Chapter 2. Stochastic Navier-Stokes equations and state dependent noise

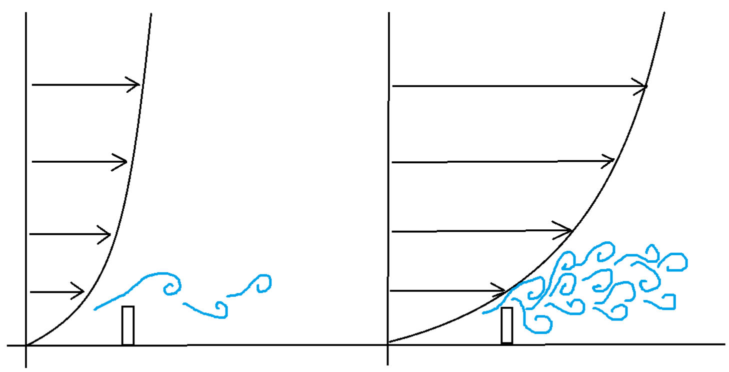
Franco Flandoli, Scuola Normale Superiore
April-May 2021, Waseda University, Tokyo, Japan

April 21, 2021

1 Introduction

Until now, although motivated by certain random input, we dealt with the Stochastic Navier-Stokes equations as if they were deterministic: given a single noise realization, we solve the equation.

This is possible in relatively few cases. The case treated above had the special feature that the random input was independent of the solution. But in real situations, as in the figure (discussed in a section below)



the noise may vary depending on the solution.

Mathematically speaking, in the previous chapter the noise entered the equation as an additive force; this was the key property which allowed us to study the linear Stokes problem first, independently of the solution of the nonlinear one. There are other cases (different from the additive case) which can be treated by similar ideas, but few.

If we have an equation of the form

$$\begin{aligned} \partial_t u + u \cdot \nabla u + \nabla p &= \nu \Delta u + f + F(u) + \sigma(u) \partial_t W \\ \operatorname{div} u &= 0 \end{aligned}$$

where the distributional derivative $\partial_t W$ is multiplied by a function of the solution, we are in trouble. The problem is not just the fact that the Stokes problem

$$\begin{aligned} \partial_t z + \nabla q &= \nu \Delta z + \sigma(u)\, \partial_t W \\ \operatorname{div} z &= 0 \end{aligned}$$

depends on u: this problem in principle could be solved by an iteration. The problem is that we cannot apply the trick of integration by parts in the mild formula for z:

$$\begin{aligned} z(t) &= e^{tA} z_0 + \int_0^t e^{(t-s)A} \sigma(u(s))\, \partial_s W(s)\, ds \\ &= e^{tA} z_0 + \left[e^{(t-s)A} \sigma(u(s))\, W(s) \right]_{s=0}^{s=t} - \int_0^t \frac{d}{ds} \left(e^{(t-s)A} \sigma(u(s)) \right) W(s)\, ds \\ &= e^{tA} z_0 + \sigma(u(t))\, W(t) - e^{tA} \sigma(u(0))\, W(0) \\ &\quad + \int_0^t A e^{(t-s)A} \sigma(u(s))\, W(s)\, ds + \int_0^t e^{(t-s)A} \frac{d}{ds} \sigma(u(s))\, W(s)\, ds \end{aligned}$$

and

$$\frac{d}{ds} \sigma(u(s)) = \langle D\sigma(u(s)), \partial_s u(s) \rangle$$

brings again into play the term $\partial_s W(s)$.

One way to escape this problem is using the theory of rough paths, which however is quite elaborated for our purposes. The most classical way is, when W is related to Brownian motions, to use stochastic calculus. The purpose of this chapter is illustrating the technique to study the Stochastic Navier-Stokes equations by stochastic calculus.

Remark 1 *The reader certainly noticed that we have introduced, in parallel to $\sigma(u)\, \partial_t W$, also a term $F(u)$. This is not for generality, which clearly is not our purpose in these notes. The reason is deep: if we introduce a term $\sigma(u)\, \partial_t W$, we also need to introduce a compensator $F(u)$, otherwise the Physics is wrong. This is Wong-Zakai principle: we shall describe it in two particular cases, in this and the next chapters.*

1.1 Filtered probability space

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A filtration indexed by $t \geq 0$ is a family $(\mathcal{F}_t)_{t\geq 0}$ of σ-algebras such that $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2} \subset \mathcal{F}$ for every $t_1 \leq t_2$. We call $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ a filtered probability space, and we abbreviate $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. A stochastic process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, taking values in a measurable space, is adapted if X_t is $\mathcal{F}_t$-measurable for every $t \geq 0$. It is progressively measurable if the map $(s, \omega) \mapsto X_s(\omega)$ is measurable on $([0,t] \times \Omega, \mathcal{B}(0,t) \otimes \mathcal{F}_t)$ for every $t \geq 0$ ($\mathcal{B}(0,t)$ being the Borel σ-algebra on $[0,t]$). When the target space is metric with the Borel σ-algebra, and the process is continuous, the concepts of adapted and progressively measurable are equivalent. When we deal with processes such that, with respect to the time variable, are equivalence classes (with respect to zero sets for the Lebesgue measure on the time interval), like $L^2(0,T;V)$, we cannot use the concept of adapted process since X_t (given t) is not well defined. In this case we always use the concept of progressively measurable: for every t, the restriction on $[0,t]$ is a well defined equivalence class and the definition applies to it.

Denote by $L^2_{\mathcal{F}_t}(\Omega, H)$ the space of random variables (in fact equivalence classes) $X : \Omega \to H$ that are $\mathcal{F}_t$-measurable and square integrable. We denote by $C_{\mathcal{F}}([0,T];H)$ the space of continuous adapted processes $(X_t)_{t\in[0,T]}$ with values in H such that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X_t\|_H^2\right] < \infty$$

and by $L^2_{\mathcal{F}}(0,T;V)$ the space of progressively measurable processes $(X_t)_{t\in[0,T]}$ with values in V such that

$$\mathbb{E}\left[\int_0^T \|X_t\|_V^2 \, dt\right] < \infty.$$

Of course we may use similar notations also with different spaces in place of H and V; this is just the most common case in the sequel.

A (real valued) Brownian motion on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ is a continuous adapted process $(W_t)_{t\geq 0}$ such that $\mathbb{P}(W_t = 0) = 1$, $W_t - W_s$ is independent of $\mathcal{F}_s$ for every $t \geq s \geq 0$, and $W_t - W_s$ is a centered Gaussian random variable with variance $t-s$ (we write $W_t - W_s \sim N(0, t-s)$). With probability one, paths are not only continuous but also locally Hölder continuous with any Hölder exponent $\alpha < \frac{1}{2}$.

The noise used in Chapter 1 had the form

$$W(t,x) := \sum_{k\in K} \sqrt{\lambda_k} \sigma_k(x) W_t^k \tag{1}$$

where K is a finite set, $\sigma_k \in D(A)$, $\left(W_t^k\right)_{t\geq 0}$ are independent Brownian motions on some filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. With probability one, the path $t \mapsto W(t,\cdot)$ is of class $C([0,T];D(A))$ (also $C^\alpha([0,T];D(A))$ for every $\alpha < \frac{1}{2}$).

In the previous chapter we have denoted by τ^k the average intertimes between creation of new eddies. Here we use the quantity

$$\lambda_k = \frac{1}{\tau^k}$$

which has the meaning of *rate* of eddy production. The reason is that, below, we modify the model with state-dependent rates and the notational analogy will be easier.

2 Additive noise under the view of stochastic calculus

Let us elaborate the result of Chapter 1 under the view of stochastic calculus. Consider the Itô type equation, in $d = 2$,

$$\begin{aligned} du + (u \cdot \nabla u + \nabla p)\, dt &= \nu \Delta u dt + \sum_{k \in K} \sqrt{\lambda_k} \sigma_k dW_t^k \qquad (2) \\ \operatorname{div} u &= 0 \end{aligned}$$

with

$$\begin{aligned} u|_{\partial D} &= 0 \\ u(0) &= u_0. \end{aligned}$$

Definition 2 *Given a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and the noise $W(t,x)$ as in (1), given $u_0 : \Omega \to H$, $\mathcal{F}_0$-measurable, we say that a process u is a solution of equation (2), if its paths are of class*

$$u \in C([0,T]; H) \cap L^2(0,T;V)$$

with probability one, it is adapted as a process in H, progressively measurable in V, and

$$\begin{aligned} & \langle u(t), \phi \rangle - \int_0^t b(u(s), \phi, u(s))\, ds \\ = \ & \langle u_0, \phi \rangle + \int_0^t \langle u(s), A\phi \rangle\, ds + \sum_{k \in K} \sqrt{\lambda_k} \langle \sigma_k, \phi \rangle W_t^k \end{aligned}$$

for every $\phi \in D(A)$.

Theorem 3 *There exists a unique solution.*

Proof. Given two solutions, with probability one their paths are two solutions in the sense of the theorem of the previous Chapter, hence they coincide. Path by path the existence of $u(\omega)$ is given by that theorem; since W is measurable, also u is measurable. But the measurability result can be applied on any subinterval $[0,t]$, the process u being always the same (namely the restriction to $[0,t]$ of the process on $[0,T]$), hence we have progressive measurability, which gives also adaptedness in H due to continuity. ■

We want now to apply Itô formula to compute

$$d \|u(t)\|_{L^2}^2 .$$

Let us recall, for comparison, that when X_t is a process in $\mathbb{R}^d$ satisfying the equation

$$dX_t^i = b_t^i dt + \sum_{k \in K} \sigma_t^{ik} dW_t^k$$

and f is a function of class $C^2\left(\mathbb{R}^d\right)$, then

$$df\left(X_t\right)=\sum_{i=1}^{d}\partial_i f\left(X_t\right)dX_t^i+\frac{1}{2}\sum_{i,j=1}^{d}\sum_{k\in K}\partial_i\partial_j f\left(X_t\right)\sigma_t^{ik}\sigma_t^{jk}dt$$

where we have to replace dX_t^i by the equation. Rigorously, all these identities have to be interpreted in integral form and the stochastic processes $X_t^i, b_t^i, \sigma_t^{ik}$ are assumed progressively measurable. In order to apply these facts we need a progressively measurable process (and this is provided by the previous theorem) and a finite dimensional reduction.

Theorem 4 *If* $\mathbb{E}\left\|u_0\right\|_{L^2}^2<\infty$ *then*

$$u\in C_{\mathcal{F}}\left([0,T];H\right)\cap L_{\mathcal{F}}^2\left(0,T;V\right)$$

and

$$\mathbb{E}\left[\left\|u\left(t\right)\right\|_{L^2}^2\right]+2\nu\int_0^t\mathbb{E}\left\|\nabla u\left(s\right)\right\|_{L^2}^2ds=\mathbb{E}\left[\left\|u_0\right\|_{L^2}^2\right]+t\sum_{k\in K}\lambda_k\left\|\sigma_k\right\|_{L^2}^2$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left\|u\left(t\right)\right\|_{L^2}^2\right]\leq\mathbb{E}\left[\left\|u_0\right\|_{L^2}^2\right]+T\sum_{k\in K}\sqrt{\lambda_k}\left\|\sigma_k\right\|_{L^2}^2+C\sum_{k\in K}\lambda_k\mathbb{E}\int_0^T\left\langle u\left(s\right),\sigma_k\right\rangle^2ds.$$

Proof. Taken a complete orthonormal system in H, (e_i), made of eigenvectors of A, with eigenvalues $(-\lambda_i)$, called H_n the finite dimensional space generated by $e_1,..,e_n$ and π_n the projection onto H_n, called $u_n\left(t\right)=\pi_n u\left(t\right)$, called finally

$$b_n\left(u\left(s\right)\right):=\sum_{i=1}^{n}b\left(u\left(s\right),u\left(s\right),e_i\right)e_i$$

we have (from the weak formulation applied to each e_i)

$$u_n\left(t\right)+\int_0^t b_n\left(u\left(s\right)\right)ds=\pi_n u_0+\int_0^t Au_n\left(s\right)ds+\pi_n W\left(t\right).$$

Taken the function $f_n\left(x\right)=\sum_{i=1}^n\left\langle x,e_i\right\rangle^2$, which has the properties $\partial_i f_n\left(x\right)=2\left\langle x,e_i\right\rangle$, $\partial_i\partial_j f_n\left(x\right)=2\delta_{ij}$, using the fact that, with $\sigma_t^{ik}=\sqrt{\lambda_k}\left\langle\sigma_k,e_i\right\rangle$, one has $\sum_{i=1}^{\infty}\left(\sigma_t^{ik}\right)^2=\lambda_k\left\|\sigma_k\right\|_{L^2}^2$, the classical Itô formula gives us

$$\begin{aligned}d\left\|u_n\left(t\right)\right\|_{L^2}^2&=2\left\langle u_n\left(t\right),du_n\left(t\right)\right\rangle+\sum_{k\in K}\lambda_k\left\|\pi_n\sigma_k\right\|_{L^2}^2dt\\&=-2\nu\left\|\nabla u_n\left(t\right)\right\|_{L^2}^2dt+\sum_{k\in K}\lambda_k\left\|\pi_n\sigma_k\right\|_{L^2}^2dt\\&\quad+2\sum_{k\in K}\sqrt{\lambda_k}\left\langle u_n\left(t\right),\pi_n\sigma_k\right\rangle dW_t^k+b\left(u\left(s\right),u\left(s\right),u_n\left(s\right)\right)dt\end{aligned}$$

where we have used

$$\langle u_n(s), b_n(u(s))\rangle = b(u(s), u(s), u_n(s)).$$

This identity has to be interpreted in integral form. Using the convergence properties of π_n and the regularity of u, it is not difficult to pass to the limit and obtain

$$\begin{aligned} \|u(t)\|_{L^2}^2 + 2\nu \int_0^t \|\nabla u(s)\|_{L^2}^2 ds &= \|u_0\|_{L^2}^2 + t\sum_{k\in K} \lambda_k \|\sigma_k\|_{L^2}^2 \\ &\quad +2\sum_{k\in K}\sqrt{\lambda_k}\int_0^t \langle u(s), \sigma_k\rangle dW_s^k \end{aligned} \tag{3}$$

where the last term is an Itô-integral. In order to take expected values we have to use a localization argument that we explain here forever, namely we omit the repetition below when it is used several times. For sake of simplicity of notations assume that u is a solution defined for all $t \geq 0$ (we can do this, T is arbitrary).For every $R > 0$, let τ_R be the stopping time defined as

$$\tau_R = \inf\{t > 0 : \|u(t)\|_{L^2} > R\}$$

or equal to $+\infty$ if the set is empty. Compute the previous identity at time $t \wedge \tau_R$ (it helps the fact that the process u is continuous in H):

$$\begin{aligned} \|u(t\wedge\tau_R)\|_{L^2}^2 + 2\nu\int_0^t 1_{s\leq\tau_R}\|\nabla u(s)\|_{L^2}^2 ds &= \|u_0\|_{L^2}^2 + (t\wedge\tau_R)\sum_{k\in K}\lambda_k\|\sigma_k\|_{L^2}^2 \\ &\quad +2\sum_{k\in K}\sqrt{\lambda_k}\int_0^t 1_{s\leq\tau_R}\langle u(s), \sigma_k\rangle dW_s^k. \end{aligned}$$

Now $\mathbb{E}\int_0^T 1_{s\leq\tau_R}\langle u(s), \sigma_k\rangle^2 ds < \infty$ hence the Itô integrals of this identity are true martingales; their expected values are thus equal to zero. Moreover, the other terms on the right-hand-side have finite expected value, hence the same is true for the sum of the two terms on the left-hand-side, and then also individually for each of them, being non-negative. We get

$$\begin{aligned} &\mathbb{E}\left[\|u(t\wedge\tau_R)\|_{L^2}^2\right] + 2\nu\mathbb{E}\int_0^t 1_{s\leq\tau_R}\|\nabla u(s)\|_{L^2}^2 ds \\ &= \mathbb{E}\left[\|u_0\|_{L^2}^2\right] + \mathbb{E}[t\wedge\tau_R]\sum_{k\in K}\lambda_k\|\sigma_k\|_{L^2}^2. \end{aligned}$$

Since $\lim_{R\to\infty}\tau_R = +\infty$, and u is continuous in H, we deduce as $R \to \infty$

$$\mathbb{E}\left[\|u(t)\|_{L^2}^2\right] + 2\nu\mathbb{E}\int_0^t \|\nabla u(s)\|_{L^2}^2 ds = \mathbb{E}\left[\|u_0\|_{L^2}^2\right] + t\sum_{k\in K}\lambda_k\|\sigma_k\|_{L^2}^2.$$

From this result, which is already part of the thesis, we deduce $u \in L^2_{\mathcal{F}}(0,T;V)$. In order to prove $u \in C_{\mathcal{F}}([0,T];H)$ we restart from (3) where now, as a consequence of the estimates just proved, we know that the Itô integrals are square integrable martingales. Let us simplify (3) into

$$\|u(t)\|_{L^2}^2 \leq \|u_0\|_{L^2}^2 + t\sum_{k\in K}\lambda_k \|\sigma_k\|_{L^2}^2 + 2\sum_{k\in K}\sqrt{\lambda_k}\int_0^t \langle u(s), \sigma_k\rangle dW_s^k.$$

By Doob's inequality,

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}\|u(t)\|_{L^2}^2\right] &\leq \mathbb{E}\left[\|u_0\|_{L^2}^2\right] + T\sum_{k\in K}\lambda_k\|\sigma_k\|_{L^2}^2 \\ &\quad + C\sum_{k\in K}\lambda_k \mathbb{E}\int_0^T \langle u(s), \sigma_k\rangle^2 ds\end{aligned}$$

and the right-hand-side is bounded as in the statement of the theorem. Hence in particular $u \in C_{\mathcal{F}}([0,T];H)$. ■

2.1 Consequences

The message we get from this theorem is manifold.

- The solution has integrability properties in ω reflecting analogous properties assumed on the data.
- In the modeling of emergence of vortices developed in the previous section we have made a mistake: creating vortices from nothing we introduce energy into the system. Therefore we have to include an extra dissipation mechanism. There is a loss of energy due to the impact of the flow with the obstacle (which, let us remember, is not included into the boundary conditions); part of this energy is given back in the form of emerging vortices. We do not have a sufficiently good solution to this mistake, which then we leave as an open problem. A possible proposal is adding a friction term $-\lambda(x)u$

  $$du + (u\cdot\nabla u + \nabla p)\,dt = (\nu\Delta u - \lambda(x)u)\,dt + \sum_{k\in K}\sqrt{\lambda_k}\sigma_k dW_t^k$$

  with a friction coefficient possibly depending on x and localized near the boundary: in this way the Physical idea is that energy of large scales is subtracted near the boundary; and re-injected through the vortices σ_k. The energy balance is now

  $$\begin{aligned}&\mathbb{E}\left[\|u(t)\|_{L^2}^2\right] + 2\nu\int_0^t \mathbb{E}\|\nabla u(s)\|_{L^2}^2 ds + 2\mathbb{E}\int_0^t\int_D \lambda(x)|u(s,x)|^2 dxds \\ =\ &\mathbb{E}\left[\|u_0\|_{L^2}^2\right] + t\sum_{k\in K}\lambda_k\|\sigma_k\|_{L^2}^2.\end{aligned}$$

But we should be able to choose $\lambda(x)$ in such a way that

$$2\mathbb{E}\int_0^t\int_D \lambda(x)\,|u(s,x)|^2\,dxds \sim t\sum_{k\in K}\lambda_k\,\|\sigma_k\|_{L^2}^2\,.$$

We do not know how to reach this target.

- Assume $u(t)$ is a statistically stationary solution; this implies that $\mathbb{E}\,\|u(t)\|_{L^2}^2 = \mathbb{E}\,\|u_0\|_{L^2}^2$ and $\mathbb{E}\,\|\nabla u(s)\|_{L^2}^2$ is independent s, which then we denote by $\mathbb{E}\,\|\nabla u\|_{L^2}^2$. Then, stressing the dependence of u on ν,

$$\epsilon := \nu\mathbb{E}\,\|\nabla u_\nu\|_{L^2}^2\,ds = \frac{1}{2}\sum_{k\in K}\lambda_k\,\|\sigma_k\|_{L^2}^2\,.$$

  The dissipation ϵ of energy due to viscosity remains constant in the inviscid limit $\epsilon\to 0$ (it is a statement of K41 theory), if the energy injection is constant.

- We may use a small variant of the previous result to study state-dependent noise by iterations, see below.

2.2 Example of state-dependent noise

In Chapter 1 we have introduced a noise modeling the emergence of vortices at a boundary due to instability. However, when the fluid is at rest, certainly no vortex is created; similarly, we do not expect frequent creations if the velocity of the flow is very small. The rate of creation of vortices hence should depend on some feature of the flow itself. This doesn't mean that the model of the previous Chapter is useless: it is reasonable when the mean flow is roughly constant, and the rates τ^k should be taken appropriately with respect to the constant mean flow value.

When the state $u(t,\cdot)$ affects the rate of creation, we may use the concept of non-homogeneous Poisson process with random time-dependent rate: we introduce (corresponding to each k) an instantaneous rate $\lambda_k(u(t))$ depending on an average intensity of $u(t,\cdot)$, e.g.

$$\lambda_k(u(t)) = \chi^2\left(\frac{1}{|B(x_k,r)|}\int_{B(x_k,r)}|u(t,y)|\,dy\right)$$

where χ^2 is a nondecreasing non-negative function, equal to zero in zero and $r>0$ is a length scale relevant to the problem. Then we introduce the cumulative rate

$$\Lambda_k(t) = \int_0^t \lambda_k(u(s))\,ds$$

and finally we modify the Poisson process N_t^k by this rate, namely we consider the process

$$N^k_{\Lambda_k(t)}.$$

The case previously considered was simply

$$\lambda_k(u(t)) = \lambda_k, \qquad \Lambda_k(t) = \lambda_k, \qquad N^k_{\lambda_k t}.$$

The jump times of the noise in the equation will be the jump times of this processes, which are delayed or accelerated depending on the average intensity of $u(t)$:

$$\partial_t u + u \cdot \nabla u + \nabla p = \nu \Delta u + f + F(u) + \sum_{k \in K} \sigma_k \partial_t N^k_{\Lambda_k(t)} \tag{4}$$

or

$$\partial_t u + u \cdot \nabla u + \nabla p = \nu \Delta u + f + F(u) + \sum_{k \in K} \frac{1}{\sqrt{2}} \sigma_k \partial_t \left(N^{k,1}_{\Lambda_k(t)} - N^{k,2}_{\Lambda_k(t)} \right) \tag{5}$$

depending whether we assume that both vortices $\sigma_k(x)$ and $-\sigma_k(x)$ appear and are equally likely.

This is already a very interesting model which could deserve investigation. Otherwise, in the case of (5), we may rescale the noise as

$$\sum_{k \in K} \frac{1}{n\sqrt{2}} \sigma_k(x) \left(N^{k,1}_{n^2 \Lambda_k(t)} - N^{k,2}_{n^2 \Lambda_k(t)} \right). \tag{6}$$

Notice that, in order to increase the rate at time t, we have to use the instantaneous rate $n^2 \lambda_k(t)$, whence the expression $n^2 \Lambda_k(t)$ (instead of $\Lambda_k(n^2 t)$ which has a completely different and wrong meaning).

Recalling the convergence of rescaled Poisson processes to Brownian motion discussed in Chapter 1, it can be proved that the limit process of (6), in law, is

$$\sum_{k \in K} \sigma_k(x) B^k_{\Lambda_k(t)}$$

where B^k_t are independent Brownian motions. Then, by a deep theorem on martingales, there exists (possibly on a larger probability space) independent Brownian motions W^k_t such that, in law

$$B^k_{\Lambda_k(t)} = \int_0^t \sqrt{\lambda_k(u(s))} dW^k_s$$

(jointly in k). This result in undoubtedly advanced and not trivial even at the heuristic level but notice at least the analogy with the coefficients $\sqrt{\lambda_k}$ in the case of constant rate: when $\lambda_k(u(s)) = \lambda_k$, $\Lambda_k(t) = \lambda_k$, the previous identity reads

$$B^k_{\lambda_k t} = \int_0^t \sqrt{\lambda_k} dW^k_s = \sqrt{\lambda_k} W^k_t$$

and it is well known that $\lambda_k^{-1/2} B^k_{\lambda_k t}$ is a new Brownian motion.

The final equation is

$$\partial_t u + u \cdot \nabla u + \nabla p = \nu \Delta u + f + F(u) + \sum_{k\in K} \sigma_k \sqrt{\lambda_k(u)} \partial_t W_t^k.$$

We write it in the form

$$\partial_t u + u \cdot \nabla u + \nabla p = \nu \Delta u + f + F(u) + \sum_{k\in K} \sigma_k(u) \partial_t W_t^k \tag{7}$$

by introducing the maps $\sigma_k : H \to H$ given by

$$\sigma_k(u)(x) = \sigma_k(x) \sqrt{\lambda_k(u)}.$$

2.3 The Wong-Zakai corrector

Equations (4)-(5) are mathematically correct (whether they are physically relevant, it should be investigated more deeply). On the contrary, equation (7) requires a special choice of $F(u)$ to be the right one:

$$F(u) = \frac{1}{2} \sum_{k\in K} D\sigma_k(u) \sigma_k(u).$$

Here by $D\sigma_k(u)$ we mean the Frechét Jacobian of $\sigma_k(u)$, which is a linear bounded operator from H to H, under suitable assumptions, and $D\sigma_k(u)\sigma_k(u)$ is the application of the linear map $D\sigma_k(u)$ to the element $\sigma_k(u)$ of H. We do not know whether a full proof of this fact has been given and under which assumptions. We assume this is the correct result by heuristic extension of a known argument for finite dimensional equations. Let us explain it in the simple case of a one-dimensional equation.

Consider the one dimensional equation, with $\sigma(x) \geq \nu > 0$,

$$\frac{dX_t^\epsilon}{dt} = \sigma(X_t^\epsilon) \frac{dW_t^\epsilon}{dt}$$

where W_t^ϵ is an approximation of a Brownian motion W_t. It is an equation with separated variables. Then

$$\frac{\frac{dX_t^\epsilon}{dt}}{\sigma(X_t^\epsilon)} = \frac{dW_t^\epsilon}{dt}$$

$$\int_0^T \frac{\frac{dX_t^\epsilon}{dt}}{\sigma(X_t^\epsilon)} dt = \int_0^T \frac{dW_t^\epsilon}{dt} dt$$

$$\Phi(X_T^\epsilon) - \Phi(x_0) = W_T^\epsilon, \qquad \Phi'(x) = \frac{1}{\sigma(x)}$$

$$X_t^\epsilon = \Phi^{-1}(\Phi(x_0) + W_t^\epsilon)$$

Hence $X^{\epsilon}_{\cdot}$ converges weakly to $X_{\cdot}$ given by

$$X_t = \Phi^{-1}\left(\Phi\left(x_0\right) + W_t\right).$$

From Ito formula, since

$$\begin{aligned}
D\Phi^{-1}\left(x\right) &= \frac{1}{\Phi'\left(\Phi^{-1}\left(x\right)\right)} = \sigma\left(\Phi^{-1}\left(x\right)\right) \\
D^2\Phi^{-1}\left(x\right) &= D\left[\sigma\left(\Phi^{-1}\left(x\right)\right)\right] = \sigma'\left(\Phi^{-1}\left(x\right)\right) D\Phi^{-1}\left(x\right) \\
&= \sigma'\left(\Phi^{-1}\left(x\right)\right)\sigma\left(\Phi^{-1}\left(x\right)\right)
\end{aligned}$$

$$\begin{aligned}
dX_t &= \sigma\left(\Phi^{-1}\left(\Phi\left(x_0\right) + W_t\right)\right) dW_t + \frac{1}{2}\sigma'\left(\Phi^{-1}\left(\Phi\left(x_0\right) + W_t\right)\right)\sigma\left(\Phi^{-1}\left(\Phi\left(x_0\right) + W_t\right)\right) dt \\
&= \sigma\left(X_t\right) dW_t + \frac{1}{2}\sigma'\left(X_t\right)\sigma\left(X_t\right) dt.
\end{aligned}$$

We have found the corrector above.

Our conclusion, supported by the previous heuristic evidences, is that the right stochastic equations is

$$\partial_t u + u \cdot \nabla u + \nabla p = \nu \Delta u + f + \frac{1}{2}\sum_{k\in K} D\sigma_k\left(u\right)\sigma_k\left(u\right) + \sum_{k\in K}\sigma_k\left(u\right)\partial_t W_t^k.$$

Remark 5 *There is a notion of stochastic integral, different from the Itô one, called Stratonovich integral and denoted by* $\int_0^t \sigma_k\left(u\left(s\right)\right) \circ dW_s^k$*, such that*

$$\int_0^t \sigma_k\left(u\left(s\right)\right) \circ dW_s^k = \int_0^t \sigma_k\left(u\left(s\right)\right) dW_s^k + \frac{1}{2}\int_0^t D\sigma_k\left(u\left(s\right)\right)\sigma_k\left(u\left(s\right)\right) ds$$

when u *solves equation above. Therefore, with such notion, the equation has the form*

$$\partial_t u + u \cdot \nabla u + \nabla p = \nu \Delta u + f + \sum_{k\in K}\sigma_k\left(u\right) \circ \partial_t W_t^k.$$

3 2D Stochastic Navier-Stokes equations

Consider now the equations

$$\begin{aligned}
\partial_t u + u \cdot \nabla u + \nabla p &= \nu \Delta u + f + F\left(u\right) + \sum_{k\in K}\sigma_k\left(u\right)\partial_t W_t^k \qquad (8) \\
\operatorname{div} u &= 0
\end{aligned}$$

with

$$\begin{aligned} u|_{\partial D} &= 0 \\ u(0) &= u_0. \end{aligned}$$

Assume

$$\begin{aligned} F &\in Lip(H,H) \\ \sigma_k &\in Lip(H,H) \cap C(H, D(A)), \text{ bounded in } H, \qquad k \in K. \end{aligned}$$

With some additional elements of stochastic analysis (Itô formula for $\|u(t)\|_{L^2}^p$ and Burkholder-Davis-Gundy inequality) one can drop the assumption that σ_k are bounded, so it is made here only for simplicity of exposition. The assumption $C(H, D(A))$ is also made just for simplicity, but it is clear from the estimates below that it is absolutely unessential.

Definition 6 *Given $u_0 \in H$ and $f \in L^2(0,T;V')$, we say that*

$$u \in C_{\mathcal{F}}([0,T];H) \cap L^2_{\mathcal{F}}(0,T;V)$$

is a weak solution of equation (8) if

$$\begin{aligned} &\langle u(t), \phi\rangle - \int_0^t b(u(s), \phi, u(s))\, ds \\ = &\langle u_0, \phi\rangle + \int_0^t \langle u(s), A\phi\rangle\, ds + \int_0^t \langle f(s), \phi\rangle\, ds \\ &+ \int_0^t \langle F(u(s)), \phi\rangle\, ds + \sum_{k\in K} \int_0^t \langle \sigma_k(u(s)), \phi\rangle\, dW_s^k \end{aligned}$$

for every $\phi \in D(A)$.

Theorem 7 *For every $u_0 \in L^2_{\mathcal{F}_0}(\Omega, H)$ and $f \in L^2_{\mathcal{F}}(0,T;V')$, there exists a unique weak solution of equation (8). It satisfies*

$$\begin{aligned} &\mathbb{E}\left[\|u(t)\|_{L^2}^2\right] + 2\nu\mathbb{E}\int_0^t \|\nabla u(s)\|_{L^2}^2\, ds \\ = &\mathbb{E}\left[\|u_0\|_{L^2}^2\right] + 2\mathbb{E}\int_0^t \langle u(s), f(s) + F(u(s))\rangle\, ds \\ &+ \sum_{k\in K} \mathbb{E}\int_0^t \|\sigma_k(u(s))\|_{L^2}^2\, ds. \end{aligned}$$

Remark 8 *We have taken deterministic data for simplicity, but extensions to random data are possibile. Uniqueness holds under the natural assumption* $u_0 \in L^2_{\mathcal{F}_0}(\Omega, H)$ *and* $f \in L^2_{\mathcal{F}}(0, T; V')$. *Existence requires the same assumption plus the additional integrability*

$$\mathbb{E}\left[\|u_0\|_H^r\right] < \infty, \qquad \mathbb{E}\int_0^t \|f(s)\|_{V'}^r \, ds < \infty \tag{9}$$

for some $r > 2$.

3.1 Proof of uniqueness

Let $u^{(i)}$ be two solutions. Then $w = u^{(1)} - u^{(2)}$ satisfies

$$\begin{aligned} & \langle w(t), \phi\rangle - \int_0^t \left(b\left(u^{(1)}, \phi, u^{(1)}\right) - b\left(u^{(2)}, \phi, u^{(2)}\right)\right)(s)\, ds \\ = & \int_0^t \langle w(s), A\phi\rangle\, ds + \int_0^t \left\langle F\left(u^{(1)}(s)\right) - F\left(u^{(2)}(s)\right), \phi\right\rangle ds \\ & + \sum_{k\in K} \int_0^t \left\langle \sigma_k\left(u^{(1)}(s)\right) - \sigma_k\left(u^{(2)}(s)\right), \phi\right\rangle dW_s^k \end{aligned}$$

and since

$$\begin{aligned} & b\left(u^{(1)}, \phi, u^{(1)}\right) - b\left(u^{(2)}, \phi, u^{(2)}\right) - b(w, \phi, w) \\ = & \; b\left(u^{(2)}, \phi, w\right) + b\left(w, \phi, u^{(2)}\right) \end{aligned}$$

we get

$$\begin{aligned} & \langle w(t), \phi\rangle - \int_0^t (b(w(s), \phi, w(s)))\, ds \\ = & \int_0^t \langle w(s), A\phi\rangle\, ds + \int_0^t \left\langle F\left(u^{(1)}(s)\right) - F\left(u^{(2)}(s)\right), \phi\right\rangle ds \\ & + \sum_k \int_0^t \left\langle \sigma_k\left(u^{(1)}(s)\right) - \sigma_k\left(u^{(2)}(s)\right), \phi\right\rangle dW_s^k \\ & - \int_0^t \left(b\left(u^{(2)}, \phi, w\right) + b\left(w, \phi, u^{(2)}\right)\right)(s)\, ds. \end{aligned}$$

We need the Itô formula to continue; it can be proved similarly to Theorem 4. It gives us

$$\begin{aligned}\|w(t)\|_H^2 + 2\nu \int_0^t \|\nabla w(s)\|_H^2 \, ds &= 2\int_0^t \left\langle F\left(u^{(1)}(s)\right) - F\left(u^{(2)}(s)\right), w(s)\right\rangle ds \\ &\quad -2\int_0^t \left(b\left(u^{(2)}, w, w\right) + b\left(w, w, u^{(2)}\right)\right)(s)\, ds \\ &\quad + \sum_{k\in K}\int_0^t \left\|\sigma_k\left(u^{(1)}(s)\right) - \sigma_k\left(u^{(2)}(s)\right)\right\|_{L^2}^2 ds \\ &\quad + M_t\end{aligned}$$

where

$$M_t := \sum_k \int_0^t \left\langle \sigma_k\left(u^{(1)}(s)\right) - \sigma_k\left(u^{(2)}(s)\right), w(s)\right\rangle dW_s^k.$$

Therefore, if L_F and L_k are the Lipschitz constants of F and σ_k respectively, using estimates of Chapter 1 we get

$$\begin{aligned}\|w(t)\|_H^2 + \nu \int_0^t \|\nabla w(s)\|_H^2 \, ds &\leq \left(2L_F + \sum_{k\in K} L_k^2\right)\int_0^t \|w(s)\|_H^2 \, ds \\ &\quad + C\int_0^t \|w(s)\|_H^2 \left(1 + \left\|u^{(2)}(s)\right\|_{\mathbb{L}^4}^2\right) ds \\ &\quad + M_t.\end{aligned}$$

We need now a very interesting trick that we have learned from Bjorn Schmalfuss: introduced

$$\rho_t = \exp\left(-C\int_0^t \left(1 + \left\|u^{(2)}(s)\right\|_{\mathbb{L}^4}^2\right) ds\right)$$

we have, from Itô formula again,

$$\|w(t)\|_H^2 \rho_t + \nu \int_0^t \|\nabla w(s)\|_H^2 \rho_s ds \leq \left(2L_F + \sum_{k\in K} L_k^2\right)\int_0^t \|w(s)\|_H^2 \rho_s ds + \widetilde{M}_t$$

where

$$\widetilde{M}_t := \sum_{k\in K}\int_0^t \left\langle \sigma_k\left(u^{(1)}(s)\right) - \sigma_k\left(u^{(2)}(s)\right), w(s)\right\rangle \rho_s dW_s^k.$$

Omitting the necessary localization argument entirely similar to the one used in Theorem 4, we get

$$\begin{aligned}&\mathbb{E}\left[\|w(t)\|_H^2 \rho_t\right] + \nu\mathbb{E}\int_0^t \|\nabla w(s)\|_H^2 \rho_s ds \\ \leq &\left(2L_F + \sum_{k\in K} L_k^2\right)\int_0^t \mathbb{E}\left[\|w(s)\|_H^2 \rho_s\right] ds\end{aligned}$$

which leads to $\mathbb{E}\left[\|w(t)\|_H^2 \rho_t\right] = 0$ by Gronwall lemma. But, thanks to the regularity of $u^{(2)}$, $\mathbb{P}(\rho_t > 0) = 1$. Hence $\mathbb{P}(w(t) = 0) = 1$. Since this is true for all t, the processes $u^{(1)}$ and $u^{(2)}$ are modifications; but they are continuous, hence they are indistinguishable.

4 Proof of existence

4.1 Introduction

Existence for differential equations is a wide subject with many ideas. More or less, all methods consist in the construction of a sequence, based on some approximation or iteration method which allows to define the sequence by means of easier equations than the one object of investigation. Then one has to prove convergence in a topology which allows one to pass to the limit in the approximate equations. Linear terms pass to the limit under very weak convergences, hence the demanding part for the limit step are the nonlinear terms. When they have suitable monotonicity properties, again weak convergence is sufficient, but the Navier-Stokes nonlinearity does not have such properties. Strong convergence in a topology like H is needed. Weak convergence does not suffice to take the limit in a quadratic expression; the weak limit of the square is not the square of the weak limit, in general.

We have insisted on this classification of ideas because the existence of weakly convergent subsequences of an approximating scheme is an excellent property also in the stochastic case, it applies for instance to spaces like $L^2(\Omega, B)$ with a Banach space B. But the existence of *strongly convergent subsequences* of an approximating scheme is very demanding, in the stochastic case. And for the Navier-Stokes equations we are faced with this demanding problem.

Essentially there are two ways to get strong convergence: one is related to contraction principle arguments and consists in the proof of the Cauchy property of the sequence, in the strong topology, usually in expected value. This kind of argument is not easy to be implemented for the Navier-Stokes equations. In the deterministic setting we have seen an example of this technique in Chapter 1: for the auxiliary Navier-Stokes equations we have constructed a sequence (v_n) and proved it was Cauchy. In the stochastic case, performing similar proofs is very difficult because of the problem of *closure of moments*: we have to take expected values but the nonlinearity increases the order of the moment. Inspection into the proof of Chapter 1 reveals we have used uniform bounds on the iterates to close a certain inequality in the proof of the Cauchy property; in the deterministic case such bounds are deterministic; in the stochastic case they are in expected value and thus not easily applicable.

The alternative strategy to have strong convergence of subsequences is by compactness theorems. However, here there is a structural problem: compactness in spaces like $L^2(\Omega, B)$ is essentially impossible to be proven (except for criteria based on Malliavin calculus, which however did not prove to be competitive, until now). Thus one goes to compactness of the laws, because compactness in spaces of measures is very well characterized.

But then the problem becomes that we have only subsequences of laws, which converge in strong topologies. Namely, it is not strong convergence of the original stochastic processes, only of their laws. How to identify a limit stochastic process and pass to the limit in the equations?

Here there are several strategies, each one with advantages depending on a certain feature of the problem; or, if not advantages, it is the only one we can use.

- When we can prove the so-called pathwise uniqueness, as above in the 2D case, there is a brilliant criterion of Gyongy and Krylov which proves the convergence in probability of the approximating sequence of stochastic processes, hence upgrading the pure convergence in law. We shall explain this below.

- Alternative to this method, when pathwise uniqueness is known, is proving weak convergence of the laws, construct a solution on an auxiliary probability space and then use a theorem of Yamada-Watanabe type (which requires pathwise uniqueness) to prove that a solution on the original probability space exists. This strategy looks longer than the previous one, hence we prefer to describe Gyongy-Krylov approach.

- When pathwise uniqueness is not known or it is false, there is no way to upgrade the weak convergence of laws to some kind of stronger convergence of the processes. In this case Skorohod representation theorem allows one to reformulate the approximating sequence on a new, auxiliary probability space, where it converges also almost surely, not only in law. Then one can pass to the limit. But the limit process lives in an auxiliary probability space. This is the same strategy used in the previous item, but not followed by a Yamada-Watanabe step. Hence the final result is just existence on an auxiliary space.

- For special noise, like the additive one, when pathwise uniqueness is not known, there is a trick to pass to the limit in the equation using just the weak convergence of the laws, without performing the Skorohod representation theorem step. The limit law is a solution of the equation, in a suitable sense. We shall describe this procedure below. It applies for instance to the 3D Navier-Stokes equations with additive noise.

One may add several comments to the previous list, related for instance to the concept of martingale solutions, but we limit ourselves to the previous discussion and show some of the computations for the first and the last item.

4.2 Gyongy-Krylov convergence criterium

If (E, d) is a metric space we denote by $\left(E^2, d^2\right)$ the product space with the metric $d^2\left((x,y),(x',y')\right) = d(x,y) + d(x',y')$; we understand that on every one of these metric spaces the σ-field is the Borel one; and we denote by D the diagonal:

$$D = \left\{(x,x) \in E^2; x \in E\right\}.$$

Lemma 9 *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to a complete separable metric space (E,d). Assume that, for every pair of subsequences $((n_1(k), n_2(k)))_{k\in\mathbb{N}}$, with $n_1(k) \geq n_2(k)$ for every $k \in \mathbb{N}$, there is a subsequence $(k(h))_{h\in\mathbb{N}}$ such that the random variables $\left(X_{n_1(k(h))}, X_{n_2(k(h))}\right)_{h\in\mathbb{N}}$ from $(\Omega, \mathcal{F}, \mathbb{P})$ to $\left(E^2, d^2\right)$ converge in law to a measure μ on E^2 such that $\mu(D) = 1$. Then there exists a random variable X from $(\Omega, \mathcal{F}, \mathbb{P})$ to (E,d) such that X_n converges to X in probability.*

Proof. It is sufficient to prove that $(X_n)_{n\in\mathbb{N}}$ is Cauchy in probability: given $\epsilon > 0$ we have to find n_0 such that for all $n, m > n_0$ one has

$$\mathbb{P}(d(X_n, X_m) \geq \epsilon) < \epsilon.$$

Let us prove this by contradiction: we assume that there exists $\epsilon_0 > 0$ such that for every k there are $n_1(k) \geq n_2(k) > k$ such that

$$\mathbb{P}\left(d\left(X_{n_1(k)}, X_{n_2(k)}\right) \geq \epsilon_0\right) \geq \epsilon_0.$$

We may perfection the construction in order to have that $n_1(k), n_2(k)$ are strictly increasing, hence they are subsequences. But by assumption there exists a subsequence $k(h)$ such that $\left(X_{n_1(k(h))}, X_{n_2(k(h))}\right)$ converges in law to μ, hence its probability of taking values in a closed set is upper semicontinuous:

$$\mu((x,y) : d(x,y) \geq \epsilon_0) \geq \limsup \mathbb{P}\left(d\left(X_{n_1(k(h))}, X_{n_2(k(h))}\right) \geq \epsilon_0\right) \geq \epsilon_0.$$

This inequality is incompatible with $\mu(D^c) = 0$, hence we have reached a contradiction. ■

4.3 Compactness criteria

4.3.1 Deterministic Ascoli-Arzelà theorem

A version of Ascoli-Arzelà theorem claims that, given two Banach spaces $X \overset{\text{compact}}{\subset} Y$, a family $F \subset C([0,T]; Y)$ with the following two properties is relatively compact in $C([0,T]; Y)$:

i) $\{f(t); f \in F\}$ is bounded in X

ii) F is uniformly equicontinuous in $C([0,T]; Y)$, namely for every $\epsilon > 0$ there is $\delta > 0$ such that $\|f(t) - f(s)\|_Y \leq \epsilon$ for every $f \in F$ and $t, s \in [0,T]$ such that $|t-s| \leq \delta$.

In particular:

Proposition 10 *If $p > 1$ and $X \overset{compact}{\subset} Y$, then*

$$W^{1,p}(0,T;X) \overset{compact}{\subset} C([0,T];Y).$$

Indeed, if $F \subset W^{1,p}(0,T;X)$ is bounded, and $t \in \left[\frac{T}{2}, T\right]$ (similarly for $t \in \left[0, \frac{T}{2}\right]$)

$$f(t) - f(s) = \int_s^t f'(r)\, dr$$

$$f(t) = \frac{2}{T}\int_0^{T/2} f(s)\, ds + \frac{2}{T}\int_0^{T/2}\int_s^t f'(r)\, dr ds$$

$$\begin{aligned}\|f(t)\|_X &\leq \frac{2}{T}\int_0^{T/2} \|f(s)\|_X\, ds + \frac{2}{T}\int_0^{T/2}\int_s^t \left\|f'(r)\right\|_X dr ds \\ &\leq \frac{2}{T}\|f\|_{L^1(0,T;X)} + \left\|f'\right\|_{L^1(0,T;X)} \leq C\end{aligned}$$

and

$$\|f(t) - f(s)\|_X \leq \int_s^t \left\|f'(r)\right\|_X dr \leq \left\|f'\right\|_{L^p(0,T;X)} |t-s|^q \leq C|t-s|^q$$

where $\frac{1}{p} + \frac{1}{q} = 1$ and the constant C is independent of $f \in F$. So F satisfies the assumptions of Ascoli-Arzelà theorem.

4.3.2 Deterministic Aubin-Lions type theorems

Given two Banach spaces $X \subset Y$, we say that the embedding $X \subset Y$ is compact if bounded sets of X are relatively compact in Y.

Theorem 11 *Let $X \subset Y \subset Z$ be three Banach spaces, with continuous dense embeddings. Assume that the embedding $X \subset Y$ is compact. Let $p \in [1, \infty)$ be given. Then the embedding*

$$L^p(0,T;X) \cap W^{1,1}(0,T;Z) \subset L^p(0,T;Y)$$

is compact.

Remark 12 *The previous theorem, when applied to functions spaces $X \subset Y \subset Z$, treats the problem of compactness of functions of space-time. Heuristically, one needs a condition of compactness for the space variable and one for the time variable and, a priori, one could expect the need of some sort of joint compactness in the two variables. By Ascoli-Arzelà theorem, the space or real-valued functions $W^{1,2}(0,T;\mathbb{R})$ is compactly embedded into $L^2(0,T;\mathbb{R})$. The remarkable feature of the previous theorem is that the compactness in the time variable does not require a simultaneous compactness in the space variable: the space Z can be much larger than Y. Said differently, the two compactness requirements, in space and time, are quite decoupled.*

Remark 13 *The consequence in examples is that the only key assumption turns out to be $L^p(0,T;X)$, the other being a technical consequence based on the differential equation.*

Remark 14 *Assume $p > 1$ and also assume the bound is in $W^{1,r}(0,T;Z)$ with $r > 1$. The previous result means that, if we have a sequence of functions (u_n) (usually solutions of an approximate equation) such that*

$$\int_0^T \|u_n(t)\|_X^p \, dt + \int_0^T \left\| \frac{du_n(t)}{dt} \right\|_Z^r dt \le C$$

then there exists a subsequence (u_{n_k}) and a function $u \in L^p(0,T;Y)$ such that

$$\lim_{n\to\infty} \int_0^T \|u_{n_k}(t) - u(t)\|_Y^p \, dt = 0.$$

Moreover, $u \in L^p(0,T;X) \cap W^{1,r}(0,T;Z)$ and (u_{n_k}) can be chosen so that it converges weakly to u in $L^p(0,T;X)$ and in $W^{1,r}(0,T;Z)$ (it is here that we use $p, r > 1$). The weak convergence in these topologies is a consequence of general theory of reflexive Banach spaces; that it can be done for a unique subsequence is easy; that the limit in the strong topology of $L^p(0,T;Y)$ and weak topologies of $L^p(0,T;X)$ and $W^{1,r}(0,T;Z)$ is the same function u, it requires some arguments that we omit (for instance: weak convergence in $L^p(0,T;X)$ implies weak convergence in $L^p(0,T;Y)$, hence the weak limit in these topologies is the same as the strong limit in $L^p(0,T;Y)$, by uniqueness between weak and strong limit in $L^p(0,T;Y)$). Moreover, in most examples we shall prove also a bound of the form

$$\sup_{t\in[0,T]} \|u_n(t)\|_Y \le C.$$

By the same arguments, one may have that (u_{n_k}) converges also weak-star to u in $L^\infty(0,T;Y)$. Finally, If $Y \overset{compact}{\subset} Z$, by Proposition 10 we may also add strong convergence of (u_{n_k}) to u in $C([0,T];Z)$.

Essential for the stochastic case is the following generalization (see Simon [**?**], Corollary 5):

Theorem 15 *If $\alpha r > 1 - \frac{r}{p}$ $(p, r \ge 1)$ then*

$$L^p(0,T;X) \cap W^{\alpha,r}(0,T;Z) \overset{compact}{\subset} L^p(0,T;Y)$$

Here $\alpha \in (0,1)$ and $W^{\alpha,r}(0,T;Z)$ is the space of functions $f \in L^r(0,T;Z)$ such that

$$\int_0^T \int_0^T \frac{\|f(t) - f(s)\|_Z^r}{|t-s|^{1+\alpha r}} ds dt < \infty.$$

Recall also that $W^{\alpha,r}(0,T;Z) \subset C([0,T];Z)$ if $\alpha r > 1$. The reason for asking this generalization is that we do not have true time derivatives in the stochastic case, but we have fractional time regularity.

The property of continuity in time in Y of solutions sometimes follows a posteriori, from the (S)PDE. Alternatively, we may try to prove convergence of the approximating scheme in the uniform topology. Obviously Ascoli-Arzelà theorem provides uniform convergence but the assumptions are too difficult to be checked in (S)PDEs like those of fluid mechanics (let us remark, however, that Ascoli-Arzelà theorem is at the foundation of most proofs of the compactness results illustrated here). To this purpose we may use the following result of Simon [**?**], Corollary 9:

Theorem 16 *Assume in addition ($\theta \in (0,1)$)*

$$\|v\|_Y \leq C \|v\|_X^{1-\theta} \|v\|_Z^{\theta} \qquad \theta \in (0,1)$$

$$\alpha r > 1 \text{ and } p > \frac{1-\theta}{\theta} \frac{r}{\alpha r - 1} \qquad (p, r \geq 1).$$

Then

$$L^p(0,T;X) \cap W^{\alpha,r}(0,T;Z) \overset{compact}{\subset} C([0,T];Y).$$

4.3.3 Stochastic theory

Consider now a differential equation where the solution depends also on a random parameter,

$$u = u(\omega, t, x).$$

The principle that compactness can be investigated separately in the three arguments, in principle, could still hold. However, the obstacle is that compactness in the random parameter ω is not an easy matter. The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is always infinite dimensional in our examples and compactness criteria in $L^p(\Omega)$ are not natural (although something can be done by weighted Sobolev spaces and Malliavin calculus, when $(\Omega, \mathcal{F}, \mathbb{P})$ has a special structure).

The natural approach is to consider the laws of the random objects and apply compactness arguments to these laws. It is easier due to the following basic theorem. Let (X, d) be a complete metric space and $\mathcal{B}$ the Borel σ-field. Recall we say that a family $\mathcal{G}$ of probability measures on $(X, \mathcal{B})$ is *tight* if for every $\epsilon > 0$ there is a compact set $K \subset X$ such that

$$\mu(K) \geq 1 - \epsilon$$

for all $\mu \in \mathcal{G}$.

Theorem 17 (Prohorov) *A family $\mathcal{G}$ of probability measures on $(X, \mathcal{B})$ is tight if and only if it is relatively compact.*

Corollary 18 *Assume* (u_N) *is a sequence of random functions from* $(\Omega, \mathcal{F}, \mathbb{P})$ *to* $L^p(0,T;Y)$. *Assume* $\alpha r > 1 - \frac{r}{p}$ *and that for every* $\epsilon > 0$ *there are* $R_1, R_2 > 0$ *such that*

$$\begin{aligned} \mathbb{P}\left(\|u_N\|_{L^p(0,T;X)} \geq R_1\right) &\leq \epsilon \\ \mathbb{P}\left(\|u_N\|_{W^{\alpha,r}(0,T;Z)} \geq R_1\right) &\leq \epsilon. \end{aligned}$$

Then there exists a subsequence (u_{N_k}) *which converges in law, in the strong topology of* $L^p(0,T;Y)$, *to a random function* $\widetilde{u}$ *from a probability space* $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ *to* $L^p(0,T;Y)$. *Moreover, if* $p, r > 1$, *we may chose* (u_{N_k}) *so that* $\widetilde{u}$ *takes also values in* $L^p(0,T;X)$ *and* $W^{\alpha,r}(0,T;Z)$.

If u_N *are* $(\mathcal{F}_t)$*-progressively measurable, there exists* $\left(\widetilde{\mathcal{F}}_t\right)$ *such that* $\widetilde{u}$ *is* $\left(\widetilde{\mathcal{F}}_t\right)$*-progressively measurable.*

Recall that the convergence in law stated above means

$$\lim_{k\to\infty} \mathbb{E}\left[\Phi\left(u_{N_k}\right)\right] = \widetilde{\mathbb{E}}\left[\Phi\left(\widetilde{u}\right)\right]$$

for every bounded continuous function $\Phi : L^p(0,T;Y) \to \mathbb{R}$. Here $\mathbb{E}$ and $\widetilde{\mathbb{E}}$ are the expected values on $(\Omega, \mathcal{F}, \mathbb{P})$ and $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ respectively.

Remark 19 *Sufficient conditions for the applicability of the Corollary are uniform in* N *estimates of the form*

$$\begin{aligned} \mathbb{E}\left[\|u_N\|_{L^p(0,T;X)}\right] &\leq C \\ \mathbb{E}\left[\|u_N\|_{W^{\alpha,r}(0,T;Z)}\right] &\leq C. \end{aligned}$$

Indeed, by Markov inequality,

$$\mathbb{P}\left(\|u_N\|_{L^p(0,T;X)} \geq R_1\right) \leq \frac{C}{R_1}$$

and similarly for the second inequality, hence given $\epsilon > 0$ *we can find* $R_1, R_2 > 0$ *with the required properties.*

Remark 20 *The consequence of the peculiar feature of the previous Corollary that the process* $\widetilde{u}$ *may be defined on a new probability space* $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ *is the emergence of the concept of "weak solution in the probabilistic sense". This means that the probability space over which we find a solution is not necessarily prescribed a priori. If we are only interested in statistical properties, this is not bad, but sometimes for special investigation it is very restrictive.*

4.4 Application to Galerkin approximations: 2D case

4.4.1 Estimates and compactness

Step 1 (preparation). Let us use the definitions introduced in the proof of Theorem 4: (e_i) is a complete orthonormal system in H made of eigenvectors of A, with eigenvalues $(-\lambda_i)$, H_n and π_n are consequently defined, and we introduce the bilinear operator $B_n : H_n \times H_n \to H_n$ definend as

$$B_n(u,v) = \pi_n P(u \cdot \nabla v)$$

(we omit the verification that $u, v \in H_n$ imply $u \cdot \nabla v \in H$, so that P is well defined on $u \cdot \nabla v$). Then we consider the finite dimensional equation

$$du_n = Au_n dt - B_n(u_n, u_n)\,dt + f_n + F_n(u_n) + \sum_k \sigma_k^n(u_n)\,dW_t^k$$

where $f_n = \pi_n f$, $F_n(u) = \pi_n F(u)$, $\sigma_k^n(u_n) = \pi_n \sigma_k(u_n)$; with initial condition $u_0^n = \pi_n u_0$. It is easy to check that

$$\langle B_n(u_n, u_n), u_n \rangle = 0.$$

Step 2 (estimates in square norms). Therefore, from Itô formula (in finite dimensions) we get

$$\begin{aligned}\|u_n(t)\|_H^2 + 2\nu \int_0^t \|\nabla u_n(s)\|_H^2\,ds &= 2\int_0^t \langle f_n(s) + F(u_n(s)), u_n(s)\rangle\,ds \\ &\quad + \sum_{k\in K}\int_0^t \|\sigma_k^n(u_n(s))\|_{L^2}^2\,ds + M_t^n\end{aligned}$$

where

$$M_t^n = 2\sum_{k\in K}\int_0^t \langle \sigma_k^n(u_n(s)), u_n(s)\rangle\,dW_t^k.$$

After having seen above various proofs, it is a simple exercise to deduce (see also Step 3 below)

$$\begin{aligned}\mathbb{E}\int_0^T \|u_n(s)\|_V^2\,ds &\leq C \\ \mathbb{E}\left[\sup_{t\in[0,T]} \|u_n(t)\|_H^2\right] &\leq C.\end{aligned} \tag{10}$$

Then we investigate the $W^{\alpha,p}(0,T;V')$ norm of u_n. In a sense, this is the most technical part but the reader will recognize that the key properties are (10), the rest of the proof are technicalities. For $s < t$

$$\begin{aligned}\|u_n(t) - u_n(s)\|_{V'} &\leq \int_s^t \|Au_n(r)\|_{V'}\,dr + \int_s^t \|B_n(u_n, u_n)(r)\|_{V'}\,dr \\ &\quad + \int_s^t \|f_n(r) + F_n(u_n)(r)\|_{V'}\,dr + \left\|\sum_k \int_s^t \sigma_k^n(u_n(r))\,dW_r^k\right\|_{V'}.\end{aligned}$$

We have

$$\mathbb{E}\int_s^t \|Au_n(r)\|_{V'}\,dr \leq \sqrt{t-s}\left(\mathbb{E}\int_s^t \|Au_n(r)\|_{V'}^2\,dr\right)^{1/2} \leq C\sqrt{t-s}$$

by (10), and similarly

$$\mathbb{E}\int_s^t \|f_n(r) + F_n(u_n)(r)\|_{V'}\,dr \leq C\sqrt{t-s}.$$

Moreover,

$$\begin{aligned}\mathbb{E}\left\|\sum_k \int_s^t \sigma_k^n(u_n(r))\,dW_r^k\right\|_{V'} &\leq \left(\mathbb{E}\left[\left\|\sum_k \int_s^t \sigma_k^n(u_n(r))\,dW_r^k\right\|_{V'}^2\right]\right)^{1/2} \\ &= \left(\mathbb{E}\sum_k \int_s^t \|\sigma_k^n(u_n(r))\|_{V'}^2\,dr\right)^{1/2} \leq C\sqrt{t-s}\end{aligned}$$

because we assume σ_k^n bounded. Finally, from the usual inequalities,

$$\begin{aligned}\int_s^t \|B_n(u_n, u_n)(r)\|_{V'}\,dr &\leq C\int_s^t \|u_n(r)\|_H \|u_n(r)\|_V\,dr \\ &\leq C\sup_{r\in[0,T]} \|u_n(r)\|_H \int_s^t \|u_n(r)\|_V\,dr\end{aligned}$$

hence

$$\begin{aligned}\mathbb{E}\int_s^t \|B_n(u_n, u_n)(r)\|_{V'}\,dr &\leq C\mathbb{E}\left[\sup_{r\in[0,T]} \|u_n(r)\|_H^2\right]^{1/2} \mathbb{E}\left[\left(\int_s^t \|u_n(r)\|_V\,dr\right)^2\right]^{1/2} \\ &\leq C\sqrt{t-s}.\end{aligned}$$

Putting together all these pieces,

$$\mathbb{E}\|u_n(t) - u_n(s)\|_{V'} \leq C\sqrt{t-s}$$

which implies

$$\mathbb{E}\int_0^T\int_0^T \frac{\|u_n(t)-u_n(s)\|_{V'}}{|t-s|^{1+\alpha}}dsdt \leq \int_0^T\int_0^T \frac{C}{|t-s|^{\frac{1}{2}+\alpha}}dsdt =: C<\infty$$

if $\alpha\in\left(0,\frac{1}{2}\right)$. The condition $\alpha r>1-\frac{r}{p}$ of Theorem 15 is fulfilled for $1-\frac{1}{p}<\frac{1}{2}$, namely for $p<2$. This result is not so good for the sequel: when passing to the limit in the nonlinear term we have

$$\int_0^t \langle B_n(u_n(s),u_n(s)),\phi\rangle ds = -\int_0^t b(u_n(s),\pi_n\phi,u_n(s))\,ds$$

so, taking $\phi\in D(A)\subset C_b(D)$, it is sufficient to have strong convergence of u_n in $L^2(0,T;H)$, but not in $L^p(0,T;H)$ with $p<2$. Perhaps there are arguments to overcome this difficulty thanks to the uniform in time bound of estimate (10), but it is interesting to show how to upgrade the integrability of solutions and thus let us develop this in the next step.

Step 3 (estimates in L^r). Take $r>2$. Assume

$$\mathbb{E}\left[\|u_0\|_H^r\right]<\infty, \qquad \mathbb{E}\int_0^t \|f(s)\|_{V'}^r\,ds<\infty.$$

Consider the function

$$f(x)=\|x\|^r$$

for $x\in\mathbb{R}^n$. We have, for $x\neq 0$,

$$\partial_i f(x) = r\|x\|^{r-1}\frac{x_i}{\|x\|} = r\|x\|^{r-2}x_i$$

$$\begin{aligned}\partial_j\partial_i f(x) &= rx_i\partial_j\|x\|^{r-2}+r\|x\|^{r-2}\delta_{ij}\\ &= r(r-2)\|x\|^{r-4}x_ix_j+r\|x\|^{r-2}\delta_{ij}\end{aligned}$$

and we may include $x=0$ for $r\geq 4$. Treating rigorously the case $r\in(2,4)$ requires some more details that we omit. Then from Itô formula we have

$$\begin{aligned}d\|u_n(t)\|_H^r &= r\|u_n(t)\|^{r-2}\langle u_n(t),du_n(t)\rangle\\ &\quad+\frac{1}{2}r(r-2)\sum_{k\in K}\|u_n(t)\|^{r-4}\langle u_n(t),\sigma_k^n(u_n(t))\rangle^2dt\\ &\quad+\frac{1}{2}r\|u_n(t)\|^{r-2}\sum_{k\in K}\|\sigma_k^n(u_n(t))\|_{L^2}^2dt\end{aligned}$$

hence

$$\begin{aligned} & d\left\|u_n(t)\right\|_H^r + r\nu \left\|u_n(t)\right\|^{r-2} \left\|\nabla u_n(t)\right\|_{L^2}^2 \\ \leq\ & r\left\|u_n(t)\right\|^{r-2} \left\langle u_n(t), f_n + F_n(u_n)\right\rangle dt + M_t^{n,r} \\ & +\frac{1}{2} r(r-1)\left\|u_n(t)\right\|^{r-2} \sum_k \left\|\sigma_k^n(u_n(t))\right\|_{L^2}^2 dt \end{aligned}$$

where

$$M_t^{n,r} = r \sum_{k\in K} \int_0^t \left\|u_n(s)\right\|^{r-2} \left\langle \sigma_k^n(u_n(s)), u_n(s)\right\rangle dW_t^k.$$

From the usual localization argument,

$$\begin{aligned} & \mathbb{E}\left[\left\|u_n(t)\right\|_H^r\right] + r\nu \int_0^t \left\|u_n(s)\right\|^{r-2} \left\|\nabla u_n(s)\right\|_{L^2}^2 ds \\ \leq\ & C_r \mathbb{E}\int_0^t \left(\left\|u_n(s)\right\|_H^r + 1\right) ds + C_r \mathbb{E}\int_0^t \left\|u_n(s)\right\|_H^{r-2} \left\|f(s)\right\|_{V'}^2 ds \\ & +\nu \int_0^t \left\|u_n(s)\right\|_H^{r-2} \left\|u_n(s)\right\|_V^2 ds. \end{aligned}$$

We need, from $ab \leq c_r \left(a^{\frac{r}{r-2}} + b^{\frac{r}{2}}\right)$ $\left(\frac{r-2}{r} + \frac{2}{r} = 1\right)$

$$\mathbb{E}\int_0^t \left\|u_n(s)\right\|_H^{r-2} \left\|f(s)\right\|_{V'}^2 ds \leq \mathbb{E}\int_0^t \left\|u_n(s)\right\|_H^r ds + \mathbb{E}\int_0^t \left\|f(s)\right\|_{V'}^r ds$$

hence the additional assumption on f. From Gronwall lemma,

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left\|u_n(t)\right\|_H^r\right] \leq C.$$

Using this preliminary estimate and Burkholder-Davis-Gundy inequality (we omit the details) we get

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left\|u_n(t)\right\|_H^r\right] \leq C. \tag{11}$$

Repeating the arguments above, one can check that

$$\mathbb{E}\left[\left\|u_n(t) - u_n(s)\right\|_{V'}^r\right] \leq C(t-s)^{r/2}.$$

It follows

$$\mathbb{E}\int_0^T \int_0^T \frac{\left\|u_n(t) - u_n(s)\right\|_{V'}^r}{|t-s|^{1+\alpha r}} dsdt \leq \int_0^T \int_0^T \frac{C}{|t-s|^{\frac{2-r}{2}+\alpha r}} dsdt =: C_r < \infty$$

if $\alpha r < \frac{r}{2}$. The condition $\alpha r > 1 - \frac{r}{p}$ of Theorem 15 is fulfilled for $p = 2$ if $\alpha r > 1 - \frac{r}{2}$. Thus if

$$1 - \frac{r}{2} < \alpha r < \frac{r}{2}$$

both conditions are satisfied. For $r = 1$ this is impossible, as seen in the previous step, but for every $r > 1$ there exists $\alpha \in \left(0, \frac{1}{2}\right)$ with such property.

The conclusion is;

Theorem 21 *There exist* (α, r) *with* $\alpha r > 1 - \frac{r}{2}$ *and* $C > 0$ *such that*

$$\mathbb{E}\left[\|u_n\|_{W^{\alpha,r}(0,T;V')}\right] \leq C.$$

Form the previous results:

Corollary 22 *The family of laws of* u_n *is tight in* $L^2(0,T;H)$.

4.4.2 Application Gyongy-Krylov criterion and conclusion of the proof of existence

In the sequel we assume u_0 and f deterministic but a simple variant of the argument covers the random case.

Let u_n be the Galerkin sequence. Assume we have a subsequence u_{n_k} and a process u with the following properties:

1. u has the regularity prescribed by the theorem
2. u_{n_k} converges to u in probability in $L^2(0,T;H)$
3. u_{n_k} converges weakly to u in $L^2_{\mathcal{F}}(0,T;V)$ and weak star in $C_{\mathcal{F}}([0,T];H)$.

Then with some work we can pass to the limit in the weak formulation of the equations; property 2 is needed to pass to the limit in the quadratic term. The existence of a subsequence with properties 1-3 comes from (10) (and a variant of the argument of Remark 14 to identify the limit as the same function). From this subsequence, from the bounds of the previous section and the compactness theorem, we may also extract another one such that u_{n_k} converges in law, in the strong topology of $L^2(0,T;H)$, to the law of u (again we identify the limit by a variant of the argument of Remark 14). The convergence in law implies convergence in probability, in the strong topology of $L^2(0,T;H)$, by Gyongy-Krylov criterium, which is applicable as shown below in this section.

Hence we have to show that Gyongy-Krylov criterium applies. Take any pair of subsequences $(n_1(k), n_2(k))$ and consider the sequence of pairs $\left(u_{n_1(k)}, u_{n_2(k)}\right)$. Since (u_n) is tight in $L^2(0,T;H)$, it is very easy to check that also $\left(u_{n_1(k)}, u_{n_2(k)}\right)$ is tight in $L^2(0,T;H)^2$. Let $k(h)$ be a subsequence such that $\left(u_{n_1(k(h))}, u_{n_2(k(h))}\right)$ converges in law to some μ. We only need to prove that $\mu(D) = 1$. It is sufficient to prove that, given $\epsilon > 0$, $\mu(D_\epsilon^c) = 0$, where

$$D_\epsilon^c = \left\{(u, v) \in L^2(0, T; H)^2 ; \|u - v\|_{L^2(0,T;H)} > \epsilon\right\}.$$

Since this is an open set, where weakly convergent probabilities are lower semicontinuous, it is sufficient to prove that

$$\lim_{h\to\infty} \mathbb{P}\left(\left\|u_{n_1(k(h))} - u_{n_2(k(h))}\right\|_{L^2(0,T;H)} > \epsilon\right) = 0.$$

To shorten the notations, let us denote the subsequences $u_{n_1(k(h))}, u_{n_2(k(h))}$ simply by $u_{n(h)}, u_{m(h)}$.

Consider the triple $\left(u_{n(h)}, u_{m(h)}, W\right)$ and call Q_h its law. It converges weakly to a measure Q on $L^2(0, T; H)^2 \times C([0, T]; H)$. By Skorohod representation theorem there exists a new probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ and random variables $\left(\widetilde{u}_{n(h)}, \widetilde{u}_{m(h)}, \widetilde{W}_h\right)$ with laws Q_h, and a random variable $\left(\widetilde{u}^{(1)}, \widetilde{u}^{(2)}, \widetilde{W}\right)$ with law Q, such that $\left(\widetilde{u}_{n(h)}, \widetilde{u}_{m(h)}, \widetilde{W}_h\right)$ converges $\widetilde{\mathbb{P}}$-a.s. to $\left(\widetilde{u}^{(1)}, \widetilde{u}^{(2)}, \widetilde{W}\right)$ in $L^2(0, T; H)^2 \times C([0, T]; H)$. With some work that we do not develop (this is the most difficult part of the argument), one can introduce a filtration $\widetilde{\mathcal{F}}_t$ and show that the processes are progressively measurable, $\widetilde{W}_h$ has the form $\widetilde{W}_h(t) := \sum_k \sigma_k \widetilde{W}_t^{k,h}$ where $\widetilde{W}_t^{k,h}$ are (for each h) independent Brownian motions and we have

$$\begin{aligned}
&\left\langle \widetilde{u}_{n(h)}(t), \phi\right\rangle - \int_0^t b\left(\widetilde{u}_{n(h)}(s), \pi_{n(h)}\phi, \widetilde{u}_{n(h)}(s)\right) ds \\
= \;& \langle u_0, \phi\rangle + \int_0^t \left\langle \widetilde{u}_{n(h)}(s), A\phi\right\rangle ds \\
&+ \int_0^t \left\langle f_n(s) + F_n\left(\widetilde{u}_{n(h)}(s)\right), A\phi\right\rangle ds \\
&+ \sum_{k\in K} \int_0^t \left\langle \sigma_k^n\left(\widetilde{u}_{n(h)}(s)\right), \phi\right\rangle d\widetilde{W}_s^{k,h}
\end{aligned}$$

and similarly for $\widetilde{u}_{m(h)}$. From the strong convergence and minor work we deduce

$$\begin{aligned}
&\left\langle \widetilde{u}^{(i)}(t), \phi\right\rangle - \int_0^t b\left(\widetilde{u}^{(i)}(s), \phi, \widetilde{u}^{(i)}(s)\right) ds \\
= \;& \langle u_0, \phi\rangle + \int_0^t \left\langle \widetilde{u}^{(i)}(s), A\phi\right\rangle ds \\
&+ \int_0^t \left\langle f(s) + F\left(\widetilde{u}^{(i)}(s)\right), A\phi\right\rangle ds \\
&+ \sum_{k\in K} \int_0^t \left\langle \sigma_k\left(\widetilde{u}^{(i)}(s)\right), \phi\right\rangle d\widetilde{W}_s^{k}
\end{aligned}$$

for $i=1,2$, where $\widetilde{W}_t^k$ are independent Brownian motions. From this we already know that we deduce $\widetilde{u}^{(1)}=\widetilde{u}^{(2)}$ and this means that the projection of Q on $L^2(0,T;H)^2$ is concentrated on the diagonal.

4.5 3D Navier-Stokes equations with additive noise

Let us add a few remarks on the 3D Navier-Stokes equations in a domain D, just with additive noise, which we shortly write in abstract form

$$du = Audt + B(u,u)\,dt + f + F(u) + \sum_k \sigma_k dW_t^k. \tag{12}$$

Writing the theory of 3D Navier-Stokes equations in the same detail as above is not coherent with the format of these lectures. Therefore we shall limit ourselves to an outline of ideas. The definition of weak solution is similar to the 2D case. However, two new elements are present. The first one is that we just require *weak continuity in* H, namely continuity in the weak topology of H:

$$u \in C([0,T];H_w) \cap L^\infty(0,T;H) \cap L^2(0,T;V). \tag{13}$$

For every test function $\phi \in H$, the function $t \mapsto \langle u(t),\phi\rangle$ is continuous. Since we assume $u \in L^\infty(0,T;H)$, a property like

$$u \in C\left([0,T];D(A)'\right)$$

implies $u \in C([0,T];H_w)$.

The second detail is that now we cannot prove the energy identity; and if u is a weak solution (in the sense of weak regularity plus the weak formulation of the equation), we cannot even prove an energy inequality. We have to include it in the definition, if we want to use it; and the existence of weak solutions satisfying the energy inequality can be established. Sometimes the weak solutions which have an energy inequality are called Leray solutions.

The other aspect which drastically changes are the interpolation inequalities. The property (b,B,P etc. are refined as in the 2D case)

$$b(u,v,w) \le \|v\|_V \|u\|_{\mathbb{L}^4} \|w\|_{\mathbb{L}^4}$$

is always true, being given by Hölder inequality. But then, recall Sobolev embedding theorem in dimension d: $W^{\alpha,p}(D) \subset L^q(D)$ if $\frac{1}{q}=\frac{1}{p}-\frac{\alpha}{d}$. We thus have:

$$\|f\|_{L^4} \overset{d=2}{\le} \|f\|_{W^{\frac{1}{2},2}} \le \|f\|_{L^2}^{1/2} \|f\|_{W^{1,2}}^{1/2}$$

$$\|f\|_{L^4} \overset{d=3}{\le} \|f\|_{W^{\frac{3}{4},2}} \le \|f\|_{L^2}^{1/4} \|f\|_{W^{1,2}}^{3/4}.$$

This increase of the power of $\|f\|_{W^{1,2}}$ has tremendous consequences. In particular, from the regularity

$$u \in L^{\infty}(0,T;H) \cap L^{2}(0,T;V)$$

we cannot deduce anymore $u \in L^{4}\left(0,T;\mathbb{L}^{4}\right)$, property that we have used in essential way in $d=2$. Now we only have $u \in L^{8/3}\left(0,T;\mathbb{L}^{4}\right)$:

$$\int_0^T \|u(t)\|_{\mathbb{L}^4}^{8/3} dt \leq C \int_0^T \|u(t)\|_H^{2/3} \|u(t)\|_V^2 dt \leq C \sup_{t\in[0,T]} \|u(t)\|_H^{2/3} \int_0^T \|u(t)\|_V^2 dt.$$

4.5.1 The problem of uniqueness

Let us illustrate the problem in the particular case $F=0$, $\sigma_k=0$. If $u^{(i)}$ are two solutions and we set $w=u^{(1)}-u^{(2)}$, we have

$$\langle w(t),\phi\rangle - \int_0^t \left(b\left(u^{(1)},\phi,u^{(1)}\right) - b\left(u^{(2)},\phi,u^{(2)}\right)\right)(s)\,ds = \int_0^t \langle w(s),A\phi\rangle\,ds$$

and since

$$\begin{aligned} & b\left(u^{(1)},\phi,u^{(1)}\right) - b\left(u^{(2)},\phi,u^{(2)}\right) - b(w,\phi,w) \\ = \; & b\left(u^{(2)},\phi,w\right) + b\left(w,\phi,u^{(2)}\right) \end{aligned}$$

we get

$$\begin{aligned} & \langle w(t),\phi\rangle - \int_0^t \left(b\left(w(s),\phi,w(s)\right)\right)ds \\ = \; & \int_0^t \langle w(s),A\phi\rangle\,ds - \int_0^t \left(b\left(u^{(2)},\phi,w\right) + b\left(w,\phi,u^{(2)}\right)\right)(s)\,ds. \end{aligned}$$

Up to details (in particular the next fact requires Leray solutions), we have

$$\begin{aligned} \|w(t)\|_H^2 + 2\nu\int_0^t \|\nabla w(s)\|_H^2\,ds &\leq -2\int_0^t \left(b\left(u^{(2)},w,w\right) + b\left(w,w,u^{(2)}\right)\right)(s)\,ds \\ &= -2\int_0^t b\left(w,w,u^{(2)}\right)(s)\,ds. \end{aligned}$$

But now

$$\begin{aligned} \left|b\left(w,w,u^{(2)}\right)\right| &\leq C\|w\|_V \|w\|_{\mathbb{L}^4} \left\|u^{(2)}\right\|_{\mathbb{L}^4} \\ &\leq C\|w\|_V^{7/4} \|w\|_H^{1/4} \left\|u^{(2)}\right\|_{\mathbb{L}^4}. \end{aligned}$$

We may use Young's inequality $ab \leq \nu a^{8/7} + C_\nu b^8$:

$$\left|b\left(w, w, u^{(2)}\right)\right| \leq \nu \|w\|_V^2 + C_\nu \|w\|_H^2 \left\|u^{(2)}\right\|_{\mathbb{L}^4}^8$$

so that

$$\|w(t)\|_H^2 + \nu \int_0^t \|\nabla w(s)\|_H^2 \, ds \leq C_\nu \int_0^t \|w(s)\|_H^2 \left(\left\|u^{(2)}(s)\right\|_{\mathbb{L}^4}^8 + 1\right) ds.$$

Gronwall this time does not apply because we do not know that $u^{(2)}$ is of class $u \in L^8\left(0, T; \mathbb{L}^4\right)$; we only know $u \in L^{8/3}\left(0, T; \mathbb{L}^4\right)$.

4.5.2 Estimates on Galerkin and tightness

The definition of Galerkin approximations is the same as in 2D and the first energy inequalities are proved in the same way. We get the same bounds (10)-(11). With due work we deduce that laws of u_n are tight in $L^2(0, T; H)$. A little additional work gives tightness in

$$L^2(0, T; H) \cap C\left([0, T]; D(A)'\right).$$

Moreover, we have weak convergence in the topologies of (10), hence any limit measure of subsequences is supported on the regularity space of the definition of weak solution. It remains to prove that such limit measures (which exist) correspond to solutions of the 3d Navier-Stokes equations.

4.5.3 Definition of solution and convergence

Until now a solution has been a stochastic process. However, the previous construction provides only a probability measure on certain function spaces. One can always introduce a stochastic process with such measure as law, but it is just an artefact, it is not defined on the original probability space where the problem was formulated. Therefore we give the following definition, which is called weak in a double sense: weak probabilistically and weak analytically.

Definition 23 *Let $u_0 \in H$ be given. A weak solution of the 3D Navier-Stokes equations (12) with initial condition u_0 is a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, a family of independent Brownian motions W_t^k, $k \in K$, over such space, and a stochastic process u, with paths of class (13), progressively measurable (adapted in H, being weakly continuous), which satisfies*

$$\begin{aligned} & \langle u(t), \phi \rangle - \int_0^t b(u(s), \phi, u(s)) \, ds \\ = \quad & \langle u_0, \phi \rangle + \int_0^t \langle u(s), A\phi \rangle \, ds + \sum_{k \in K} \sqrt{\lambda_k} \langle \sigma_k, \phi \rangle W_t^k \end{aligned}$$

for every $\phi \in D(A)$. We also require

$$\mathbb{E}\left[\|u(t)\|_{L^2}^2\right] + 2\nu \int_0^t \mathbb{E}\|\nabla u(s)\|_{L^2}^2 \, ds = \|u_0\|_{L^2}^2 + t \sum_{k \in K} \lambda_k \|\sigma_k\|_{L^2}^2 .$$

Notice that assuming u_0 random provokes a problem: that a probability space should be defined in advance; this is not compatible with the construction. An alternative then is to prescribe the law of u_0 on H.

Let us sketch the proof of existence of such solutions. Let u_n be the Galerkin approximations defined above. In fact consider for each n the pair

$$(u_n, W_n)$$

where $W_n(t) := \sum_k \sigma_k^n W_t^k$, which is a random variable with values in

$$L^2(0,T;H) \times C([0,T];H). \tag{14}$$

Call Q_n its law. The family $(Q_n)_{n\in\mathbb{N}}$ is tight in this space (the tightness of the second component follows from its convergence to $W(t) := \sum_k \sigma_k W_t^k$). Let us extract a subsequence (Q_{n_k}) which weakly converges to a probability measure Q. Then, for every smooth compact support divergence free test vector field $\phi(t,x)$, consider the functional

$$\begin{aligned} J_\phi(u,w) \; : \; &= 1 \wedge \\ &\left| \int_0^T \langle u, (\partial_s + A)\phi\rangle \, ds + \int_0^T b(u,\phi,u)\, ds + \int_0^T \langle f + F(u), \phi\rangle - \int_0^T \langle w, \partial_s \phi\rangle \, ds \right| . \end{aligned}$$

Notice that, if a sequence of functions $(u_n) \subset L^2(0,T;H)$ converges strongly to u, and ϕ is bounded, then $b(u_n,\phi,u_n)$ converges to $b(u,\phi,u)$. Thus the functional J_ϕ is continuous on the product space (14), and bounded. Hence

$$\lim_{k\to\infty} \int J_\phi(u,w)\, Q_{n_k}(du,dw) = \int J_\phi(u,w)\, Q(du,dw).$$

But

$$\begin{aligned} &\int J(u,w)\, Q_{n_k}(du,dw) \\ = \;& \mathbb{E}\left[1 \wedge \left| \int_0^T \langle u_{n_k}, (\partial_s + A)\phi\rangle \, ds + \int_0^T b(u_{n_k},\phi,u_{n_k})\, ds \right.\right. \\ &\left.\left. + \int_0^T \langle f + F(u_{n_k}), \phi\rangle - \int_0^T \langle W_{n_k}, \partial_s\phi\rangle \, ds \right| \right] . \end{aligned}$$

The equation satisfied by u_{n_k} may be rewritten for time-dependent test functions ϕ as we did in Chapter one when dealing with the Stokes problem:

$$\int_0^T \langle u_{n_k}, (\partial_s + A)\phi\rangle \, ds + \int_0^T b(u_{n_k}, \pi_{n_k}\phi, u_{n_k})\, ds + \int_0^T \langle f + F(u_{n_k}), \pi_{n_k}\phi\rangle - \int_0^T \langle W_{n_k}, \partial_s\phi\rangle \, ds.$$

Hence

$$\begin{aligned} & \int J(u,w)\, Q_{n_k}(du,dw) \\ = \; & \mathbb{E}\left[1 \wedge \left|\int_0^T b\left(u_{n_k}, \phi - \pi_{n_k}\phi, u_{n_k}\right) ds + \int_0^T \left\langle f + F\left(u_{n_k}\right), \phi - \pi_{n_k}\phi\right\rangle\right|\right]. \end{aligned}$$

Let us prove it goes to zero:

$$\begin{aligned} \mathbb{E}\left|\int_0^T b\left(u_{n_k}, \phi - \pi_{n_k}\phi, u_{n_k}\right) ds\right| & \le \left\|\phi - \pi_{n_k}\phi\right\|_V \mathbb{E}\int_0^T \left\|u_{n_k}\right\|_H \left\|u_{n_k}\right\|_V ds \\ & \le \left\|\phi - \pi_{n_k}\phi\right\|_V \mathbb{E}\left[\sup_{t\in[0,T]} \left\|u_{n_k}(t)\right\|_H \int_0^T \left\|u_{n_k}\right\|_V ds\right] \end{aligned}$$

and $\left\|\phi - \pi_{n_k}\phi\right\|_V \to 0$ (using $\phi \in V$ and the commutativity of π_{n_k} with A),

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left\|u_{n_k}(t)\right\|_H \int_0^T \left\|u_{n_k}\right\|_V ds\right] \le C$$

by the bounds (10); and

$$\begin{aligned} & \mathbb{E}\left|\int_0^T \left\langle f + F\left(u_{n_k}\right), \phi - \pi_{n_k}\phi\right\rangle\right| \\ \le \; & \left\|\phi - \pi_{n_k}\phi\right\|_V \left(\mathbb{E}\int_0^T \|f\|_{V'} ds + C\mathbb{E}\int_0^T \left(1 + \left\|u_{n_k}\right\|_H\right) ds\right) \end{aligned}$$

and the argument is similar and easier.

It follows that Q satisfies

$$\int J_\phi(u,w)\, Q(du,dw) = 0$$

for every ϕ. Realize Q as law of $\left(\widetilde{u}, \widetilde{W}\right)$. The second marginal of Q is the law of $W := \sum_k \sigma_k dW_t^k$, being the weak limit of the second marginal of Q_{n_k}, which is the law of W_n which converges a.s. to W; hence $\widetilde{W}$ has the same law of W. Working a little bit with Gaussianity, we may check that $\widetilde{W}$ is represented as $\sum_k \sigma_k d\widetilde{W}_t^k$ where $\widetilde{W}_t^k$ are independent Brownian motions.

We have

$$\begin{aligned} & \widetilde{\mathbb{E}}\left[1 \wedge \left|\int_0^T \left\langle \widetilde{u}, \left(\partial_s + A\right)\phi\right\rangle ds + \int_0^T b\left(\widetilde{u}, \phi, \widetilde{u}\right) ds \right.\right. \\ & \left.\left. + \int_0^T \left\langle f + F\left(\widetilde{u}\right), \phi\right\rangle - \int_0^T \left\langle \widetilde{W}, \partial_s\phi\right\rangle ds\right|\right] = 0 \end{aligned}$$

hence $\widetilde{\mathbb{P}}$-a.s.

$$\int_0^T \langle \widetilde{u}, (\partial_s + A)\,\phi\rangle\, ds + \int_0^T b\,(\widetilde{u}, \phi, \widetilde{u})\, ds + \int_0^T \langle f + F\,(\widetilde{u})\,, \phi\rangle - \int_0^T \left\langle \widetilde{W}, \partial_s \phi \right\rangle ds = 0$$

for every given ϕ (the negligible set where this may not hold depends on ϕ). Taking first a dense countable set of ϕ's, so that we can inverte the quantifiers and then a convergence argument based on pathwise regularity, we deduce that, $\widetilde{\mathbb{P}}$-a.s., we have

$$\int_0^T \langle \widetilde{u}, (\partial_s + A)\,\phi\rangle\, ds + \int_0^T b\,(\widetilde{u}, \phi, \widetilde{u})\, ds + \int_0^T \langle f + F\,(\widetilde{u})\,, \phi\rangle - \int_0^T \left\langle \widetilde{W}, \partial_s \phi \right\rangle ds = 0$$

for all ϕ, which is the definition of weak solution.

5 Summary

The main open problem outlined in this Chapter is the continuation of the one posed in the previous chapter, namely the link between a real irregular boundary and stochastic models of fluids; here the problem is enriched of the dependence on the flow intensity, a very realistic feature, which poses a new technical issue, namely the presence of the Wong-Zakai corrector in the limit equation. We have also seen that noise introduces energy, in the average, hence the model should be corrected by an energy loss.

The main techniques illustrated in this Chapter are the use of Itô formula, an interesting idea for uniqueness, its consequence through a criterion of Gyongy and Kryolov, and expecially the method of compactness, quite universal and useful in many fields.