\subset\mathcal{R}^1L^2\subset H^1$, we get $f\in L^\infty\cap H^1$. Let $f_n=\rho_n * f$. Clearly $f_n\to f$ in L^2, and by assumption $\mathcal{L}f_n\to\mathcal{L}f$ in H^{-1}. Moreover, $f_n\in\mathcal{C}$ and thus by the definition of energy solutions the process

$$M^{f_n}_t=f_n(Y_t)-f_n(Y_0)-I(\mathcal{L}f_n)_t, \qquad t\geqslant 0,$$

is a continuous martingale. By incompressibility of Y and L^2-convergence $f_n\to f$ we get

$$f_n(Y_t)-f_n(Y_0)\overset{L^1(\Omega)}{\longrightarrow}f(Y_t)-f(Y_0),$$

while the energy estimate yields the convergence

$$I(\mathcal{L}f_n)_t\overset{L^1(\Omega)}{\longrightarrow}I(\mathcal{L}f)_t.$$

The $L^1(\Omega)$-limit of martingales is a martingale, which proves the claim for $f\in\mathcal{D}^\infty(\mathcal{L})$.

For general $f=\mathcal{R}_1g\in\mathcal{D}(\mathcal{L})$ (with $g\in L^2$ but not necessarily $g\in L^\infty$), let $g_n=(g\vee(-n))\wedge n\in L^\infty$ and $f_n=\mathcal{R}_1g_n\in\mathcal{D}^\infty(\mathcal{L})$. Then $(1-\mathcal{L})f_n=g_n\to g$ in L^2 and $f_n=\mathcal{R}_1g_n\to f$ in H^1 by continuity of $\mathcal{R}_1\in L(L^2,H^1)$. Thus, also $\mathcal{L}f_n\to\mathcal{L}f$ in L^2 and it follows from the incompressibility of Y that M^f is a martingale. □

Lemma 3.8. *Let $b\in L^2+B^{-1}_{\infty,1}\supset\bigcup_{p\in[1,\infty]}B^{-\gamma}_{\frac{2}{1-\gamma},1}$ be divergence-free. Then the condition of the previous lemma is satisfied.*

Proof. The claim $L^2+B^{-1}_{\infty,1}\supset\bigcup_{p\in[1,\infty]}B^{-\gamma}_{\frac{2}{1-\gamma},1}$ follows by interpolation space type arguments, see [GP24]. Let us show that for $b_1\in L^2$ and for $b_2\in B^{-1}_{\infty,1}$ the condition from the previous lemma is satisfied, then it clearly also holds for $b=b_1+b_2$.

For $b_1\in L^2$ we bound

$$\|b_1\cdot\nabla(f-f_n)\|_{H^{-1}}=\|\nabla\cdot(b_1(f-f_n))\|_{H^{-1}}\lesssim\|b_1(f-f_n)\|_{L^2},$$

and since

$$|b_1(f-f_n)| \leqslant |b_1|(\|f\|_\infty + \|f_n\|_\infty) \lesssim |b_1| \|f\|_\infty \in L^2$$

and f_n converges in measure to f, we get from the dominated convergence theorem

$$\|b_1(f-f_n)\|_{L^2} \to 0.$$

For $b_2 \in B^{-1}_{\infty,1}$ we get $A(b_2) \in B^0_{\infty,2} \subset L^\infty$ by Lemma 2.2 in [BCD11] and in particular (assuming $\mathcal{F}b_2(0) = 0$ for simplicity)

$$\|b_2 \cdot \nabla(f-f_n)\|_{H^{-1}} = \|\nabla \cdot (A \nabla(f-f_n))\|_{H^{-1}} \lesssim \|A \nabla(f-f_n)\|_{L^2} \lesssim \|A\|_{L^\infty} \|f-f_n\|_{H^1},$$

which converges to 0. □

The idea of using the maximum principle of $\mathcal{R}_1$ and the associated a priori L^∞ bound for solutions to the resolvent equation with right hand side in L^∞ to go beyond subcriticality is due to Zhang-Zhao [ZZ21].

Remark 3.9. In [MS18], Theorem 1.9, the authors use convex integration techniques to construct for every $d \geqslant 4$ and $\tilde{p} \in [1, \infty)$ with

$$\frac{1}{\tilde{p}} > \frac{1}{2} + \frac{1}{d-1},$$

a divergence-free drift $b \in C([0,T]; W^{1,\tilde{p}}(\mathbb{T}^d) \cap L^2(\mathbb{T}^d))$ and an initial density $\rho_0 \in L^2(\mathbb{T}^d)$ such that there are multiple weak solutions ρ to the Fokker-Planck equation

$$\partial_t \rho = \Delta \rho - \nabla \cdot (b\rho), \qquad \rho(0) = \rho_0.$$

Morally, this corresponds to weak non-uniqueness for the SDE. But by a time-dependent extension of our results (which is possible) we get weak uniqueness of energy solutions for $b \in C([0,T]; L^2(\mathbb{T}^d))$ even without assuming $b \in C([0,T]; W^{1,\tilde{p}}(\mathbb{T}^d))$. Therefore, the energy estimate in the definition of energy solutions can be interpreted as a selection principle, an additional assumptions which guarantees uniqueness of weak solutions. See Definition 1 on p.54 of [LL19] for a different type of selection principle for this problem.

Looking directly at weak solutions of the SDE instead of the Fokker-Planck equation, [KP23b] constructs a one-dimensional $b \in B^{-1/2-\varepsilon}_{\infty,\infty}$ (not divergence-free, which in $d=1$ would mean constant) such that weak solutions to the SDE with drift b are not unique in law. I expect that this example can be extended to higher dimensions and to divergence-free b, for example by considering a shear flow drift, $b(x,y) = (0, b(x))$ with the same b as in [KP23b]. But again the energy solution for an SDE with divergence free $b \in B^{-1/2-\varepsilon}_{\infty,\infty}$ is weakly well-posed and thus the energy estimate serves as a selection principle.

4 Lecture 4: Applications

4.1 Hairer-Quastel universality

This is from a joint work with Massimiliano Gubinelli [GP16], 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 only 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 $\varepsilon^{1/2}\mathcal{P}_{\varepsilon^{-1}}u_\varepsilon(x)$ 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 $D_x W_n(\varphi_n) = nW_{n-1}(\varphi_n(x,\cdot))$ 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 tedious[4.1], 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 2.12, we are only going to verify that limit points of u_ε satisfy the condition i'. from Remark 2.13 following the theorem, i.e. we only treat linear test functions instead of cylinder functions, and keep track manually of the quadratic variation. We will also verify ii. and iii., so that by the discussion in Remark 2.13 we get that the limit is the unique in law energy solution of the stochastic Burgers equation.

4.1. But a more general bound is derived in the still unpublished Master's thesis of Da Li at FU Berlin.

Thus, let $\varphi(u)=u(f)$ for $f\in C^\infty$. Then

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

and taking the Fourier transform:

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

so that

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

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 $\varphi(u)=u(f)$:

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

where in the last step we used that $\left(\frac{1}{\sqrt{m!}}H_m\right)$ 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

$$\partial_t u=\Delta u+c_2(F)\partial_x u^2+\sqrt{2}(-\Delta)^{1/2}\xi.$$

4.2 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 [GK24] using the Cole-Hopf transform. Using energy solutions and the generator $\mathcal{L}$, we could extend them to fractional, multi-component KPZ equations by the same arguments, by a small extension of the classical Kipnis-Varadhan martingale approach [KLO12].

Definition 4.1. *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 4.2. *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 rise to Gaussian fluctuations. 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 $\big(M_s = M_s^{p_{t-\cdot}(x-\cdot)}\big)_{s \in [0,t]}$ 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

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

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

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

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 4.3. *We have*

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

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 [GP18]. 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 4.4. *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 $\varphi \in L^2(\mu)$

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

Proof. Let $\varphi \in \mathcal{D}(\mathcal{L})$ with $\varphi_0 = 0$. We showed in Theorem 1.4 that

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

and we have by Parseval's identity 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 bound on the semigroup we simply differentiate $\|T_t\varphi\|_{L^2}^2$: Since $T_t\varphi_0 = \varphi_0 = 0$, we get

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

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 4.5. *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. For $\varphi \in \mathcal{H}$ it is classical that there exists a unique solution $\psi \in \mathcal{D}(\mathcal{L})$ to the Poisson equation. We can simply set

$$\psi = \int_0^\infty T_t\varphi \mathrm{d}t,$$

which converges by the spectral gap estimate and which satisfies (these formal arguments can be justified with a little bit of work, see for example the first Chapter of [EK86]):

$$\mathcal{L}\psi = \int_0^\infty \mathcal{L}T_t\varphi \mathrm{d}t = \int_0^\infty \partial_t T_t\varphi \mathrm{d}t = T_\infty\varphi - T_0\varphi = -\varphi.$$

Since we know that

$$\langle (-\mathcal{L}_0)\psi, \psi\rangle = \langle (-\mathcal{L})\psi, \psi\rangle = \langle \varphi, \psi\rangle \leqslant \|\varphi\|_{\mathcal{H}_0^1} \|\psi\|_{\mathcal{H}_0^{-1}},$$

and once more by the spectral gap estimate (this time for $\mathcal{L}_0$), we get

$$\|\psi\|_{\mathcal{H}_0^1}^2 \simeq \langle (-\mathcal{L}_0)\psi, \psi\rangle,$$

so that

$$\|\psi\|_{\mathcal{H}_0^1} \lesssim \|\varphi\|_{\mathcal{H}_0^{-1}},$$

and there exists a unique continuous extension of the map $\mathcal{R}_0\varphi = \psi$ to $\mathcal{H}_0^{-1}$. 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 4.2) 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:

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

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 $t = 1$. □

Remark 4.6. With a similar but simpler variant of the same arguments we can also show a singular periodic homogenization result: If

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

where b is a divergence-free periodic distribution as in Section 3 but X is now $\mathbb{R}^d$-valued and not interpreted as a periodic process, then

$$t^{-1/2}(X_t - t\mathcal{F}b(0))$$

converges weakly to a normal distribution (and there is also a functional limit result). This partly recovers the result of [KP23a] with a simpler proof. But [KP23a] also allows Lévy noise and non-divergence free b.

Appendix A Chaos expansion and Fock space

Here we consider the white noise over $L^2(\mathbb{T})$ for concreteness, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the one-dimensional torus, but with a slightly different setup the following constructions works for the white noise over any separable real Hilbert space. A white noise over H is a centered Gaussian process $(\eta(h))_{h \in H}$ with covariance $\mathbb{E}[\eta(h)\eta(g)] = \langle h, g \rangle$.

Let now η be a white noise over $L^2(\mathbb{T})$.

Definition A.1. (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 $\Pi \colon L^2(\mathbb{T}^n) \to L_s^2(\mathbb{T}^n)$,*

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

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 $L^2(\mathbb{T}^n)$-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 A.2. (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 A.3. *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 A.4. *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

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

where

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

We will only sketch the argument.

Definition A.5. *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 $H_n(x, t) := t^{n/2} H_n\left(\frac{x}{\sqrt{t}}\right)$ for $t > 0, x \in \mathbb{R}$ 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 A.6. *Let* M *be a continuous local martingale with* $M_0 = 0$*. Then*

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

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

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

□

Corollary A.7. *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 A.6:

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

□

Corollary A.8. *We have the chaos representation property*

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

where

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

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

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

Proof. It suffices to apply Corollary A.7 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 $\bigoplus_{n \geqslant 0} \mathcal{H}_n$ is orthogonal to all polynomial of $\eta(\varphi)$, for all $\varphi \in L^2(\mathbb{T})$. See Theorem 1.1.1 of [Nua06] for details. □

The following result is very powerful and it has many applications.

Theorem A.9. (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 $\varphi \in L_s^2(\mathbb{T}^n)$:*

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

Proof. For $p < 2$ we can take $C_p = 1$ and apply Jensen's inequality, so let $p \geqslant 2$ and let $\varphi \in L_s^2(\mathbb{T}^n)$. 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

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

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

Bibliography

[ABK24] Scott Armstrong, Ahmed Bou-Rabee, and Tuomo Kuusi. Superdiffusive central limit theorem for a Brownian particle in a critically-correlated incompressible random drift. *ArXiv preprint arXiv:2404.01115*, 2024.

[BCD11] Hajer Bahouri, Jean-Yves Chemin, and Raphael Danchin. *Fourier analysis and nonlinear partial differential equations*. Springer, 2011.

[BG97] Lorenzo Bertini and Giambattista Giacomin. Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.*, 183(3):571–607, 1997.

[CGT23] Giuseppe Cannizzaro, Massimiliano Gubinelli, and Fabio Toninelli. Gaussian fluctuations for the stochastic Burgers equation in dimension ≥ 2. *ArXiv preprint arXiv:2304.05730*, 2023.

[CHT22] Giuseppe Cannizzaro, Levi Haunschmid-Sibitz, and Fabio Toninelli. $\sqrt{\log t}$-superdiffusivity for a Brownian particle in the curl of the 2D GFF. *Ann. Probab.*, 50(6):2475–2498, 2022.

[CMOW22] Georgiana Chatzigeorgiou, Peter Morfe, Felix Otto, and Lihan Wang. The gaussian free-field as a stream function: asymptotics of effective diffusivity in infra-red cut-off. *ArXiv preprint arXiv:2212.14244*, 2022.

[DD16] François Delarue and Roland Diel. Rough paths and 1d SDE with a time dependent distributional drift: application to polymers. *Probab. Theory Related Fields*, 165(1-2):1–63, 2016.

[DL89] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.

[EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes: Characterization and convergence*. John Wiley & Sons, 1986.

[FIR17] Franco Flandoli, Elena Issoglio, and Francesco Russo. Multidimensional stochastic differential equations with distributional drift. *Trans. Amer. Math. Soc.*, 369(3):1665–1688, 2017.

[GIP15] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics, Pi*, 3(e6), 2015.

[GJ13] Massimiliano Gubinelli and Milton Jara. Regularization by noise and stochastic Burgers equations. *Stochastic Partial Differential Equations: Analysis and Computations*, 1(2):325–350, 2013.

[GJ14] Patrícia Gonçalves and Milton Jara. Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Arch. Ration. Mech. Anal.*, 212(2):597–644, 2014.

[GJ18] Patrícia Gonçalves and Milton Jara. Density fluctuations for exclusion processes with long jumps. *Probab. Theory Related Fields*, 170(1-2):311–362, 2018.

[GK24] Yu Gu and Tomasz Komorowski. KPZ on torus: Gaussian fluctuations. *Ann. Inst. Henri Poincaré Probab. Stat.*, 60(3):1570–1618, 2024.

[GP16] Massimiliano Gubinelli and Nicolas Perkowski. The Hairer–Quastel universality result at stationarity. *RIMS Kôkyûroku Bessatsu*, B59, 2016.

[GP18] Massimiliano Gubinelli and Nicolas Perkowski. Energy solutions of KPZ are unique. *J. Amer. Math. Soc.*, 31(2):427–471, 2018.

[GP20] Massimiliano Gubinelli and Nicolas Perkowski. The infinitesimal generator of the stochastic Burgers equation. *Probab. Theory Related Fields*, 178(3-4):1067–1124, 2020.

[GP24] Lukas Gräfner and Nicolas Perkowski. Weak well-posedness of energy solutions to singular SDEs with supercritical distributional drift. *ArXiv preprint arXiv:2407.09046*, 2024.

[Hai11] Martin Hairer. Rough stochastic PDEs. *Comm. Pure Appl. Math.*, 64(11):1547–1585, 2011.

[Hai13] Martin Hairer. Solving the KPZ equation. *Ann. Math.*, 178(2):559–664, 2013.

[Hai14] Martin Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[HLZZ24] Martina Hofmanová, Xiaoyutao Luo, Rongchan Zhu, and Xiangchan Zhu. Surface quasi-geostrophic equation perturbed by derivatives of space-time white noise. *Mathematische Annalen*, pages 1–42, 2024.

[HQ18] Martin Hairer and Jeremy Quastel. A class of growth models rescaling to KPZ. *Forum Math. Pi*, 6:0, 2018.

[HZ23] Zimo Hao and Xicheng Zhang. SDEs with supercritical distributional drifts. *ArXiv preprint arXiv:2312.11145*, 2023.

[KLO12] Tomasz Komorowski, Claudio Landim, and Stefano Olla. *Fluctuations in Markov processes*, volume 345 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2012. Time symmetry and martingale approximation.

[KNV07] 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.

[KP23a] Helena Kremp and Nicolas Perkowski. Periodic homogenization for singular Lévy SDEs. *ArXiv preprint arXiv:2309.16225*, 2023.

[KP23b] Helena Kremp and Nicolas Perkowski. Rough weak solutions for singular Lévy SDEs. *ArXiv preprint arXiv:2309.15460*, 2023.

[KPZ86] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang. Dynamic scaling of growing interfaces. *Physical Review Letters*, 56(9):889–892, 1986.

[LL19] Claude Le Bris and Pierre-Louis Lions. *Parabolic equations with irregular data and related issues—applications to stochastic differential equations*, volume 4 of *De Gruyter Series in Applied and Numerical Mathematics*. De Gruyter, Berlin, [2019] ©2019.

[MQR21] Konstantin Matetski, Jeremy Quastel, and Daniel Remenik. The KPZ fixed point. *Acta Math.*, 227(1):115–203, 2021.

[MS18] Stefano Modena and László Székelyhidi, Jr. Non-uniqueness for the transport equation with Sobolev vector fields. *Ann. PDE*, 4(2):0, 2018.

[Nua06] David Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, Second edition, 2006.

[QS23] 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.

[YY24] Huanyu Yang and Zhilin Yang. Weak coupling limit of a Brownian particle in the curl of the 2d GFF. *ArXiv:2405.05778*, 2024.

[ZZ21] 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.