

# Mathematics of Two-Dimensional Turbulence

Sergei Kuksin, Armen Shirikyan

April 26, 2012

To

*Yulia, Nikita, Masha*

and

*Anna, Rafael, Gabriel*

# Contents

<i>Introduction</i> . . . . .	v
<b>1 Preliminaries</b> . . . . .	<b>1</b>
1.1 Function spaces . . . . .	1
1.1.1 Functions of the space variables . . . . .	1
1.1.2 Functions of space and time variables . . . . .	4
1.2 Basic facts from the measure theory . . . . .	5
1.2.1 $\sigma$ -algebras and measures . . . . .	5
1.2.2 Convergence of integrals . . . . .	6
1.2.3 Metrics on the space of probabilities and convergence of measures . . . . .	7
1.2.4 Couplings and maximal couplings of probability measures	16
1.2.5 Kantorovich functionals . . . . .	20
1.3 Markov processes and random dynamical systems . . . . .	21
1.3.1 Markov processes . . . . .	21
1.3.2 Random dynamical systems . . . . .	26
1.3.3 Markov RDS . . . . .	27
1.3.4 Invariant and stationary measures . . . . .	30
<i>Notes and comments</i> . . . . .	33
<b>2 Two-dimensional Navier–Stokes equations</b> . . . . .	<b>35</b>
2.1 Cauchy problem for the deterministic system . . . . .	35
2.1.1 Equations and boundary conditions . . . . .	35
2.1.2 Leray decomposition . . . . .	36
2.1.3 Properties of some multilinear maps . . . . .	38
2.1.4 Reduction to an abstract evolution equation . . . . .	40
2.1.5 Existence and uniqueness of solution . . . . .	42
2.1.6 Regularity of solutions . . . . .	47
2.1.7 Navier–Stokes process . . . . .	50
2.1.8 Foias–Prodi estimates . . . . .	52
2.1.9 Some hydrodynamical terminology . . . . .	55
2.2 Stochastic Navier–Stokes equations . . . . .	56
2.3 Navier–Stokes equations with random kicks . . . . .	59
2.3.1 Existence and uniqueness of solution . . . . .	60
2.3.2 Markov chain and RDS . . . . .	61

2.3.3	Additional results . . . . .	62
2.4	Navier–Stokes equations with white noise . . . . .	66
2.4.1	Existence and uniqueness of solution, and Markov process . . . . .	66
2.4.2	Additional results: energy balance, higher Sobolev norms, and time averages . . . . .	73
2.4.3	Universality of white noise forces . . . . .	79
2.4.4	RDS associated with Navier–Stokes equations . . . . .	81
2.5	Existence of a stationary distribution . . . . .	83
2.5.1	Bogolyubov–Krylov argument . . . . .	83
2.5.2	Application to Navier–Stokes equations . . . . .	84
2.6	Appendix: some technical proofs . . . . .	88
	<i>Notes and comments</i> . . . . .	94
<b>3</b>	<b>Uniqueness of stationary measure</b> . . . . .	<b>97</b>
3.1	Three results on uniqueness and mixing . . . . .	100
3.1.1	Decay of a Kantorovich functional . . . . .	100
3.1.2	Coupling method: uniqueness and mixing . . . . .	102
3.1.3	Coupling method: exponential mixing . . . . .	106
3.2	Dissipative RDS with bounded kicks . . . . .	110
3.2.1	Main result . . . . .	110
3.2.2	Coupling . . . . .	112
3.2.3	Proof of Theorem 3.2.5 . . . . .	116
3.2.4	Application to Navier–Stokes equations . . . . .	118
3.3	Navier–Stokes system perturbed by white noise . . . . .	121
3.3.1	Main result and scheme of its proof . . . . .	122
3.3.2	Recurrence: proof of Proposition 3.3.6 . . . . .	127
3.3.3	Stability: proof of Proposition 3.3.7 . . . . .	131
3.4	Navier–Stokes system with unbounded kicks . . . . .	135
3.4.1	Formulation of the result . . . . .	135
3.4.2	Proof of Theorem 3.4.1 . . . . .	137
3.5	Further results and generalisations . . . . .	138
3.5.1	Flandoli–Maslowski theorem . . . . .	138
3.5.2	Exponential mixing for the Navier–Stokes system with white noise . . . . .	140
3.5.3	Convergence for functionals on higher Sobolev spaces . . . . .	142
3.5.4	Mixing for Navier–Stokes equations perturbed by a com- pound Poisson process . . . . .	145
3.5.5	Description of some results on uniqueness and mixing for other PDE’s . . . . .	147
3.5.6	An alternative proof of mixing for kick force models . . . . .	148
3.6	Appendix: some technical proofs . . . . .	152
3.6.1	Proof of Lemma 3.3.11 . . . . .	152
3.6.2	Recurrence for Navier–Stokes equations with unbounded kicks . . . . .	155
3.6.3	Exponential squeezing for Navier–Stokes equations with unbounded kicks . . . . .	158

3.7	Relevance of the results for physics . . . . .	160
	<i>Notes and comments</i> . . . . .	161
<b>4</b>	<b>Ergodicity and limiting theorems</b>	<b>165</b>
4.1	Ergodic theorems . . . . .	165
4.1.1	Strong law of large numbers . . . . .	165
4.1.2	Law of iterated logarithm . . . . .	170
4.1.3	Central limit theorem . . . . .	173
4.2	Random attractors and stationary distributions . . . . .	174
4.2.1	Random point attractors . . . . .	175
4.2.2	Ledrappier–Le Jan–Crauel theorem . . . . .	179
4.2.3	Ergodic RDS and minimal attractors . . . . .	184
4.2.4	Application to the Navier–Stokes system . . . . .	189
4.3	Dependence of stationary measure on the force . . . . .	194
4.3.1	Regular dependence on parameters . . . . .	194
4.3.2	Universality of white noise perturbations . . . . .	199
4.4	Relevance of the results for physics . . . . .	200
	<i>Notes and comments</i> . . . . .	201
<b>5</b>	<b>Inviscid limit</b>	<b>203</b>
5.1	Balance relations . . . . .	203
5.1.1	Energy and enstrophy . . . . .	203
5.1.2	Balance relations . . . . .	204
5.1.3	Pointwise exponential estimates . . . . .	208
5.2	Limiting measures . . . . .	210
5.2.1	Existence of accumulation points . . . . .	210
5.2.2	Estimates for the densities of the energy and enstrophy . . . . .	219
5.2.3	Further properties of the limiting measures . . . . .	223
5.2.4	Other scalings . . . . .	228
5.2.5	Kicked Navier–Stokes system . . . . .	229
5.2.6	Inviscid limit for the complex Ginzburg–Landau equation . . . . .	231
5.3	Relevance of the results for physics . . . . .	232
	<i>Notes and comments</i> . . . . .	235
<b>6</b>	<b>Miscellanies</b>	<b>237</b>
6.1	3D Navier–Stokes system in thin domains . . . . .	237
6.1.1	Preliminaries on the Cauchy problem . . . . .	238
6.1.2	Large-time asymptotics of solutions . . . . .	239
6.1.3	The limit $\varepsilon \rightarrow 0$ . . . . .	241
6.2	Ergodicity and Markov selection . . . . .	242
6.2.1	Finite-dimensional stochastic differential equations . . . . .	243
6.2.2	Da Prato–Debussche–Odasso theorem . . . . .	247
6.2.3	Flandoli–Romito theorem . . . . .	251
6.3	Navier–Stokes with very degenerate noise . . . . .	255
6.3.1	2D Navier–Stokes equations: controllability and mixing properties . . . . .	255

6.3.2	3D Navier–Stokes equations with a degenerate noise . . .	257
<b>7</b>	<b>Appendix</b>	<b>259</b>
7.1	Monotone class theorem . . . . .	259
7.2	Standard measurable spaces . . . . .	260
7.3	Projection theorem . . . . .	261
7.4	Gaussian random variables . . . . .	262
7.5	Weak convergence of random measures . . . . .	265
7.6	Gelfand triple and Yosida approximation . . . . .	266
7.7	Itô formula in Hilbert spaces . . . . .	268
7.8	Local time for continuous Itô processes . . . . .	274
7.9	Krylov’s estimate . . . . .	276
7.10	Girsanov theorem . . . . .	278
7.11	Martingales . . . . .	279
7.12	Limit theorems for discrete-time martingales . . . . .	281
7.13	Martingale approximation for Markov processes . . . . .	283
7.14	Generalised Poincaré inequality . . . . .	285
7.15	Functions with a discrete essential range . . . . .	286
<b>8</b>	<b>Solutions to some exercises</b>	<b>287</b>
	<i>Bibliography</i>	<b>313</b>
	<i>Notation and conventions</i>	<b>314</b>
	<i>Indices</i>	<b>317</b>
	Subject Index . . . . .	317
	Author Index . . . . .	319
	Citation Index . . . . .	321

## Preface

### Equations and forces

Two-dimensional statistical hydrodynamics studies statistical properties of the velocity field  $u(t, x)$  of (imaginary) two-dimensional fluid satisfying the stochastic 2D Navier–Stokes equations

$$\begin{aligned} \dot{u}(t, x) + \langle u, \nabla \rangle u - \nu \Delta u + \nabla p &= f(t, x), \quad \operatorname{div} u = 0, \\ u &= (u^1, u^2), \quad x = (x_1, x_2). \end{aligned} \tag{0.1}$$

Here  $\nu > 0$  is the kinematic viscosity,  $u = u(t, x)$  is the velocity of the fluid,  $p = p(t, x)$  is the pressure, and  $f$  is the density of an external force applied to the fluid. The space variable  $x$  belongs to a two-dimensional domain, which in this book is supposed to be bounded. Suitable boundary conditions are assumed. For example, one may consider the case when the domain is a rectangle  $(0, a) \times (0, b)$ , where  $a$  and  $b$  are positive numbers, and the equations are supplemented with periodic boundary conditions; that is to say, the space variable  $x$  belongs to the torus  $\mathbb{R}^2 / (a\mathbb{Z} \oplus b\mathbb{Z})$  (in the case of periodic boundary conditions we will assume that space-meanvalues of the force  $f$  and the solution  $u$  vanish). Equations (0.1) are *stochastic* in the sense that the initial condition  $u_0 = u(0, x)$ , or the force  $f$ , or both of them, are random, i.e. depend on a random parameter. So the solutions  $u$  are random vector fields. The task is to study various characteristics of  $u$  averaged in ensemble, or to study their properties which hold for most values of the random parameter. In this book, we assume that the force is random and refer the reader to [FMRT01] for a mathematical treatment of the Navier–Stokes equations with zero (or deterministic) force and random initial data.

The Reynolds number  $R$  of a random velocity field  $u(t, x)$  is defined as

$$R = \frac{\langle \text{characteristic scale for } x \rangle \cdot (\mathbb{E}E(u))^{1/2}}{\nu},$$

where  $E(u) = \frac{1}{2} \int |u(x)|^2 dx$  is the kinetic energy of the fluid and  $\mathbb{E}$  denotes the average in ensemble. Since the forces we consider are smooth, then the solutions  $u$  of (0.1) are regular in  $x$  and their space-scale is of order one. So  $R \sim \nu^{-1} (\mathbb{E}E(u))^{1/2}$ . A velocity field  $u$  is called *turbulent* if  $R \gg 1$ . Turbulent solutions for (0.1) are of prime interest.

If a motion of “physical” three-dimensional fluid is parallel to the  $(x_1, x_2)$ -plane and its velocity depends only on  $(x_1, x_2)$ , i.e.  $u = u(t, x_1, x_2)$  and  $u^3 = 0$ , then  $(u^1, u^2)(t, x_1, x_2)$  satisfy (0.1). Such flows are called *two-dimensional*. Turbulent flows of real fluids never are two-dimensional (i.e., two-dimensional flows never are observed in experiments with high Reynolds number). Still, the 2D equations (0.1) and the 2D turbulence which they describe are now intensively studied by mathematicians, physicists and engineers since, firstly, they appear in physics outside the realm of classical hydrodynamics (e.g., they describe flows of

2D films, see Figure 1), secondly, they make a model<sup>1</sup> for the 3D Navier–Stokes equations and the 3D turbulence and, thirdly, the 3D statistical hydrodynamics in thin domains is approximately two-dimensional; see Section 6.1 of this book. Accordingly, two-dimensional statistical hydrodynamics is important for meteorology to model intermediate-scale flows in atmosphere (see Figures 6.1 and 6.2).

Statistical properties of the random force  $f$  are very important. It is natural and traditional to assume that

- (a) the random field  $f(t, x)$  is smooth in  $x$ ,
- (b) it is stationary in  $t$  with fast decaying correlations.

If the space domain is unbounded, we should also assume that

- (c) the space correlations of  $f$  decay fast.

However, (c) is not relevant for this book since we only consider flows in bounded domains.

In mathematics, the point of view<sup>2</sup> that turbulence in dimensions two and three should be described by the Navier–Stokes equations with a random force satisfying (a)–(c) goes back to A.N. Kolmogorov; see in [VF88]. Also see that book for some results on stochastic Navier–Stokes equations in the whole space  $\mathbb{R}^d$ ,  $d = 2$  or  $3$ , with a random force satisfying (a)–(c).

We consider three classes of random forces:

*Kick forces.* These are random fields of the form

$$f(t, x) = h(x) + \sum_{k \in \mathbb{Z}} \delta(t - \tau_k) \eta_k(x), \quad (0.2)$$

where  $h$  is a smooth deterministic function,  $\tau_k = k\tau$  with some  $\tau > 0$ , and  $\{\eta_k\}$  are independent identically distributed random vector functions, which we assume to be divergence-free. For  $t \in (\tau_{k-1}, \tau_k)$  (i.e. between two consecutive kicks) a solution  $u(t, x)$  for (0.1), (0.2) satisfies the deterministic equations (0.1) <sub>$f=h$</sub> , and at the time  $\tau_k$ , when the  $k^{\text{th}}$  kick  $\eta_k(x)$  comes, it has an instant increment equal to that kick; see Subsection 2.3. The kick forces are singular in  $t$  and are not stationary in  $t$ , but statistically periodic (the difference between the two notions is not big if the time  $t$  is much larger than the period  $\tau$  between the kicks). An advantage of this class of random forces is that the kicks  $\eta_k$  may have any statistics.

*White in time forces.* These are random fields of the form

$$f(t, x) = h(x) + \frac{d}{dt} \zeta(t, x), \quad (0.3)$$

---

<sup>1</sup>This model is not perfect since it is well known that the Navier–Stokes equations in dimensions 2 and 3 are very different. Still, it may be the best available now. Another popular model for the 3D Navier–Stokes system is the Burgers equation, see the review [BK07] by Bec and Khanin. For the stochastic 1D Burgers equation, see [Bor12].

<sup>2</sup>which is not at all a unique insight on the turbulence!



where  $h$  is as above and  $\zeta(t) = \zeta(t, \cdot)$  is a Wiener process in a space of smooth divergence-free vector functions. Such random fields are stationary and singular in  $t$ . A disadvantage is that they must be Gaussian; see Subsection 2.4.

*Compound Poisson processes.* These are kick forces (0.2) for which the periods  $\tau_k - \tau_{k-1}$  between kicks are independent exponentially distributed random variables.

A big technical advantage of these three classes of random forces is that the corresponding solution  $u(t, x)$ , regarded as a random process  $u(t, \cdot) =: u(t)$  in a space of vector fields, are Markov processes. At the moment of writing it is not clear how to extend the results of this book to arbitrary smooth random forces  $f$  satisfying (a) and (b).

### What is in this book?

We are concerned with basic problems and questions, interesting for physicists and engineers working in the theory of turbulence. Accordingly Chapters 3-5 (which form the main part of this book) end with sections, where we explain the physical relevance of the obtained results. These sections also provide brief summaries of the corresponding chapters.

In Chapters 3 and 4, our main goal is to justify, for the 2D case, the statistical properties of fluid's velocity field  $u(t, x)$  which physicists assume in their work. We refer the reader to the books [Bat82, Fri95, Gal02], written in a sufficiently rigorous way and where the underlying assumptions are formulated in clear manner.<sup>3</sup> The first postulate in the physical theory of turbulence is that *statistical properties of a turbulent flow  $u(t, x)$  converge, as time goes to infinity, to a statistical equilibrium independent of the initial data.* Mathematically speaking, it means that a process  $u(t, \cdot)$ , defined by Eq. (0.1) in a space of vector fields, has a unique stationary measure, and every solution converges to this measure in distribution. That is, the law of the random field  $x \mapsto u(t, x)$  (which is a time-dependent measure in a function space) converges, when  $t \rightarrow \infty$ , to the measure in question. Random processes possessing this property of "short-range memory" are said to be *mixing*.

In Chapter 3, we study the problem of convergence to a statistical equilibrium for Markov processes corresponding to equations with the three classes of random forces as above. We prove abstract theorems which establish the exponential mixing for certain classes of Markov processes. Next we show that these theorems apply to Eq. (0.1) if a random force  $f$  satisfies certain mild non-degeneracy assumptions. This establishes the convergence to a unique statistical equilibrium and proves that it is exponentially fast.

If the viscosity  $\nu$  and the force  $f$  continuously depend on a parameter in such a way that the former stays positive and the latter stays non-degenerate, then the stationary measure continuously depends on this parameter. For any fixed initial data  $u(0)$  the law of a corresponding solution  $u(t)$  continuously

---

<sup>3</sup>Apart from a few pages at the end, the book [Bat82] is about 3D flows. But all discussions and most of the results may be literally translated to the 2D case.

depends on the parameter as well. In Section 4.3, we show that this continuity is uniform in time  $t \geq 0$ . That is, in two space dimensions the statistical hydrodynamics is stable, no matter how big is the Reynolds number, whereas “usual” hydrodynamics of large Reynolds numbers is unstable.

The mixing has a number of important consequences, well-known in physics, but taken there for granted. Namely, consider any observable quantity  $F(u)$ , such as the first or second component of the velocity field  $u = (u^1, u^2)$ , or the energy  $E = \frac{1}{2} \int |u|^2 dx$ , or the enstrophy  $\frac{1}{2} \int (\text{curl } u)^2 dx$ . Then  $F(t) = F(u(t, \cdot))$  is an ergodic process. That is, *its time average converges to the ensemble average with respect to the stationary measure*. We show that the difference between the two mean values (in time and in ensemble) decays as  $T^{-\gamma}$ , where  $\gamma < 1/2$  and  $T$  is the time of averaging; see Section 4.1.1. Next, if the ensemble average for an observable  $F(u)$  vanishes, then the process  $F(t)$  satisfies the Central Limit Theorem: the law of the random variable

$$\frac{1}{\sqrt{T}} \int_0^T F(t) dt$$

converges, as  $T \rightarrow \infty$ , to a normal distribution  $N(0, \sigma)$ . For non-trivial observables  $F$ , the dispersion  $\sigma$  is strictly positive. In particular, for large  $T$  the random variables  $T^{-1/2} \int_0^T u^j(t, x) dt$ ,  $j = 1, 2$ , are almost Gaussian. Physicists say that *on large time scales a turbulent velocity field is Gaussian*. These and some other related results are proved in Chapter 4.

In Chapter 5 we study velocity fields  $u(t, x)$ , corresponding to solutions of (0.1) with a force (0.2) or (0.3) where  $h = 0$ , when the viscosity  $\nu$  is small and the Reynolds number is large. There we only discuss stationary measures and stationary in time solutions  $u_\nu$  (i.e., solutions  $u_\nu(t, x)$  such that the law  $\mathcal{D}(u_\nu(t))$  for each  $t$  equals the stationary measure). First we observe that for a limit of order one to exist as  $\nu \rightarrow 0$ , the force  $f$  should be proportional to  $\sqrt{\nu}$ ; see Section 5.2.4. So the equations read as

$$\dot{u}(t, x) + \langle u, \nabla \rangle u - \nu \Delta u + \nabla p = \sqrt{\nu} f(t, x), \quad \text{div } u = 0,$$

where  $f$  is the force (0.2) or (0.3) with  $h \equiv 0$ . This is in sharp contrast with the 3D theory, where it is believed that a limit of order one exists for the original scaling (0.1), without the additional factor  $\sqrt{\nu}$  in the right-hand side.<sup>4</sup> In that chapter we restrict ourselves to the case when the space domain is the square torus  $\mathbb{T}^2 = \mathbb{R}^2/2\pi\mathbb{Z}^2$ . The results remain true for the non-square tori  $\mathbb{R}^2/(a\mathbb{Z} \oplus b\mathbb{Z})$ , but the argument does not apply to the equations in a bounded domain with the Dirichlet boundary condition.

Denote by  $\mu_\nu$  the unique stationary measure. We show that the set of measures  $\{\mu_\nu, 0 < \nu \leq 1\}$  is tight (i.e., relatively compact) and that any limit point  $\mu_0 = \lim_{\nu_j \rightarrow 0} \mu_{\nu_j}$  is a non-trivial invariant measure for the Euler system

$$\dot{u}(t, x) + \langle u, \nabla \rangle u + \nabla p = 0, \quad \text{div } u = 0.$$

<sup>4</sup>Note that for the small-viscosity Burgers equation the right scaling of the force also is trivial, i.e. without any additional factor; see [BK07, Bor12].

It is supported by the set of divergence-free vector fields from the Sobolev space  $H^2$  of order two. This result well agrees with the popular belief that *the Euler equation is “responsible” for the 2D turbulence*. We do not know if a limiting measure  $\mu_0$  is unique, i.e. if  $\mu_0 = \lim_{\nu \rightarrow 0} \mu_\nu$ . But we know that the measures  $\mu_\nu$  satisfy, uniformly in  $\nu > 0$ , infinitely many algebraical relations, called the *balance relations*. These relations depend only on two scalar characteristics of the force  $f$ . This indicates some *universality features of the 2D turbulence*. Such universality is another physical belief. In Section 5.1.3, we use the balance relations to prove that for any  $t$  and  $x$  the random variables  $u_\nu(t, x)$  and  $\text{curl } u_\nu(t, x)$  have finite exponential moments uniformly in  $\nu \geq 0$ . In Section 5.2, we study further properties of the limiting measures  $\mu_0$ . In particular, we establish that any  $\mu_0$  has no atoms and that its support is an infinite-dimensional set.

Results of Chapter 5 make a foundation of mathematical theory of the space-periodic 2D turbulence. In Section 5.3, we discuss relation of these results with the existing heuristic theory of 2D turbulence, originated by Batchelor and Kraichnan.

The difference between the 2D turbulence and the real physical 3D turbulence is very big. In Chapter 6, we discuss a few rigorous results on the 3D turbulence, related to the 2D theory presented in the preceding sections. Namely, in Section 6.1 we discuss (without proof) the convergence of statistical characteristics of a flow in a thin 3D layer, corresponding to the 3D Navier-Stokes system with a random kick-force, to those of a 2D flow in the limiting 2D surface. In difference with similar deterministic results, the convergence holds uniformly in time. So a class of *anisotropic 3D turbulent flows may be approximated by 2D flows* like those which we consider in our book. Section 6.2 contains an exposition of results due to Da Prato–Debussche–Odasso and Flandoli–Romito, showing that weak solutions of the stochastic 3D Navier-Stokes system perturbed by a white in time random force (which a priori are non-unique) may be arranged to a Markov process. This process is mixing if the force is rough as a function of the space variable. Finally, in Section 6.3, we evoke the methods of control theory to study further properties of stationary measures for Eq. (0.1), (0.3).

### Other equations

The abstract theorems from Chapters 3 and 4 and the methods developed there to study solutions of Eq. (0.1) apply to many other stochastic equations. For instance, one can consider the stochastic complex Ginzburg–Landau equation with a conservative nonlinearity,

$$\dot{u} + i\Delta u - i|u|^{2m}u = \Delta u - u + f(t, x), \quad (0.4)$$

where  $x \in \mathbb{T}^d$ ,  $d \leq 3$ . If  $d = 1$  or  $2$ , then  $m \geq 0$ , while if  $d = 3$ , then one can take, say,  $m \in [0, 1]$ . Such equations describe the *optical turbulence*. If  $f$  is a bounded kick force, then direct analogues of the theorems in Chapters 3 and 4 remain true for (0.4) with the same proof.

Figure 1: The onset of 2D turbulence. Pictures 1–4 represent down-motion of a soap film, punctured by a comb at the top. The Reynolds number is increasing from a picture to picture. This is a 2D turbulent motion described by the 2D Navier–Stokes system (0.1).

However, if the force  $f$  is white in time, then the methods of Chapters 3 and 4 apply only to Eq. (0.4) with  $m = 1$  if  $d = 1$  and  $m < 1$  if  $d \geq 2$  (while the equation defines a good Markov process for any  $m$  as above). That is, for some deep reason, the arguments developed to treat the stochastic Navier–Stokes equations (0.1) with white in time forces apply only to PDE’s with conservative nonlinearities of degree  $\leq 3$ ;<sup>5</sup> see Section 3.5.5.

### Readers of this book

The book is aimed at mathematicians and physicists with some background in PDE and in stochastic. Standard university courses on these subjects are sufficient since the book is provided with preliminaries on function spaces (Section 1.1), on the 2D Navier–Stokes equations (Chapter 2) and on stochastics (Sections 1.2 and 1.3). There a reader will find all needed non-standard results.

### Acknowledgements

Our interest to the stochastic 2D Navier–Stokes equations (0.1) and the problem of 2D turbulence appeared during a research on the qualitative theory of randomly forced nonlinear PDE which we made at Heriot–Watt University in Edinburgh during the years 1999–2002. Then the second author was a post-doc, supported by two EPSRC grants and by the university’s funds. We are thankful to the two institutions for the financial support. This book is a much extended version of the lecture notes [Kuk06a] for a course which the first author taught at the Mathematical Department of ETH–Zürich during the winter term of the year 2004/05. He thanks the Forschungsinstitut at ETH for the hospitality and for help in preparation the lecture notes.

---

<sup>5</sup>But the method applies to Eq. (0.4) with  $m > 1$  if we add in the right-hand side a strong nonlinear damping  $-|u|^{2m'}u$ ,  $m' \geq m$ .

# Chapter 1

## Preliminaries

### 1.1 Function spaces

#### 1.1.1 Functions of the space variables

Let  $Q$  be a domain in  $\mathbb{R}^d$  (i.e., a connected open subset of  $\mathbb{R}^d$ ) or the torus  $\mathbb{T}^d = \mathbb{R}^d/2\pi\mathbb{Z}^d$ . We shall say that a domain  $Q$  is *Lipschitz* if its boundary  $\partial Q$  is locally Lipschitz<sup>1</sup>. We shall need *Lebesgue* and *Sobolev* spaces on  $Q$  and some embedding and interpolation theorems.

#### Lebesgue spaces

We denote by  $L^p(Q; \mathbb{R}^n)$ ,  $1 \leq p \leq \infty$ , the usual Lebesgue space of vector-valued functions and abbreviate  $L^p(Q; \mathbb{R}) = L^p(Q)$ . We write  $\langle \cdot, \cdot \rangle$  for the  $L_2$  scalar product and  $|\cdot|_p$  for the standard norm in  $L^p(Q; \mathbb{R}^n)$ .

#### Sobolev spaces

We denote by  $C_0^\infty(Q; \mathbb{R}^n)$  the space of infinitely smooth functions  $\varphi : Q \rightarrow \mathbb{R}^n$  with compact support. Let  $u$  and  $v$  be two locally integrable scalar functions on  $Q$  and let  $\alpha = (\alpha_1, \dots, \alpha_d)$  be a multi-index. We say that  $v$  is the  $\alpha^{\text{th}}$  weak partial derivative of  $u$  if

$$\int_Q u D^\alpha \varphi \, dx = (-1)^{|\alpha|} \int_Q v \varphi \, dx \quad \text{for all } \varphi \in C_0^\infty(Q; \mathbb{R}),$$

where  $|\alpha| := \alpha_1 + \dots + \alpha_d$  and  $D^\alpha = \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d}$ . In this case, we write  $D^\alpha u = v$ .

Let  $m \geq 0$  be an integer. The space  $H^m(Q, \mathbb{R}^n)$  consists of all locally integrable functions  $u : Q \rightarrow \mathbb{R}^n$  such that the derivative  $D^\alpha u$  exists in the weak sense for each multi-index  $\alpha$  with  $|\alpha| \leq m$  and belongs to  $L^2(Q; \mathbb{R}^n)$ . We

---

<sup>1</sup>This means that  $\partial Q$  can be represented locally as the graph of a Lipschitz function.

write  $H^m(Q; \mathbb{R}) = H^m(Q)$  and define the norm in  $H^m(Q; \mathbb{R}^n)$  as

$$\|u\|_m := \left( \sum_{|\alpha| \leq m} |D^\alpha u|_2^2 \right)^{1/2}.$$

In the case  $Q = \mathbb{T}^d$ , it is easy to define  $H^m(\mathbb{T}^d; \mathbb{R}^n)$  for all  $m \in \mathbb{R}$ . To this end, let us expand a function  $u \in L^2(\mathbb{T}^d, \mathbb{R}^n)$  into a Fourier series:

$$u(x) = \sum_{s \in \mathbb{Z}^d} u_s e^{isx}.$$

Define the following norm, which is equivalent to  $\|\cdot\|_m$  for non-negative integers  $m$ :

$$\|u\|_m = \left( \sum_{s \in \mathbb{Z}^d} (1 + |s|^2)^m |u_s|^2 \right)^{1/2}. \quad (1.1)$$

The space  $H^m(\mathbb{T}^d; \mathbb{R}^n)$  is defined as the closure of  $C^\infty(\mathbb{T}^d, \mathbb{R}^n)$  with respect to the norm  $\|\cdot\|_m$ . It is easy to see that if  $m \geq 0$  is an integer, then the two definitions of  $H^m(\mathbb{T}^d; \mathbb{R}^n)$  give the same function space. The following result is a simple consequence of the definition of  $\|\cdot\|_m$ .

**Lemma 1.1.1.** *For any  $m \in \mathbb{R}$  and any multi-index  $\alpha$ , the linear map  $D^\alpha$  is continuous from  $H^m(\mathbb{T}^d; \mathbb{R}^n)$  to  $H^{m-|\alpha|}(\mathbb{T}^d; \mathbb{R}^n)$ . Accordingly, the Laplace operator  $\Delta : H^m(\mathbb{T}^d; \mathbb{R}^n) \rightarrow H^{m-2}(\mathbb{T}^d; \mathbb{R}^n)$  is continuous. Similar assertions are true for any open domain  $Q \subset \mathbb{R}^d$  and any integer  $m \geq 0$ .*

Now let  $u \in H^m(\mathbb{T}^d; \mathbb{R}^n)$  be a function with zero mean value, that is,

$$\langle u \rangle := (2\pi)^{-d} \int_{\mathbb{T}^d} u(x) dx = 0, \quad (1.2)$$

where the integral is understood in the sense of the theory of distributions if  $m < 0$ . In this case, the first Fourier coefficient of  $u$  is zero,  $u_0 = 0$ , and therefore the norm

$$\|u\|_m = \left( \sum_{s \neq 0} |s|^{2m} |u_s|^2 \right)^{1/2}$$

is equivalent to (1.1) on the space

$$\dot{H}^m(\mathbb{T}^d; \mathbb{R}^n) = \{u \in H^m(\mathbb{T}^d; \mathbb{R}^n) : \langle u \rangle = 0\}.$$

In particular,  $\|u\|_1^2 = |\nabla u|_2^2$  is a norm on  $\dot{H}^1(\mathbb{T}^d; \mathbb{R}^n)$ .

Finally, let us define the *Sobolev space*  $H^m(Q; \mathbb{R}^n)$  in a bounded Lipschitz domain  $Q \subset \mathbb{R}^d$  for an arbitrary  $m \geq 0$ . Namely, without loss of generality, we can assume that  $Q \subset \mathbb{T}^d$ . We shall say that a function  $u \in L^2(Q, \mathbb{R}^n)$  belongs to  $H^m(Q; \mathbb{R}^n)$  if there is a function  $\tilde{u} \in \dot{H}^m(\mathbb{T}^d; \mathbb{R}^n)$  whose restriction to  $Q$  coincides with  $u$ . In this case, we define  $\|u\|_m$  as the infimum of  $\|\tilde{u}\|_m$  over all possible extensions  $\tilde{u} \in \dot{H}^m(\mathbb{T}^d; \mathbb{R}^n)$  for  $u$ .

**Property 1.1.2.** *Sobolev Embeddings.* Let  $Q$  be either a Lipschitz domain in  $\mathbb{R}^d$  or the torus  $\mathbb{T}^d$ .

1. If  $m \leq \frac{d}{2}$  and  $2 \leq q \leq \frac{2d}{d-2m}$ ,  $q < \infty$ , then

$$H^m(Q; \mathbb{R}^n) \subset L^q(Q; \mathbb{R}^n). \quad (1.3)$$

2. If  $m \geq \frac{d}{2} + \alpha$  with  $0 < \alpha < 1$ , then

$$H^m(Q; \mathbb{R}^n) \subset C_b^\alpha(Q; \mathbb{R}^n), \quad (1.4)$$

where  $C_b^\alpha(Q)$  denotes the space of functions that are bounded and Hölder continuous with the exponent  $\alpha$ . In particular, if  $m > \frac{d}{2}$ , then  $H^m(Q; \mathbb{R}^n)$  is continuously embedded into the space  $C_b(Q; \mathbb{R}^n)$  of bounded continuous functions.

3. If  $Q$  is bounded, then we have the compact embedding

$$H^{m_1}(Q; \mathbb{R}^n) \Subset H^{m_2}(Q; \mathbb{R}^n) \quad \text{for } m_1 > m_2. \quad (1.5)$$

It follows that embeddings (1.3) and (1.4) are compact for  $q < \frac{2d}{d-2m}$  and  $m > \frac{d}{2} + \alpha$ , respectively.

**Property 1.1.3.** *Duality.* The spaces  $H^m(\mathbb{T}^d; \mathbb{R}^n)$  and  $H^{-m}(\mathbb{T}^d; \mathbb{R}^n)$  are dual with respect to the  $L^2$ -scalar product  $\langle \cdot, \cdot \rangle$ . That is,

$$\|u\|_m = \sup_v |\langle u, v \rangle| \quad \text{for any } u \in C^\infty(\mathbb{T}^d; \mathbb{R}^n), \quad (1.6)$$

where the supremum is taken over all  $v \in C^\infty(\mathbb{T}^d; \mathbb{R}^n)$  such that  $\|v\|_{-m} \leq 1$ . Relation (1.6) implies that the scalar product in  $L^2$  extends to a continuous bilinear map from  $H^m(\mathbb{T}^d; \mathbb{R}^n) \times H^{-m}(\mathbb{T}^d; \mathbb{R}^n)$  to  $\mathbb{R}$ .

**Property 1.1.4.** *Interpolation inequality.* Let  $Q \subset \mathbb{R}^d$  be a Lipschitz domain, let  $a < b$  be non-negative integers, and let  $0 \leq \theta \leq 1$  be a constant. Then

$$\|u\|_{\theta a + (1-\theta)b} \leq \|u\|_a^\theta \|u\|_b^{1-\theta} \quad \text{for any } u \in H^b(Q; \mathbb{R}^n). \quad (1.7)$$

In the case of the torus, inequality (1.7) holds for any real numbers  $a < b$  and any  $\theta \in [0, 1]$ .

*Proof for the case of a torus.* We have

$$\begin{aligned} \|u\|_{\theta a + (1-\theta)b}^2 &= \sum_{s \in \mathbb{Z}^d} (1 + |s|^2)^{\theta a + (1-\theta)b} |u_s|^2 \\ &= \sum_{s \in \mathbb{Z}^d} \left( (1 + |s|^2)^{\theta a} |u_s|^{2\theta} \right) \left( (1 + |s|^2)^{(1-\theta)b} |u_s|^{2(1-\theta)} \right) \\ &\leq \left( \sum_{s \in \mathbb{Z}^d} (1 + |s|^2)^a |u_s|^2 \right)^\theta \left( \sum_{s \in \mathbb{Z}^d} (1 + |s|^2)^b |u_s|^2 \right)^{1-\theta}, \end{aligned}$$

where we used Hölder's inequality in the last step.  $\square$

*Example 1.1.5.* Let  $Q$  be either a Lipschitz domain in  $\mathbb{R}^2$  or the torus  $\mathbb{T}^2$ . Then the Sobolev embedding (1.3), with  $m = 1/2$  and  $q = 4$ , and the interpolation inequality (1.6) with  $a = 0$ ,  $b = 1$  and  $\theta = \frac{1}{2}$  imply that

$$|u|_4 \leq C_1 \|u\|_{1/2} \leq C_2 \sqrt{|u|_2 \|u\|_1} \quad \text{for any } u \in H^1(Q; \mathbb{R}^n). \quad (1.8)$$

This is *Ladyzhenskaya's inequality*.

A proof of Properties 1.1.2 – 1.1.4 can be found in [BIN79, Ste70, Tay97].

### 1.1.2 Functions of space and time variables

Solutions of the equations mentioned in the introduction are functions depending on the time  $t$  and the space variables  $x$ . We fix any  $T > 0$  and view a solution  $u(t, x)$  with  $0 \leq t \leq T$  as a map

$$[0, T] \longrightarrow \text{“space of functions of } x\text{”}, \quad t \mapsto u(t, \cdot).$$

Let us introduce corresponding functional spaces.

For a Banach space  $X$ , we denote by  $C(0, T; X)$  the space of continuous functions  $u : [0, T] \rightarrow X$  and endow it with the norm

$$\|u\|_{C(0, T; X)} = \sup_{0 \leq t \leq T} \|u(t)\|_X,$$

where  $\|\cdot\|_X$  stands for the norm in  $X$ . We denote by  $S(0, T; X)$  the space of functions of the form

$$u(t) = \sum_{k=1}^N u_k \mathbb{I}_{\Gamma_k}(t),$$

where  $N \geq 1$  is an integer depending on the function,  $u_k \in X$  are some vectors,  $\Gamma_k$  are Borel-measurable subsets of  $[0, T]$  (see Subsection 1.2.1), and  $\mathbb{I}_\Gamma$  stands for the indicator function of  $\Gamma$ . If  $X$  is separable, then for  $p \in [1, \infty]$  define  $L^p(0, T; X)$  as the completion of the space  $S(0, T; X)$  with respect to the norm

$$\|u\|_{L^p(0, T; X)} = \begin{cases} \left( \int_0^T \|u(t)\|_X^p dt \right)^{1/p} & \text{for } 1 \leq p < \infty, \\ \text{ess sup}_{0 \leq t \leq T} \|u(t)\|_X & \text{for } p = \infty. \end{cases}$$

Note that, in view of Fubini's theorem, we have

$$L^p(0, T; L^p(Q; \mathbb{R}^n)) = L^p((0, T) \times Q; \mathbb{R}^n) \quad \text{for } p < \infty.$$

A more detailed discussion of these spaces can be found in [Lio69, Yos95].

We shall also need the space of continuous functions on an interval with range in a metric space. Namely, let  $J \subset \mathbb{R}$  be a closed interval and let  $X$  be a Polish space, that is, a complete separable metric space with a distance  $\text{dist}_X$ .



We denote by  $C(J; X)$  the space of continuous functions from  $J$  to  $X$ . When  $J$  is bounded,  $C(J; X)$  is a Polish space with respect to the distance

$$\|u - v\|_{C(J; X)} = \max_{t \in J} \text{dist}_X(u(t), v(t)).$$

In the case of an unbounded interval  $J$ , we endow  $C(J; X)$  with the metric

$$\text{dist}(u, v) = \sum_{k=1}^{\infty} 2^{-k} \frac{\|u - v\|_{C(J_k; X)}}{1 + \|u - v\|_{C(J_k; X)}}, \quad (1.9)$$

where  $J_k = J \cap [-k, k]$ . Note that, for a sequence  $\{u_j\} \subset C(J, X)$ , we have  $\text{dist}(u_j, u) \rightarrow 0$  as  $j \rightarrow \infty$  if and only if  $\|u_j - u\|_{C(J_k; X)}$  for each  $k$ . That is, (1.9) is the *metric of uniform convergence on bounded intervals*. When  $J = \mathbb{Z}$  (or  $J$  is a countable subset of  $\mathbb{Z}$ ), formula (1.9) may be used to define a distance on  $X^J$ . This distance corresponds to the *Tikhonov topology* on  $X^J$ .

*Exercise 1.1.6.* Prove that if  $J \subset \mathbb{R}$  is an unbounded closed interval, then  $C(J; X)$  is a Polish space. Prove also that if  $X$  is a separable Banach space, then  $C(J; X)$  is a separable Fréchet space.

## 1.2 Basic facts from the measure theory

In this section, we first recall the concept of a  $\sigma$ -algebra, together with some related definitions, and formulate without proof three standard results on the passage to the limit under Lebesgue's integral. We next discuss various metrics on the space of probability measures on a Polish space and establish some results on (maximal) couplings of measures.

### 1.2.1 $\sigma$ -algebras and measures

Let  $\Omega$  be an arbitrary set and let  $\mathcal{F}$  be a family of subsets of  $\Omega$ . Recall that  $\mathcal{F}$  is called a  $\sigma$ -algebra if it contains the sets  $\emptyset$  and  $\Omega$ , and is invariant under taking the complement and countable union of its elements. Any pair  $(\Omega, \mathcal{F})$  is called a *measurable space*. If  $(\Omega_i, \mathcal{F}_i)$ ,  $i = 1, 2$ , are measurable spaces, then a mapping  $f : \Omega_1 \rightarrow \Omega_2$  is said to be *measurable* if  $f^{-1}(\Gamma) \in \mathcal{F}_1$  for any  $\Gamma \in \mathcal{F}_2$ . If  $\mu$  is a (positive) measure on  $(\Omega_1, \mathcal{F}_1)$ , then its *image* under  $f$  is the measure  $f_*(\mu)$  on  $(\Omega_2, \mathcal{F}_2)$  defined by  $f_*(\mu)(\Gamma) = \mu(f^{-1}(\Gamma))$  for any  $\Gamma \in \mathcal{F}_2$ . Note that  $f_*$  is a linear mapping on the space of positive measures:

$$f_*(c_1\mu_1 + c_2\mu_2) = c_1f_*(\mu_1) + c_2f_*(\mu_2) \quad \text{for any } c_1, c_2 \geq 0.$$

The *product* of two measurable spaces  $(\Omega_i, \mathcal{F}_i)$ ,  $i = 1, 2$ , is defined as the set  $\Omega_1 \times \Omega_2$  endowed with the minimal  $\sigma$ -algebra  $\mathcal{F}_1 \otimes \mathcal{F}_2$  generated by subsets of the form  $\Gamma_1 \times \Gamma_2$  with  $\Gamma_i \in \mathcal{F}_i$ . The product of finitely or countably many  $\sigma$ -algebras is defined in a similar way.

Given a probability measure  $\mu$  on a measurable space  $(\Omega, \mathcal{F})$ , we denote by  $\mathcal{N}_\mu$  the family of subsets  $A \subset \Omega$  such that  $A \subset B$  for some  $B \in \mathcal{F}$  with

$\mu(B) = 0$ . A  $\sigma$ -algebra  $\mathcal{F}$  is said to be *complete* with respect to a measure  $\mu$  if it contains all sets from  $\mathcal{N}_\mu$ . The *completion of  $\mathcal{F}$  with respect to  $\mu$*  is defined as the minimal  $\sigma$ -algebra generated by  $\mathcal{F} \cup \mathcal{N}_\mu$  and is denoted by  $\mathcal{F}_\mu$ . This is the minimal complete  $\sigma$ -algebra which contains  $\mathcal{F}$ . A subset  $\Gamma \subset \Omega$  is said to be *universally measurable* if it belongs to  $\mathcal{F}_\mu$  for any probability measure  $\mu$  on  $(\Omega, \mathcal{F})$ . If  $\mu$  is a measure on  $(\Omega_1, \mathcal{F}_1)$ ,  $\mathcal{F}_1$  is complete with respect to  $\mu$ , and a map  $f : \Omega_1 \rightarrow \Omega_2$  is a  $\mu$ -almost sure limit of a sequence of measurable maps, then  $f$  is measurable.

Now let  $X$  be a Polish space, that is, a complete separable metric space. We denote by  $\text{dist}_X$  the metric on  $X$ . The *Borel  $\sigma$ -algebra*  $\mathcal{B} = \mathcal{B}(X)$  is defined as the minimal  $\sigma$ -algebra containing all open subsets of  $X$ . The pair  $(X, \mathcal{B}(X))$  is called a *measurable Polish space*. If  $X_1$  and  $X_2$  are Polish spaces, then a map  $f : X_1 \rightarrow X_2$  is said to be *measurable* if  $f^{-1}(\Gamma) \in \mathcal{B}(X_1)$  for any  $\Gamma \in \mathcal{B}(X_2)$ . In particular, a function  $f : X \rightarrow \mathbb{R}$  is called *measurable* if it is measurable with respect to the Borel  $\sigma$ -algebras on  $X$  and  $\mathbb{R}$ . An important property of Polish spaces is that any probability measure on it is *regular*. Namely, *Ulam's theorem* says that, for any probability measure  $\mu$  on a Polish space  $X$  and any  $\varepsilon > 0$ , there is a compact set  $K \subset X$  such that  $\mu(K) \geq 1 - \varepsilon$ . A proof of this result can be found in [Dud02] (see Theorem 7.1.4).

Recall that, for any probability measure  $\mathbb{P}$  on a measurable space  $(\Omega, \mathcal{F})$ , the triple  $(\Omega, \mathcal{F}, \mathbb{P})$  is called a probability space. A probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  is said to be *complete* if  $\mathcal{F}_\mathbb{P} = \mathcal{F}$ . We shall often consider a probability space together with a family  $\{\mathcal{F}_t \subset \mathcal{F}\}$  of  $\sigma$ -algebras that depend on a parameter  $t$  varying either in  $\mathbb{R}_+$  or in  $\mathbb{Z}_+$ . In this case, we shall always assume that  $\mathcal{F}_t$  is non-decreasing with respect to  $t$ . The quadruple  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  is called a *filtered probability space*. We shall say that  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  satisfies the *usual hypotheses* if  $(\Omega, \mathcal{F}, \mathbb{P})$  is complete and  $\mathcal{F}_t$  contains all  $\mathbb{P}$ -null sets of  $\mathcal{F}$ .

If  $X$  is a Polish space, then an  $X$ -valued *random variable* is a measurable map  $\xi$  from a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  into  $X$ . The *law* or the *distribution* of  $\xi$  is defined as the image of  $\mathbb{P}$  under  $\xi$  and is denoted by  $\mathcal{D}(\xi)$ , i.e.,  $\mathcal{D}(\xi) = \xi_*(\mathbb{P})$ . If we need to emphasise that the distribution of a random variable is considered with respect to a probability measure  $\mu$ , then we write  $\mathcal{D}_\mu(\xi)$ . An  $X$ -valued *random process* is defined as a collection of a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a family of  $X$ -valued random variables  $\{\xi_t\}$  on  $\Omega$  (where  $t$  varies in  $\mathbb{R}_+$  or  $\mathbb{Z}_+$ ). If the underlying probability space is equipped with a filtration  $\mathcal{F}_t$ , then we shall say that the process  $\xi_t$  is *adapted to  $\mathcal{F}_t$*  if  $\xi_t$  is  $\mathcal{F}_t$ -measurable for any  $t \geq 0$ . Finally, a random process  $\xi_t$  defined on a filtered probability space  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  is said to be *progressively measurable* if for any  $t \geq 0$  the map  $(s, \omega) \mapsto \xi_s(\omega)$  from  $[0, t] \times \Omega$  to  $X$  is measurable. It is clear that if  $t$  varies in  $\mathbb{Z}_+$ , then these two concepts coincide.

### 1.2.2 Convergence of integrals

In what follows, we shall systematically use well-known results on the passage to the limit under Lebesgue's integrals. For the reader's convenience, we state them here without proofs, referring the reader to Section 4.3 of [Dud02].

Let  $(\Omega, \mathcal{F})$  be a measurable space, let  $\mu$  be an arbitrary  $\sigma$ -finite measure on it (so  $\mu(\Omega) \leq \infty$ ), and let  $f_n : \Omega \rightarrow \mathbb{C}$  be a sequence of integrable functions. The following result called *Lebesgue's theorem on dominated convergence* gives a sufficient condition for the convergence of the integrals of  $f_n$  to that of the limit function.

**Theorem 1.2.1.** *Assume that  $\{f_n\}_{n \geq 1}$  is a sequence of functions that converge  $\mu$ -almost surely and satisfy the inequality*

$$|f_n(\omega)| \leq g(\omega) \quad \text{for } \mu\text{-almost every } \omega \in \Omega, \quad (1.10)$$

where  $g : \Omega \rightarrow \mathbb{R}_+$  is a  $\mu$ -integrable function. Then

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n d\mu = \int_{\Omega} \left( \lim_{n \rightarrow \infty} f_n \right) d\mu. \quad (1.11)$$

In the case when the functions  $f_n$  are real-valued and form a monotone sequence, bound (1.10) can be replaced by a weaker condition, which a posteriori turns out to be equivalent to the former. Namely, we have the following *monotone convergence theorem*.

**Theorem 1.2.2.** *Let  $f_n : \Omega \rightarrow \mathbb{R}$  be a non-decreasing (or non-increasing) sequence that converges  $\mu$ -almost surely and satisfies the condition*

$$\sup_{n \geq 1} \left| \int_{\Omega} f_n d\mu \right| < \infty.$$

Then relation (1.11) holds.

Finally, the following result called *Fatou's lemma* is useful when estimating the integral of the limit for a sequence of non-negative functions.

**Theorem 1.2.3.** *Let  $f_n : \Omega \rightarrow \mathbb{R}_+$  be an arbitrary sequence of  $\mu$ -integrable functions. Then*

$$\int_{\Omega} \left( \liminf_{n \rightarrow \infty} f_n \right) d\mu \leq \liminf_{n \rightarrow \infty} \int_{\Omega} f_n d\mu.$$

In particular, the three theorem above apply if  $\Omega$  is the set  $\mathbb{N}$  of non-negative integers with the counting measure. In this case, they describe passage to the limit for sums of infinite series.

### 1.2.3 Metrics on the space of probabilities and convergence of measures

In what follows, we denote by  $X$  a Polish space with a metric  $d_X$ . Define  $C_b(X)$  as the space of bounded continuous functions  $f : X \rightarrow \mathbb{R}$  endowed with the norm

$$\|f\|_{\infty} = \sup_{u \in X} |f(u)|,$$

and denote by  $L_b(X)$  the space of bounded Lipschitz functions on  $X$ . That is, of functions  $f \in C_b(X)$  for which

$$\text{Lip}(f) := \sup_{u_1, u_2 \in X} \frac{|f(u_1) - f(u_2)|}{\text{dist}_X(u_1, u_2)} < \infty.$$

The space  $L_b(X)$  is endowed with the norm

$$\|f\|_L = \|f\|_\infty + \text{Lip}(f).$$

Note that  $C_b(X)$  and  $L_b(X)$  are Banach spaces with respect to the corresponding norms. The following exercise summarises some further properties of these spaces.

*Exercise 1.2.4.* Let  $X$  be a Polish space.

- (i) Prove that  $C_b(X)$  is separable if and only if  $X$  is compact.
- (ii) Prove that  $L_b(X)$  is not separable for the space  $X = [0, 1]$  with the usual metric.

*Hint:* To prove that  $C_b(X)$  is separable for a compact metric space  $X$ , use the existence of a finite  $\varepsilon$ -net and a partition of unity on  $X$ . To show that if  $X$  is not compact, then  $C_b(X)$  is not separable, use the existence of a sequence  $\{x_k\} \subset X$  such that  $\text{dist}_X(x_k, x_m) \geq \varepsilon > 0$ . Finally, to prove (ii), construct a continuum  $\{\varphi_\alpha\} \subset L^\infty(X)$  such that the distance between any two functions is equal to 1, and use the integrals of  $\varphi_\alpha$ .

Let us denote by  $\mathcal{P}(X)$  the set of probability measures on  $(X, \mathcal{B}(X))$  and by  $\mathcal{P}_1(X)$  the subset of those measures  $\mu \in \mathcal{P}(X)$  for which

$$\mathbf{m}_1(\mu) := \int_X \text{dist}_X(u, u_0) \mu(du) < \infty, \quad (1.12)$$

where  $u_0 \in X$  is an arbitrary point. The triangle inequality implies that the class  $\mathcal{P}_1(X)$  does not depend on the choice of  $u_0$ . We shall need the following three metrics.

**Total variation distance :**

$$\|\mu_1 - \mu_2\|_{\text{var}} := \frac{1}{2} \sup_{\substack{f \in C_b(X) \\ \|f\|_\infty \leq 1}} |(f, \mu_1) - (f, \mu_2)|, \quad \mu_1, \mu_2 \in \mathcal{P}(X). \quad (1.13)$$

This is the distance induced on  $\mathcal{P}(X)$  by its embedding into the space dual to  $C_b(X)$ . It can be extended to probability measures on an *arbitrary* measurable space; see Remark 1.2.8 below.

**Dual-Lipschitz distance :**

$$\|\mu_1 - \mu_2\|_L^* := \sup_{\substack{f \in L_b(X) \\ \|f\|_L \leq 1}} |(f, \mu_1) - (f, \mu_2)|, \quad \mu_1, \mu_2 \in \mathcal{P}(X). \quad (1.14)$$

This is the distance induced on  $\mathcal{P}(X)$  by its embedding into the space dual to  $L_b(X)$ .

**Kantorovich distance :**

$$\|\mu_1 - \mu_2\|_K := \sup_{\substack{f \in L_b(X) \\ \text{Lip}(f) \leq 1}} |(f, \mu_1) - (f, \mu_2)|, \quad \mu_1, \mu_2 \in \mathcal{P}_1(X). \quad (1.15)$$

*Exercise 1.2.5.* Show that the symmetric functions (1.13) – (1.15) define metrics on the sets  $\mathcal{P}(X)$  and  $\mathcal{P}_1(X)$ . *Hint:* The only non-trivial point is that if measures  $\mu_1$  and  $\mu_2$  satisfies the relation  $\|\mu_1 - \mu_2\|_L^* = 0$ , then  $\mu_1 = \mu_2$ . This can be done with the help of monotone class technique; see Corollary 7.1.3 in the Appendix.

An immediate consequence of definitions (1.13) – (1.15) and the inequalities  $\|f\|_\infty \leq \|f\|_L$  and  $\text{Lip}(f) \leq \|f\|_L$  is that

$$\|\mu_1 - \mu_2\|_L^* \leq 2\|\mu_1 - \mu_2\|_{\text{var}} \quad \text{for } \mu_1, \mu_2 \in \mathcal{P}(X), \quad (1.16)$$

$$\|\mu_1 - \mu_2\|_L^* \leq \|\mu_1 - \mu_2\|_K \quad \text{for } \mu_1, \mu_2 \in \mathcal{P}_1(X). \quad (1.17)$$

Furthermore, if the space  $X$  is bounded, that is, there is an element  $u_0 \in X$  and a constant  $d_0 > 0$  such that

$$\text{dist}_X(u, u_0) \leq d_0 \quad \text{for all } u \in X,$$

then, for any function  $f \in C_b(X)$  vanishing at  $u_0 \in X$ , we have

$$\|f\|_\infty \leq d_0 \text{Lip}(f),$$

where the right-hand side may be infinite. It follows that in this case

$$\|\mu_1 - \mu_2\|_K \leq 2d_0\|\mu_1 - \mu_2\|_{\text{var}} \quad \text{for } \mu_1, \mu_2 \in \mathcal{P}_1(X).$$

It turns out that the distance  $\|\cdot\|_L^*$  is equivalent to the one obtained by replacing  $L_b(X)$  in (1.14) with the space of bounded Hölder-continuous functions. Namely, for  $\gamma \in (0, 1)$  we denote by  $C_b^\gamma(X)$  the space of continuous functions  $f : X \rightarrow \mathbb{R}$  such that

$$|f|_\gamma := \|f\|_\infty + \sup_{0 < \text{dist}_X(u, v) \leq 1} \frac{|f(u) - f(v)|}{\text{dist}_X(u, v)^\gamma} < \infty.$$

Let us set

$$\|\mu_1 - \mu_2\|_\gamma^* := \sup_{\substack{f \in C_b^\gamma(X) \\ |f|_\gamma \leq 1}} |(f, \mu_1) - (f, \mu_2)|, \quad \mu_1, \mu_2 \in \mathcal{P}(X). \quad (1.18)$$

**Proposition 1.2.6.** *For any  $\gamma \in (0, 1)$  and  $\mu_1, \mu_2 \in \mathcal{P}(X)$ , we have*

$$\|\mu_1 - \mu_2\|_L^* \leq \|\mu_1 - \mu_2\|_\gamma^* \leq 5(\|\mu_1 - \mu_2\|_L^*)^{\frac{1}{2-\gamma}}.$$

*Proof.* The lower bound of the inequality is obvious, and therefore we shall confine ourselves to the proof of the upper bound. For any continuous function  $f : X \rightarrow \mathbb{R}$ , we define an approximation for it by the relation

$$f_\varepsilon(u) = \inf_{v \in X} (\varepsilon^{-1}d(u, v) + f(v)), \quad u \in X, \quad (1.19)$$

where  $\varepsilon > 0$  is an arbitrary constant. It is a matter of direct verification to show that if  $f \in C_b^\gamma(X)$  and  $\|f\|_\gamma \leq 1$ , then

$$\|f_\varepsilon\|_L \leq 1 + \varepsilon^{-1}, \quad 0 \leq f(u) - f_\varepsilon(u) \leq \varepsilon^{\frac{1}{1-\gamma}} \quad \text{for } u \in X. \quad (1.20)$$

We now fix  $\delta > 0$  and find a function  $f \in C_b^\gamma(X)$  with  $\|f\|_\gamma \leq 1$  such that

$$\|\mu_1 - \mu_2\|_\gamma^* \leq |(f, \mu_1) - (f, \mu_2)| + \delta. \quad (1.21)$$

It follows from (1.20) that, for any  $\varepsilon > 0$ , we have

$$\begin{aligned} |(f, \mu_1) - (f, \mu_2)| &\leq |(f_\varepsilon - f, \mu_1)| + |(f_\varepsilon - f, \mu_2)| + |(f_\varepsilon, \mu_1) - (f_\varepsilon, \mu_2)| \\ &\leq 2\varepsilon^{\frac{1}{1-\gamma}} + (1 + \varepsilon^{-1})\|\mu_1 - \mu_2\|_L^*. \end{aligned}$$

Choosing  $\varepsilon = (\|\mu_1 - \mu_2\|_L^*)^{\frac{1-\gamma}{2-\gamma}}$  and noting that  $\|\mu_1 - \mu_2\|_L^* \leq 2$ , we get

$$|(f, \mu_1) - (f, \mu_2)| \leq 5(\|\mu_1 - \mu_2\|_L^*)^{\frac{1}{2-\gamma}}.$$

Combining this with (1.21) and recalling that  $\delta > 0$  was arbitrary, we arrive at the required assertion.  $\square$

The following proposition gives an alternative description of the total variation distance and provides some formulas for calculating it.

**Proposition 1.2.7.** *For any  $\mu_1, \mu_2 \in \mathcal{P}(X)$ , we have*

$$\|\mu_1 - \mu_2\|_{\text{var}} = \sup_{\Gamma \in \mathcal{B}(X)} |\mu_1(\Gamma) - \mu_2(\Gamma)|. \quad (1.22)$$

*Furthermore, if  $\mu_1$  and  $\mu_2$  are absolutely continuous with respect to a given measure  $m \in \mathcal{P}(X)$ , then*

$$\|\mu_1 - \mu_2\|_{\text{var}} = \frac{1}{2} \int_X |\rho_1(u) - \rho_2(u)| dm = 1 - \int_X (\rho_1 \wedge \rho_2)(u) dm, \quad (1.23)$$

where  $\rho_i(u)$  is the density of  $\mu_i$  with respect to  $m$ .

*Remark 1.2.8.* Let us note that a measure  $m \in \mathcal{P}(X)$  with respect to which  $\mu_1$  and  $\mu_2$  are absolutely continuous always exists. For instance, we can take  $m = \frac{1}{2}(\mu_1 + \mu_2)$ . Furthermore, relation (1.22) enables one to extend the definition of the *total variation distance* to an arbitrary measurable space.

*Proof of Proposition 1.2.7. Step 1.* Let us denote by  $\|\mu_1 - \mu_2\|'_{\text{var}}$  the right-hand side of (1.22) and show that relation (1.23) is true for it. Setting  $\rho = \rho_1 \wedge \rho_2$  and integrating the obvious relation

$$\frac{1}{2} |\rho_1 - \rho_2| = \frac{1}{2} (\rho_1 + \rho_2) - \rho$$

over  $X$  with respect to  $m$ , we obtain the second equality in (1.23).

We now show that

$$\|\mu_1 - \mu_2\|'_{\text{var}} \leq 1 - \int_X \rho(u) dm. \quad (1.24)$$

Let us define the set  $Y = \{u \in X : \rho_1(u) > \rho_2(u)\}$ . Since  $\rho = \rho_2$  on  $Y$ , we have

$$\begin{aligned} \mu_1(\Gamma) - \mu_2(\Gamma) &= \int_{\Gamma} (\rho_1 - \rho_2) dm \leq \int_{\Gamma \cap Y} (\rho_1 - \rho_2) dm \\ &= \int_{\Gamma \cap Y} (\rho_1 - \rho) dm \leq \int_X (\rho_1 - \rho) dm = 1 - \int_X \rho(u) dm \end{aligned}$$

for any  $\Gamma \in \mathcal{B}(X)$ . By symmetry, this inequality implies (1.24).

To prove the converse inequality, we note that  $\rho = \rho_1$  on  $Y^c$  and  $\rho = \rho_2$  on  $Y$ . It follows that

$$\begin{aligned} \mu_1(Y) - \mu_2(Y) &= \int_Y (\rho_1 - \rho_2) dm \\ &= \left( \int_Y \rho_1 dm + \int_{Y^c} \rho dm \right) - \left( \int_Y \rho_2 dm + \int_{Y^c} \rho dm \right) \\ &= \left( \int_Y \rho_1 dm + \int_{Y^c} \rho_1 dm \right) - \left( \int_Y \rho dm + \int_{Y^c} \rho dm \right) \\ &= 1 - \int_X \rho dm. \end{aligned}$$

This completes the proof of (1.23) for  $\|\mu_1 - \mu_2\|'_{\text{var}}$ .

*Step 2.* We now prove (1.22). Using Step 1, for any  $f \in C_b(X)$  with  $\|f\|_{\infty} \leq 1$ , we derive

$$|(f, \mu_1) - (f, \mu_2)| \leq \int_X |f(u)(\rho_1(u) - \rho_2(u))| dm \leq 2 \|\mu_1 - \mu_2\|'_{\text{var}},$$

which implies that

$$\|\mu_1 - \mu_2\|_{\text{var}} \leq \|\mu_1 - \mu_2\|'_{\text{var}}.$$

To establish the converse inequality, let us consider a function  $f(u)$  that is equal to 1 on  $Y$  and to  $-1$  on  $Y^c$ . We have

$$\begin{aligned} (f, \mu_1) - (f, \mu_2) &= \int_X f(u)(\rho_1(u) - \rho_2(u)) dm \\ &= \int_X |\rho_1(u) - \rho_2(u)| dm = 2 \|\mu_1 - \mu_2\|'_{\text{var}}, \end{aligned} \quad (1.25)$$

where we used the first relation in (1.23). To complete the proof of (1.22), let us choose a sequence  $f_n \in C_b(X)$  such that (see Exercise 1.2.12 below)

$$\begin{aligned} \|f_n\|_\infty &\leq 1 \quad \text{for all } n \geq 1, \\ f_n(u) &\rightarrow f(u) \quad \text{as } n \rightarrow \infty \text{ for } m\text{-a.e. } u \in X. \end{aligned}$$

It is easy to see that the difference  $(f_n, \mu_1) - (f_n, \mu_2)$  tends to the left-hand side of (1.25) as  $n \rightarrow \infty$ . This completes the proof of the proposition.  $\square$

*Exercise 1.2.9.* Analysing the proof of relation (1.23), show that it remains valid for any positive Borel measure  $m$  (that is, we no longer require that  $m(X) = 1$ ) with respect to which  $\mu_1$  and  $\mu_2$  are absolutely continuous.

*Exercise 1.2.10.* Let  $X$  be a Polish space and let  $L^\infty(X)$  be the space of bounded measurable functions endowed with the norm  $\|\cdot\|_\infty$ . Prove that for any measures  $\mu_1, \mu_2 \in \mathcal{P}(X)$  we have (cf. (1.13))

$$\|\mu_1 - \mu_2\|_{\text{var}} := \frac{1}{2} \sup_{\substack{f \in L^\infty(X) \\ \|f\|_\infty \leq 1}} |(f, \mu_1) - (f, \mu_2)|. \quad (1.26)$$

Two measures  $\mu, \nu \in \mathcal{P}(X)$  are said to be *mutually singular* if there is a Borel subset  $A \subset X$  such that  $\mu(A) = 1$  and  $\nu(A) = 0$ . Relation (1.23) of Proposition 1.2.7 implies the following result.

**Corollary 1.2.11.** *Two measures  $\mu, \nu \in \mathcal{P}(X)$  are mutually singular if and only if  $\|\mu - \nu\|_{\text{var}} = 1$ .*

*Exercise 1.2.12.* Let  $X$  be a Polish space and let  $m \in \mathcal{P}(X)$ . Show that for any bounded measurable function  $f: X \rightarrow \mathbb{R}$  there is a sequence of continuous functions that are uniformly bounded by  $\|f\|_\infty$  and converge to  $f$  for  $m$ -almost all  $u \in X$ . *Hint:* It suffices to prove that  $f$  can be approximated by continuous functions in the space  $L^1(X, m)$ . To do this, show that any bounded measurable function can be approximated (in the sense of uniform convergence) by finite linear combinations of indicator functions and that the indicator function of any measurable set can be approximated (in the sense of convergence in  $L^1(X, m)$ ) by bounded continuous functions. For the latter property, one could use Ulam's theorem on interior regularity of Borel measures; see Section 1.2.1.

*Exercise 1.2.13.* (i) Let  $(\Omega_i, \mathcal{F}_i)$ ,  $i = 1, 2$ , be two measurable spaces and let  $f: \Omega_1 \rightarrow \Omega_2$  be a measurable mapping. Prove that, for any measures  $\mu_1, \mu_2$  on  $(\Omega_1, \mathcal{F}_1)$ ,

$$\|f_*(\mu_1) - f_*(\mu_2)\|_{\text{var}} \leq \|\mu_1 - \mu_2\|_{\text{var}};$$

see Remark 1.2.8 and relation (1.22) for the definition of the total variation distance in the case of measurable spaces.



- (ii) Let  $X$  be a Polish space and let  $\xi_1, \xi_2$  be two  $X$ -valued random variables defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  such that  $\xi_1 = \xi_2 \circ \Phi$  almost surely, where  $\Phi : \Omega \rightarrow \Omega$  is a measurable transformation. Show that

$$\|\mathcal{D}(\xi_1) - \mathcal{D}(\xi_2)\|_{\text{var}} \leq \|\mathbb{P} - \Phi_*(\mathbb{P})\|_{\text{var}}.$$

*Hint:* Use (i) with  $f = \xi_2$ ,  $\mu_2 = \mathbb{P}$ , and  $\mu_1 = \Phi_*(\mathbb{P})$ .

We shall say that a sequence  $\{\mu_k\} \subset \mathcal{P}(X)$  converges weakly to  $\mu \in \mathcal{P}(X)$  if

$$(f, \mu_k) \rightarrow (f, \mu) \quad \text{as } k \rightarrow \infty \quad (1.27)$$

for any  $f \in C_b(X)$ . In this case, we write  $\mu_k \rightarrow \mu$ . Recall that a family  $\{\mu_\alpha, \alpha \in \mathcal{A}\} \subset \mathcal{P}(X)$  is said to be *tight* in  $X$  if for any  $\varepsilon > 0$  there is a compact set  $K_\varepsilon \subset X$  such that  $\mu_\alpha(K_\varepsilon) \geq 1 - \varepsilon$  for any  $\alpha \in \mathcal{A}$ . In what follows, we shall need the following well-known result called *Prokhorov theorem*, which gives a necessary and sufficient condition for the compactness of a family of measures in the weak topology. Its proof can be found in [Dud02, Theorem 11.5.4].

**Theorem 1.2.14.** *Let  $X$  be a Polish space and let  $\{\mu_\alpha, \alpha \in \mathcal{A}\}$  be a family of probability measures on  $X$ . Then  $\{\mu_\alpha\}$  is relatively compact in the weak topology of  $\mathcal{P}(X)$  if and only if it is tight.*

The following result of fundamental importance shows, in particular, that the weak convergence of measures is equivalent to the convergence in the dual-Lipschitz distance.

**Theorem 1.2.15.** (i) *The set  $\mathcal{P}(X)$  endowed with the total variation distance is a complete metric space. Furthermore, a sequence  $\{\mu_k\} \subset \mathcal{P}(X)$  converges to a measure  $\mu$  in this space if and only if (1.27) holds uniformly in  $f \in C_b(X)$  with  $\|f\|_\infty \leq 1$ . In particular,  $\mathcal{P}(X)$  is naturally embedded in the dual space of  $C_b(X)$  as its closed subspace.*

- (ii) *The set  $\mathcal{P}(X)$  endowed with the dual-Lipschitz distance is a complete metric space. Furthermore, a sequence  $\{\mu_k\} \subset \mathcal{P}(X)$  converges to a measure  $\mu$  in this space if and only if either  $\{\mu_k\}$  converges weakly to  $\mu$  or (1.27) holds for any  $f \in L_b(X)$ .*

*Proof.* Assertion (i) follows easily from the definition and basic properties of the metric  $\|\cdot\|_{\text{var}}$ . Indeed, if  $\{\mu_k\} \subset \mathcal{P}(X)$  is a Cauchy sequence, then by (1.22) for each Borel set  $\Gamma \subset X$  there exists a limit  $\lim_{k \rightarrow \infty} \mu_k(\Gamma) =: \mu(\Gamma)$ . If we show that  $\mu$  is a probability measure, then it will be the limit of  $\{\mu_k\}$  for the total variation distance.

It is obvious that the mapping  $\Gamma \mapsto \mu(\Gamma)$  is additive and that  $\mu(\Gamma) = 1$ . To complete the proof of (i), it remains to establish the  $\sigma$ -additivity of  $\mu$ . To this end, it suffices to verify that, for any decreasing sequence  $\{\Gamma_n\}$  of Borel subsets in  $X$ , we have  $\mu(\cap_n \Gamma_n) = \lim_{n \rightarrow \infty} \mu(\Gamma_n)$ . However, this relation follows from (1.22) since each  $\mu_k$  is a measure. The second claim of (i) is an immediate consequence of (1.13).

We now prove assertion (ii). To simplify the presentation, we carry out the proof for compact metric spaces  $X$ , referring the reader to [Dud02, Chapter 11] for the general case.

We first show that if (1.27) holds for any  $f \in L_b(X)$ , then  $\mu_k \rightarrow \mu$  in the dual-Lipschitz distance. Indeed, using the compactness of  $X$ , it is easy to show that for any  $\varepsilon > 0$  the ball  $\{f \in L_b(X) : \|f\|_L \leq 1\}$  has a finite  $\varepsilon$ -net in  $C_b(X)$  that consists of functions of  $L_b(X)$ . Therefore convergence (1.27) for any  $f \in L_b(X)$  implies the convergence of  $\{\mu_k\}$  in the space  $\mathcal{P}(X)$  endowed with the dual-Lipschitz distance.

We now prove that if  $\mu_k \rightarrow \mu$  in the dual-Lipschitz distance, then (1.27) holds for any function  $f \in C_b(X)$ . Indeed, without loss of generality, we can assume that  $f \geq 0$ . The compactness of  $X$  implies that the function  $f_\varepsilon \in L_b(X)$  defined by (1.19) converges to  $f$  in the norm  $\|\cdot\|_\infty$  as  $\varepsilon \rightarrow 0$ . Furthermore,

$$|(f, \mu_k) - (f, \mu)| \leq |(f, \mu_k) - (f_\varepsilon, \mu_k)| + |(f_\varepsilon, \mu_k) - (f_\varepsilon, \mu)| + |(f_\varepsilon, \mu) - (f, \mu)|.$$

The first and third terms on the right-hand side of this inequality go to zero as  $\varepsilon \rightarrow 0$  uniformly in  $k$ , while the second can be made arbitrarily small by choosing a sufficiently large  $k$ .

Finally, let us prove that the space  $\mathcal{P}(X)$  endowed with the dual-Lipschitz distance is complete. Let  $\{\mu_k\} \subset \mathcal{P}(X)$  be a Cauchy sequence. Since  $X$  is a compact space, by Prokhorov's theorem we can find a subsequence  $\{\mu_{k_j}\}$  that converges weakly to a measure  $\mu \in \mathcal{P}(X)$ , that is, (1.27) holds for any  $f \in C_b(X)$ . If we show that the limiting measure  $\mu$  does not depend on the subsequence, we can conclude that the entire sequence converges to  $\mu$  weakly and, hence, in the space  $\mathcal{P}(X)$  as well.

Since  $\{\mu_k\}$  is a Cauchy sequence, for any  $f \in L_b(X)$  we have

$$\lim_{k \rightarrow \infty} (f, \mu_k) = \lim_{k_j \rightarrow \infty} (f, \mu_{k_j}) = (f, \mu).$$

Thus, the limiting measure is uniquely defined. This completes the proof of the theorem in the case of compact metric spaces.  $\square$

Let us emphasise that if  $\mu_k \rightarrow \mu$  weakly in  $\mathcal{P}(X)$ , then, in general, it is *not* true that

$$\mu_k(\Gamma) \rightarrow \mu(\Gamma) \quad \text{as } k \rightarrow \infty \quad (1.28)$$

for any  $\Gamma \in \mathcal{B}(X)$ . However, the well-known *portmanteau theorem* claims that  $\mu_k \rightarrow \mu$  if and only if one of the following conditions is satisfied:

$$\liminf_{k \rightarrow \infty} \mu_k(G) \geq \mu(G) \quad \text{for any open set } G \subset X, \quad (1.29)$$

$$\liminf_{k \rightarrow \infty} \mu_k(F) \leq \mu(F) \quad \text{for any closed set } F \subset X. \quad (1.30)$$

It is also equivalent to convergence (1.28) for any Borel subset  $\Gamma \subset X$  such that  $\mu(\partial\Gamma) = 0$ , where  $\partial\Gamma$  stands for the boundary of  $\Gamma$ . We refer to Theorem 11.1.1 in the book [Dud02] for a proof of these results.

A simple, but important consequence of the portmanteau theorem is the following description of the weak convergence of measures on the real line. Given a probability measure  $\mu \in \mathcal{P}(\mathbb{R})$ , we denote by  $F_\mu(x)$  its *distribution function*, defined by  $F_\mu(x) = \mu((-\infty, x])$  for  $x \in \mathbb{R}$ . Note that the distribution function of a measure is always non-decreasing and right-continuous at any point, and if the measure has no atoms, then its distribution function is continuous.

**Lemma 1.2.16.** *Let  $\{\mu_n\} \subset \mathcal{P}(\mathbb{R})$  be a sequence. Then  $\{\mu_n\}$  converges weakly to a measure  $\mu \in \mathcal{P}(\mathbb{R})$  if and only if*

$$F_{\mu_n}(x) \rightarrow F_\mu(x) \quad \text{as } n \rightarrow \infty,$$

where  $x \in \mathbb{R}$  is an arbitrary point of continuity for  $F_\mu$ .

Another remarkable property of weak convergence is that, under some assumptions, one can pass to the limit under the integrals, even if the integrand is not continuous. To formulate the corresponding result, recall that a continuous mapping  $\pi : X \rightarrow X$  acting in a Polish space  $X$  is called a *projection* if  $\pi \circ \pi = \pi$ .

**Lemma 1.2.17.** *Let  $X$  be a Polish space, let  $\pi_n : X \rightarrow X$ ,  $n \geq 1$ , be continuous projections, and let  $\{\mu_k\} \subset \mathcal{P}(X)$  be a sequence converging weakly to a measure  $\mu \in \mathcal{P}(X)$ . Assume that  $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is a Borel functional such that  $\{f \circ \pi_n\}$  is a sequence of bounded continuous functions converging to  $f$  pointwise and*

$$(f \circ \pi_n, \mu_k) \leq C \quad \text{for any } k, n \geq 1. \quad (1.31)$$

Then  $(f, \mu) \leq C$ , provided that either  $f \geq 0$  or  $\{f \circ \pi_n\}$  is non-decreasing.

Note that if  $\{f \circ \pi_n\}$  is non-decreasing, then inequality (1.31) will be satisfied if we assume that  $(f, \mu_k) \leq C$  for any  $k \geq 1$ .

*Proof.* We can pass to the limit in inequality (1.31) as  $k \rightarrow \infty$ . This results in  $(f \circ \pi_n, \mu) \leq C$ . Now the required assertion follows from Fatou's lemma in the first case and from the monotone convergence theorem in the second case.  $\square$

We complete this subsection by two exercises establishing some further properties of the space of probability measures and of the weak convergence.

*Exercise 1.2.18.* (i) Let  $X$  be a Polish space. Prove that the complete metric space  $(\mathcal{P}(X), \|\cdot\|_L^*)$  is separable.

(ii) Prove that the space  $(\mathcal{P}(\mathbb{R}), \|\cdot\|_{\text{var}})$  is not separable.

*Exercise 1.2.19.* Let  $X$  be a Polish space and let  $\zeta_m^n, \zeta_m, \zeta^n, \zeta$  be some  $X$ -valued random variables such that

$$\begin{aligned} \mathcal{D}(\zeta_m^n) &\rightarrow \mathcal{D}(\zeta^n) \quad \text{as } m \rightarrow \infty \text{ for any } n \geq 1, \\ \sup_{m \geq 1} \mathbb{E} (\text{dist}_X(\zeta_m, \zeta_m^n) + \text{dist}_X(\zeta^n, \zeta)) &\rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Show that  $\mathcal{D}(\zeta_m) \rightarrow \mathcal{D}(\zeta)$  as  $m \rightarrow \infty$ .

### 1.2.4 Couplings and maximal couplings of probability measures

**Definition 1.2.20.** Let  $\mu_1, \mu_2 \in \mathcal{P}(X)$ . A pair of random variables  $(\xi_1, \xi_2)$  defined on the same probability space is called a *coupling* for  $(\mu_1, \mu_2)$  if

$$\mathcal{D}(\xi_j) = \mu_j \quad \text{for } j = 1, 2. \quad (1.32)$$

The law  $\mathcal{D}(\xi_1, \xi_2) =: \mu$  is a measure on the product space  $X \times X$ . If we denote by  $\pi_1$  and  $\pi_2$  the projections of  $X \times X$  to the first and second factor, respectively, then

$$(\pi_1)_* \mu = \mu_1, \quad (\pi_2)_* \mu = \mu_2. \quad (1.33)$$

Other way round, if  $\mu$  is a measure on  $X \times X$  satisfying conditions (1.33), then the random variables  $\xi_1 = \pi_1$  and  $\xi_2 = \pi_2$  defined on the probability space  $(X \times X, \mathcal{B}(X \times X), \mu)$  meet (1.32). So a measure  $\mu$  on  $X \times X$  satisfying (1.33) is an alternative definition of the coupling. In this form, the coupling was systematically used by L. Kantorovich starting from late 1930's, e.g., see in [KA82].

In what follows, an important role is played by the *maximal coupling* of measures. Let  $(\xi_1, \xi_2)$  be a coupling for  $(\mu_1, \mu_2)$ . For any  $\Gamma \in \mathcal{B}(X)$ , we have

$$\begin{aligned} \mu_1(\Gamma) - \mu_2(\Gamma) &= \mathbb{E}(\mathbb{I}_\Gamma(\xi_1) - \mathbb{I}_\Gamma(\xi_2)) \\ &= \mathbb{E}(\mathbb{I}_{\{\xi_1 \neq \xi_2\}}(\mathbb{I}_\Gamma(\xi_1) - \mathbb{I}_\Gamma(\xi_2))) \leq \mathbb{P}\{\xi_1 \neq \xi_2\}. \end{aligned}$$

Therefore,

$$\mathbb{P}\{\xi_1 \neq \xi_2\} \geq \|\mu_1 - \mu_2\|_{\text{var}}.$$

**Definition 1.2.21.** A coupling  $(\xi_1, \xi_2)$  is said to be *maximal* if

$$\mathbb{P}\{\xi_1 \neq \xi_2\} = \|\mu_1 - \mu_2\|_{\text{var}},$$

and the random variables  $\xi_1$  and  $\xi_2$  conditioned on the event  $N = \{\xi_1 \neq \xi_2\}$  are independent. The latter condition means that, for any  $\Gamma_1, \Gamma_2 \in \mathcal{B}(X)$ , we have<sup>2</sup>

$$\mathbb{P}\{\xi_1 \in \Gamma_1, \xi_2 \in \Gamma_2 \mid N\} = \mathbb{P}\{\xi_1 \in \Gamma_1 \mid N\} \mathbb{P}\{\xi_2 \in \Gamma_2 \mid N\}.$$

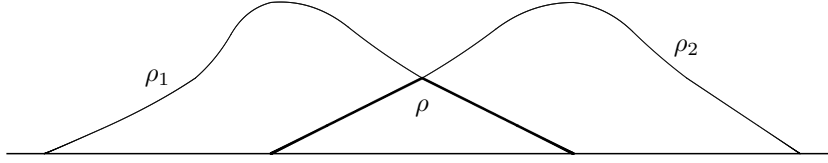
*Exercise 1.2.22.* Let  $\mu_1, \mu_2 \in \mathcal{P}(X)$  be such that  $\|\mu_1 - \mu_2\|_{\text{var}} = 1$ . Show that a coupling  $(\xi_1, \xi_2)$  for  $(\mu_1, \mu_2)$  is maximal if and only if  $\xi_1$  and  $\xi_2$  are independent.

*Exercise 1.2.23.* Let  $(\xi_1, \xi_2)$  be any pair of random variables that are independent on the event  $N = \{\xi_1 \neq \xi_2\}$ . Show that

$$\mathbb{P}\{\xi_1 \in \Gamma, \xi_2 \in \Gamma\} \geq \mathbb{P}\{\xi_1 \in \Gamma\} \mathbb{P}\{\xi_2 \in \Gamma\} \quad \text{for any } \Gamma \in \mathcal{B}(X). \quad (1.34)$$

The following result is often referred to as the *coupling lemma* or *Dobrushin's lemma*. It makes an effective tool to study the total variation distance between measures.

<sup>2</sup>In the case  $\mathbb{P}(N) = 0$ , this condition should be omitted.

Figure 1.1: Densities  $\rho_1$ ,  $\rho_2$ , and  $\rho$  marked by thin and thick graphs

**Lemma 1.2.24.** *For any two measures  $\mu_1, \mu_2 \in \mathcal{P}(X)$  there exists a maximal coupling  $(\xi_1, \xi_2)$ .*

*Proof.* Let us set  $\delta := \|\mu_1 - \mu_2\|_{\text{var}}$ . If  $\delta = 1$ , then, by Exercise 1.2.22, any pair  $(\xi_1, \xi_2)$  of independent random variables with  $\mathcal{D}(\xi_i) = \mu_i$ ,  $i = 1, 2$ , is a maximal coupling for  $(\mu_1, \mu_2)$ . If  $\delta = 0$ , then  $\mu_1 = \mu_2$ , and for any random variable  $\xi$  with distribution  $\mu_1$  the pair  $(\xi, \xi)$  is a maximal coupling. Hence, we can assume that  $0 < \delta < 1$ .

Let  $m = \frac{1}{2}(\mu_1 + \mu_2)$  and let (see Figure 1.1)

$$\rho_i = \frac{d\mu_i}{dm}, \quad \rho = \rho_1 \wedge \rho_2, \quad \hat{\rho}_i = \delta^{-1}(\rho_i - \rho). \quad (1.35)$$

Direct verification shows that the measures  $\hat{\mu}_i = \hat{\rho}_i dm$  and  $\mu = (1 - \delta)^{-1} \rho dm$  are probabilities on  $X$ . Let  $\zeta_1, \zeta_2, \zeta$ , and  $\alpha$  be independent random variables defined on the same probability space such that

$$\mathcal{D}(\zeta_i) = \hat{\mu}_i, \quad \mathcal{D}(\zeta) = \mu, \quad \mathbb{P}\{\alpha = 0\} = \delta, \quad \mathbb{P}\{\alpha = 1\} = 1 - \delta. \quad (1.36)$$

We claim that the random variables  $\xi_i = \alpha\zeta + (1 - \alpha)\zeta_i$ ,  $i = 1, 2$ , form a maximal coupling for  $(\mu_1, \mu_2)$ . Indeed, for any  $\Gamma \in \mathcal{B}(X)$ , we have

$$\begin{aligned} \mathbb{P}\{\xi_i \in \Gamma\} &= \mathbb{P}\{\xi_i \in \Gamma, \alpha = 0\} + \mathbb{P}\{\xi_i \in \Gamma, \alpha = 1\} \\ &= \mathbb{P}\{\alpha = 0\}\mathbb{P}\{\zeta_i \in \Gamma\} + \mathbb{P}\{\alpha = 1\}\mathbb{P}\{\zeta \in \Gamma\} \\ &= \delta \int_{\Gamma} \hat{\rho}_i(u) dm + \int_{\Gamma} \rho(u) dm = \mu_i(\Gamma), \end{aligned} \quad (1.37)$$

where we used the independence of  $(\zeta_1, \zeta_2, \zeta, \alpha)$  and the relation  $\rho_i = \rho + \delta\hat{\rho}_i$ . Furthermore,

$$\begin{aligned} \mathbb{P}\{\xi_1 \neq \xi_2\} &= \mathbb{P}\{\xi_1 \neq \xi_2, \alpha = 0\} + \mathbb{P}\{\xi_1 \neq \xi_2, \alpha = 1\} \\ &= \mathbb{P}\{\alpha = 0\}\mathbb{P}\{\zeta_1 \neq \zeta_2\} = \delta, \end{aligned}$$

where we used again the independence of  $(\zeta_1, \zeta_2, \zeta, \alpha)$  and also the relation

$$\mathbb{P}\{\zeta_1 = \zeta_2\} = \delta^{-2} \iint_{\{u_1=u_2\}} \hat{\rho}_1(u_1)\hat{\rho}_2(u_2) m(du_1)m(du_2) = 0,$$

which follows from the identity  $\hat{\rho}_1(u)\hat{\rho}_2(u) \equiv 0$ . A similar argument shows that the random variables  $\xi_1$  and  $\xi_2$  conditioned on  $\{\xi_1 \neq \xi_2\}$  are independent. This completes the proof of Lemma 1.2.24.  $\square$

Relation (1.37), Corollary 1.2.11, and Exercise 1.2.22 imply the following alternative version<sup>3</sup> of the coupling lemma.

**Corollary 1.2.25.** *Any two measures  $\mu_1, \mu_2 \in \mathcal{P}(X)$  admit a representation*

$$\mu_j = (1 - \delta)\mu + \delta\nu_j, \quad j = 1, 2,$$

where  $\delta = \|\mu_1 - \mu_2\|_{\text{var}}$ ,  $\mu, \nu_1, \nu_2 \in \overline{\mathcal{P}(X)}$ , and the measures  $\nu_1$  and  $\nu_2$  are mutually singular.

We shall call  $(1 - \delta)\mu$  the *minimum* of  $\mu_1$  and  $\mu_2$  and denote it by  $\mu_1 \wedge \mu_2$ . Another important corollary is the following result on the conditional law of the random variables that form a maximal coupling. Recall that the *conditional law*  $\mathcal{D}(\xi | N)$  of an  $X$ -valued random variable  $\xi$  given an event  $N$  of non-zero probability is defined by the relation

$$\mathcal{D}(\xi | N)(\Gamma) = \frac{\mathbb{P}(\{\xi \in \Gamma\} \cap N)}{\mathbb{P}(N)}, \quad \Gamma \in \mathcal{B}(X).$$

**Lemma 1.2.26.** *Let  $(\xi_1, \xi_2)$  be a maximal coupling for a pair of measures  $\mu_1, \mu_2 \in \mathcal{P}(X)$  such that  $\mu_1 \wedge \mu_2(X) > 0$ . Then  $\mathbb{P}\{\xi_1 = \xi_2\} > 0$ , and we have*

$$\mathcal{D}(\xi_1 | \{\xi_1 = \xi_2\}) = \mathcal{D}(\xi_2 | \{\xi_1 = \xi_2\}) = \frac{\mu_1 \wedge \mu_2}{\mu_1 \wedge \mu_2(X)}. \quad (1.38)$$

*Proof.* The case in which  $\xi_1 = \xi_2$  almost surely is trivial, and we assume that  $\xi_1 \neq \xi_2$  with positive probability. Proposition 1.2.7 implies that

$$\mathbb{P}\{\xi_1 \neq \xi_2\} = \|\mu_1 - \mu_2\|_{\text{var}} = 1 - \mu_1 \wedge \mu_2(X) < 1, \quad (1.39)$$

whence we conclude that  $\mathbb{P}\{\xi_1 = \xi_2\} > 0$ , and the conditional laws in (1.38) are well defined. Let us set  $\mu = \mathcal{D}(\xi_1 | \{\xi_1 = \xi_2\})$  and  $\tilde{\mu}_i = \mathcal{D}(\xi_i | \{\xi_1 \neq \xi_2\})$ . Then

$$\mu_i = \mathbb{P}\{\xi_1 = \xi_2\}\mu + \mathbb{P}\{\xi_1 \neq \xi_2\}\tilde{\mu}_i, \quad i = 1, 2.$$

It follows that

$$\mu_1 \wedge \mu_2 = \mathbb{P}\{\xi_1 = \xi_2\}\mu + \mathbb{P}\{\xi_1 \neq \xi_2\}\tilde{\mu}_1 \wedge \tilde{\mu}_2.$$

Combining this with (1.39), we see that  $\tilde{\mu}_1 \wedge \tilde{\mu}_2 = 0$ , and the above relation gives immediately (1.38).  $\square$

In what follows, we deal with pairs of measures depending on a parameter, and we shall need a maximal coupling for them that depends on the parameter in a measurable manner. More precisely, let  $Z$  be a Polish space endowed with its Borel  $\sigma$ -algebra and let  $\{\mu(z, du), z \in Z\}$  be a family of probability measures on  $X$ . We shall say that  $\mu(z, du)$  is a *random probability measure on  $X$*  if for any Borel set  $\Gamma \subset X$  the function  $z \mapsto \mu(z, \Gamma)$  is measurable from  $Z$  to  $\mathbb{R}$ .

<sup>3</sup>In this form, the coupling lemma was intensively used by Dobrushin; e.g., see [Dob74].

*Exercise 1.2.27.* Prove that  $\{\mu(z, du), z \in Z\}$  is a random probability measure if and only if the mapping  $z \mapsto \mu(z, \cdot)$  is measurable from  $Z$  to the space  $\mathcal{P}(X)$  endowed with the Borel  $\sigma$ -algebra corresponding to the dual-Lipschitz metric.

**Theorem 1.2.28.** *Let  $X$  and  $Z$  be Polish spaces, and let  $\{\mu_i(z, du), z \in Z\}$ ,  $i = 1, 2$ , be two random probability measures on  $X$ . Then there is a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and two measurable functions  $\xi_i(z, \omega) : Z \times \Omega \rightarrow X$ ,  $i = 1, 2$ , such that  $(\xi_1(z, \cdot), \xi_2(z, \cdot))$  is a maximal coupling for  $(\mu_1(z, du), \mu_2(z, du))$  for any  $z \in Z$ .*

*Proof.* In view of Theorem 7.2.2 in the Appendix, any Polish space is a standard measurable space. Since the objects considered in the theorem are invariant with respect to measurable isomorphisms, we can assume from the very beginning that  $X$  coincides with one of the spaces described in Definition 7.2.1. To be precise, we shall assume that  $X$  is the interval  $[0, 1]$  endowed with its Borel  $\sigma$ -algebra.

We shall repeat the scheme used in the proof of Lemma 1.2.24, controlling the dependence of the resulting random variables on the parameter  $z$ . To this end, we first note that if  $\mu(z, du)$  is a random probability measure on  $X$ , then there is a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a measurable function  $\xi(z, \omega) : Z \times \Omega \rightarrow X$  such that the law of  $\xi(z, \cdot)$  coincides with  $\mu(z, du)$  for any  $z \in Z$ . For instance, we can take for the probability space the interval  $[0, 1]$ , endowed with its Borel  $\sigma$ -algebra and the Lebesgue measure, and define  $\xi$  by the relation

$$\xi(z, \omega) = \min\{t \in [0, 1] : F(z, t) \geq \omega\},$$

where  $F(z, t) = \mu(z, (-\infty, t])$  is the probability distribution function for the measure  $\mu$ .

We now introduce the family

$$\nu(z, du) = \mu_1(z, du) \wedge \mu_2(z, du),$$

where  $\nu_1 \wedge \nu_2$  denotes the minimum of two measures  $\nu_1, \nu_2 \in \mathcal{P}(X)$ , and define the following families of probability measures (cf. (1.35)):

$$\hat{\mu}_i(z, du) = \begin{cases} \delta(z)^{-1}(\mu_i(z, du) - \nu(z, du)) & \text{for } \delta(z) \neq 0, \\ \mu_i(z, du) & \text{otherwise,} \end{cases}$$

$$\mu(z, du) = \begin{cases} (1 - \delta(z))^{-1} \nu(z, du) & \text{for } \delta(z) \neq 1, \\ \lambda & \text{otherwise,} \end{cases}$$

where  $\delta(z) = \|\mu_1(z, \cdot) - \mu_2(z, \cdot)\|_{\text{var}}$ , and  $\lambda \in \mathcal{P}(X)$  is any fixed measure. It can be shown that  $\delta$  is a measurable function of  $z \in Z$ , and  $\hat{\mu}_i(z, \cdot)$  and  $\mu(z, \cdot)$  are random probability measures on  $X$ ; see Exercise 1.2.29 below.

What has been said implies that there is a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and measurable functions  $\zeta_1, \zeta_2, \zeta$ , and  $\alpha$  of the variable  $(z, \omega) \in Z \times \Omega$  such that, for any  $z \in Z$ , the random variables  $\zeta_i(z, \cdot)$ ,  $\zeta(z, \cdot)$ , and  $\alpha(z, \cdot)$  are independent, and

$$\mathcal{D}(\zeta_i(z, \cdot)) = \hat{\mu}_i(z, du), \quad \mathcal{D}(\zeta(z, \cdot)) = \mu(z, du),$$

$$\mathbb{P}\{\alpha(z, \cdot) = 0\} = \delta(z), \quad \mathbb{P}\{\alpha(z, \cdot) = 1\} = 1 - \delta(z);$$

cf. (1.36). The proof of the theorem can now be completed by a literal repetition of the argument used to establish Lemma 1.2.24.  $\square$

*Exercise 1.2.29.* Show that  $\delta(z)$  is a measurable function of  $z \in Z$ , and  $\hat{\mu}_i(z, \cdot)$  and  $\mu(z, \cdot)$  are random probability measures on  $X$ . *Hint:* Use a parameter version of the Radon–Nikodym theorem (e.g., see [Nov05]): if  $\lambda(z, du)$  and  $\nu(z, du)$  are random probability measures on  $X$  such that  $\lambda(z, du)$  is absolutely continuous with respect to  $\nu(z, du)$  for any  $z \in Z$ , then the Radon–Nikodym derivative  $\frac{d\lambda(z, \cdot)}{d\nu(z, \cdot)}$  can be chosen to be a measurable function of  $(z, u) \in Z \times X$ .

The existence result of the following exercise is useful when one constructs a coupling for solutions of stochastic PDE's; see [Mat02b, Oda08] and Section 3.5.1.

*Exercise 1.2.30.* Let  $X$  and  $Y$  be Polish spaces.

(i) Show that for any pair of measures  $\mu_1, \mu_2 \in \mathcal{P}(X)$  and any measurable mapping  $f : X \rightarrow Y$  there is a coupling  $(\xi_1, \xi_2)$  for  $(\mu_1, \mu_2)$  such that  $(f(\xi_1), f(\xi_2))$  is a maximal coupling for  $(f_*(\mu_1), f_*(\mu_2))$ .

(ii) Let  $Z$  be a Polish space and let  $f : X \times Z \rightarrow Y$  be a measurable mapping. Show that for any random probability measures  $\mu_1(z, du)$  and  $\mu_2(z, du)$  on  $X$  there is a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and  $X$ -valued measurable functions  $\xi_i(z, \omega)$ ,  $i = 1, 2$ , defined on  $Z \times \Omega$  such that, for any  $z \in Z$ , the pair  $(f(z, \xi_1), f(z, \xi_2))$  is a coupling for  $(f_*(z, \mu_1(z, \cdot)), f_*(z, \mu_2(z, \cdot)))$ , where  $f_*(z, \mu_i(z, \cdot))$  stands for the image of  $\mu_i(z, \cdot)$  under the mapping  $u \mapsto f(z, u)$ .

## 1.2.5 Kantorovich functionals

Let  $F$  be a measurable symmetric function on  $X \times X$  such that

$$F(u_1, u_2) \geq \text{dist}_X(u_1, u_2) \quad \text{for all } u_1, u_2 \in X. \quad (1.40)$$

We define the *Kantorovich functional* corresponding to  $F$  as the following function  $\mathcal{K}_F$  on  $\mathcal{P}(X) \times \mathcal{P}(X)$ :

$$\mathcal{K}_F(\mu_1, \mu_2) = \inf\{\mathbb{E} F(\xi_1, \xi_2)\}, \quad (1.41)$$

where the infimum is taken over all couplings  $(\xi_1, \xi_2)$  for  $(\mu_1, \mu_2)$ . The function  $F$  is called the (*Kantorovich density*) of the functional  $\mathcal{K}_F$ .

**Lemma 1.2.31.** *For any  $\mu_1, \mu_2 \in \mathcal{P}(X)$ , we have<sup>4</sup>*

$$\|\mu_1 - \mu_2\|_L^* \leq \mathcal{K}_F(\mu_1, \mu_2). \quad (1.42)$$

*Proof.* Let  $(\xi_1, \xi_2)$  be a coupling for  $(\mu_1, \mu_2)$ . Then, for any  $g \in L_b(X)$  with  $\|g\|_L \leq 1$ , we have

$$(g, \mu_1 - \mu_2) = \mathbb{E}(g(\xi_1) - g(\xi_2)) \leq \mathbb{E} \text{dist}_X(\xi_1, \xi_2) \leq \mathbb{E} F(\xi_1, \xi_2).$$

Taking first the supremum in  $g$  and then the infimum with respect to all couplings  $(\xi_1, \xi_2)$ , we obtain (1.42).  $\square$

<sup>4</sup>The right-hand side of inequality (1.42) may be infinite.



In the case when the function  $F$  coincides with  $\text{dist}_X$ , we have the following result, which is the celebrated *Kantorovich–Rubinstein theorem*.

**Theorem 1.2.32.** *For any probability measures  $\mu_1, \mu_2 \in \mathcal{P}(X)$ , we have*

$$\|\mu_1 - \mu_2\|_K = \inf\{\mathbb{E} \text{dist}_X(\xi_1, \xi_2)\} = \mathcal{K}_{\text{dist}_X}(\mu_1, \mu_2), \quad (1.43)$$

where the infimum is taken over all couplings  $(\xi_1, \xi_2)$  for  $(\mu_1, \mu_2)$ . Moreover, the infimum is attained at some coupling  $(\xi_1, \xi_2)$ .

In the case when the space  $X$  is endowed with the discrete topology (that is,  $\text{dist}_X(u_1, u_2) = 1$  for any  $u_1 \neq u_2$ ), Theorem 1.2.32 is a straightforward consequence of the existence of a maximal coupling of measures (see Lemma 1.2.24). For the general case, we refer the reader to [KA82, Section VIII.4] or [Dud02, Theorem 11.8.2].

In his celebrated research on the mass-transfer problem (for which he was given a Nobel prize in economics), L. Kantorovich interpreted  $\mathbb{E} \text{dist}_X(\xi_1, \xi_2)$  as the work needed to transport mass points  $\xi_1(\omega)$  to  $\xi_2(\omega)$ , and used relation (1.43) to estimate the work via the distance between the measures  $\mu_1$  and  $\mu_2$ . We shall use that equality other way round, that is, as a tool to estimate this distance. Accordingly, inequality (1.42) will be sufficient for our purposes.

## 1.3 Markov processes and random dynamical systems

### 1.3.1 Markov processes

Let  $X$  be a Polish space. A *Markov family of random processes in  $X$*  (or simply a *Markov process*) is defined as a collection of the following objects (e.g., see [KS91, Section 2.5]):

- a measurable space  $(\Omega, \mathcal{F})$  with a filtration  $\{\mathcal{F}_t, t \in \mathcal{T}_+\}$ ;
- a family of probability measures  $\{\mathbb{P}_v, v \in X\}$  on  $(\Omega, \mathcal{F})$  such that the mapping  $v \mapsto \mathbb{P}_v(A)$  is universally measurable<sup>5</sup> for any  $A \in \mathcal{F}$ ;
- an  $X$ -valued random process  $\{u_t, t \in \mathcal{T}_+\}$  adapted to the filtration  $\mathcal{F}_t$  and satisfying the conditions below for any  $v \in X$ ,  $\Gamma \in \mathcal{B}(X)$ , and  $t, s \in \mathcal{T}_+$ :

$$\mathbb{P}_v\{u_0 = v\} = 1, \quad (1.44)$$

$$\mathbb{P}_v\{u_{t+s} \in \Gamma \mid \mathcal{F}_s\} = P_t(u_s, \Gamma) \quad \text{for } \mathbb{P}_v\text{-almost every } \omega \in \Omega. \quad (1.45)$$

Here  $P_t$  stands for the *transition function* of the family defined as the law of  $u_t$  under the probability measure  $\mathbb{P}_v$ :

$$P_t(v, \Gamma) := \mathbb{P}_v\{u_t \in \Gamma\}, \quad v \in X, \quad \Gamma \in \mathcal{B}(X). \quad (1.46)$$

<sup>5</sup>See Section 7.3 for the definition of universal measurability.

Relation (1.45) is called *Markov property*, and the above Markov family is denoted by  $(u_t, \mathbb{P}_v)$ . If  $\mathcal{T}_+ = \mathbb{Z}_+$ , then the object we have just defined is also called a *family of Markov chains* (or a *discrete-time Markov process* or simply a *Markov chain*).

Given  $\lambda \in \mathcal{P}(X)$  and a Markov process  $(u_t, \mathbb{P}_v)$ , we define the probability measure

$$\mathbb{P}_\lambda(\Gamma) = \int_X \mathbb{P}_v(\Gamma) \lambda(dv), \quad \Gamma \in \mathcal{F}, \quad (1.47)$$

and denote by  $\mathbb{E}_\lambda$  the corresponding mean value. The following exercise gives a simple generalisation of the Markov property.

*Exercise 1.3.1.* Show that if  $(u_t, \mathbb{P}_v)$  is a Markov process, then for  $f \in L^\infty(X)$  and  $\lambda \in \mathcal{P}(X)$ , we have

$$\mathbb{E}_\lambda\{f(u_{t+s}) \mid \mathcal{F}_s\} = \mathbb{E}_{u_s} f(u_t) \quad \mathbb{P}_\lambda\text{-almost surely.} \quad (1.48)$$

More generally, for any  $m \geq 1$ , any  $0 < t_1 < \dots < t_m$ , and any bounded measurable function  $f : X \times \dots \times X \rightarrow \mathbb{R}$ , we have

$$\mathbb{E}_\lambda\{f(u_{t_1+s}, \dots, u_{t_m+s}) \mid \mathcal{F}_s\} = \mathbb{E}_{u_s} f(u_{t_1}, \dots, u_{t_m}) \quad \mathbb{P}_\lambda\text{-almost surely.} \quad (1.49)$$

We now establish the so-called *Kolmogorov–Chapman relation*, which will imply, in particular, that every Markov process generates an evolution in the space of probability measures.

**Lemma 1.3.2.** *For any  $t, s \in \mathcal{T}_+$ ,  $v \in X$ , and  $\Gamma \in \mathcal{B}(X)$ , the transition function  $P_t$  satisfies the relation*

$$P_{t+s}(v, \Gamma) = \int_X P_t(v, dz) P_s(z, \Gamma). \quad (1.50)$$

*Proof.* In view of (1.45), we have

$$\begin{aligned} P_{t+s}(v, \Gamma) &= \mathbb{E}_v \mathbb{I}_\Gamma(u_{t+s}) = \mathbb{E}_v \{ \mathbb{E}_v (\mathbb{I}_\Gamma(u_{t+s}) \mid \mathcal{F}_t) \} \\ &= \mathbb{E}_v P_s(u_t, \Gamma) = \int_X P_t(v, dz) P_s(z, \Gamma), \end{aligned}$$

where we used the fact that the law of  $u_t$  under  $\mathbb{P}_v$  coincides with  $P_t(v, \cdot)$ .  $\square$

To each Markov process there correspond two families of linear operators acting in the spaces of bounded measurable functions  $L^\infty(X)$  and of probability measures  $\mathcal{P}(X)$ . They are called *Markov semigroups* and are defined in terms of the transition function by the following relations:

$$\begin{aligned} \mathfrak{P}_t : L^\infty(X) &\rightarrow L^\infty(X), & \mathfrak{P}_t f(v) &= \int_X P_t(v, dz) f(z), \\ \mathfrak{P}_t^* : \mathcal{P}(X) &\rightarrow \mathcal{P}(X), & \mathfrak{P}_t^* \mu(\Gamma) &= \int_X P_t(v, \Gamma) \mu(dv), \end{aligned}$$

where  $t \in \mathcal{T}_+$ . It is straightforward to check that  $\mathfrak{P}_t f \in L^\infty(X)$  for  $f \in L^\infty(X)$  and  $\mathfrak{P}_t^* \mu \in \mathcal{P}(X)$  for  $\mu \in \mathcal{P}(X)$ . The following exercise justifies the term *semigroup* and establishes some simple properties.

*Exercise 1.3.3.* Show that the families  $\mathfrak{P}_t$  and  $\mathfrak{P}_t^*$  form semigroups, that is,  $\mathfrak{P}_0 = \text{Id}$  and  $\mathfrak{P}_{t+s} = \mathfrak{P}_s \circ \mathfrak{P}_t$ , and similarly for  $\mathfrak{P}_t^*$ . Show also that they satisfy the following duality relation:

$$(\mathfrak{P}_t f, \mu) = (f, \mathfrak{P}_t^* \mu) \quad \text{for any } f \in L^\infty(X), \mu \in \mathcal{P}(X).$$

*Hint:* Use the Kolmogorov–Chapman relation (1.50) for the semigroup property and Fubini’s theorem for the duality relation.

*Exercise 1.3.4.* Let  $(u_t, \mathbb{P}_v)$  be a Markov process in  $X$ , let  $\lambda \in \mathcal{P}(X)$ , and let  $\mathbb{P}_\lambda$  be the measure defined by (1.47). Prove that the  $\mathbb{P}_\lambda$ -law of  $u_t$  coincides with  $\mathfrak{P}_t^* \lambda$ . Use this property to show that if  $f \in L^\infty(X)$ , then

$$\mathbb{E}_\lambda f(u_t) = \int_X \mathfrak{P}_t f(v) \lambda(dv). \quad (1.51)$$

**Definition 1.3.5.** A measure  $\mu \in \mathcal{P}(X)$  is said to be *stationary* for  $(u_t, \mathbb{P}_v)$  if  $\mathfrak{P}_t^* \mu = \mu$  for all  $t \geq 0$ .

*Example 1.3.6.* Let  $H$  be a separable Banach space, let  $S : H \rightarrow H$  be a continuous mapping, and let  $\{\eta_k, k \geq 1\}$  be a sequence of i.i.d. random variables in  $H$  defined on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . We fix  $v \in H$  and consider a sequence  $\{v_k, k \geq 0\}$  defined by the rule

$$v_0 = v, \quad v_k = S(v_{k-1}) + \eta_k, \quad k \geq 1. \quad (1.52)$$

This system defines a discrete-time Markov process in  $H$ . Indeed, let us denote by  $(\tilde{\Omega}, \tilde{\mathcal{F}})$  the product of the measurable spaces  $(H, \mathcal{B}(H))$  and  $(\Omega, \mathcal{F})$  (that is,  $\tilde{\Omega} = H \times \Omega$  and  $\tilde{\mathcal{F}} = \mathcal{B}(H) \otimes \mathcal{F}$ ) and provide it with the filtration  $\{\tilde{\mathcal{F}}_k, k \geq 0\}$ , where  $\tilde{\mathcal{F}}_k = \mathcal{B}(H) \otimes \mathcal{F}_k$  and  $\mathcal{F}_k$  is the  $\sigma$ -algebra generated by  $\eta_1, \dots, \eta_k$  (so that  $\mathcal{F}_0$  is the trivial  $\sigma$ -algebra). For any  $v \in H$ , we take  $\mathbb{P}_v = \delta_v \otimes \mathbb{P}$  and define a process  $\{u_k, k \geq 0\}$  as follows: for  $\tilde{\omega} = (v, \omega)$  we set  $u_k^{\tilde{\omega}} = v_k^\omega$ , where  $\{v_k\}$  is given by (1.52). Denote by  $P_k(v, \cdot)$  the law of  $v_k$ , where  $k \geq 0$  and  $v \in H$ . It is easy to see that the objects introduced above meet (1.44) and (1.45).

Now let  $v$  be an  $H$ -valued random variable independent of  $\{\eta_k\}$  with a law  $\mu$ . Then, by Exercise 1.3.4, we have  $\mathcal{D}_\mu(u_k) = \mathfrak{P}_k^* \mu$  for  $k \geq 0$ . In particular, if  $\mu$  is a stationary measure for the Markov process  $(u_k, \mathbb{P}_v)$ , then  $\mathcal{D}_\mu(u_k) = \mu$  for all  $k \geq 0$ . In what follows, a stationary measure for  $(u_k, \mathbb{P}_v)$  is called a *stationary measure* for Eq. (1.52) and the sequence  $\{v_k\}$  defined by (1.52) is called a *stationary solution*.

An important feature of Markov processes is the *strong Markov property*. It says, roughly speaking, that relation (1.45) remains valid if  $s$  is replaced by a random time  $\sigma$ , provided that it satisfies an addition property of the “independence of the future”. To formulate the corresponding result, we first introduce the concept of a *stopping time*, formalising that property.

**Definition 1.3.7.** A random variable  $\tau : \Omega \rightarrow \mathcal{T}_+ \cup \{\infty\}$  is called a *stopping time* for a filtration  $\{\mathcal{F}_t, t \in \mathcal{T}_+\}$  if the event  $\{\tau \leq t\}$  belongs to  $\mathcal{F}_t$  for any  $t \in \mathcal{T}_+$ .

*Exercise 1.3.8.* (i) Show that if  $\tau$  and  $\sigma$  are stopping times, then the random variables  $\tau + \sigma$ ,  $\tau \wedge \sigma$ , and  $\tau \vee \sigma$  are also stopping times.

(ii) Suppose that the function  $t \mapsto u_t$  is continuous from  $\mathbb{R}_+$  to  $X$ . Show that for any closed subset  $A \subset X$ , the random variable<sup>6</sup>

$$\tau(A) = \min\{t \in \mathcal{T}_+ : u_t \in A\}$$

is a stopping time. *Hint:* See Exercise 2.7 in [KS91].

For any stopping time  $\tau$ , we shall denote by  $\mathcal{F}_\tau$  the  $\sigma$ -algebra of the events  $\Gamma \in \mathcal{F}$  such that

$$\Gamma \cap \{\tau \leq t\} \in \mathcal{F}_t \quad \text{for any } t \in \mathcal{T}_+.$$

*Exercise 1.3.9.* Show that any stopping time  $\tau$  is  $\mathcal{F}_\tau$ -measurable.

In what follows, we shall always assume the Markov processes we deal with satisfy the following additional hypotheses:

*Feller property.* For any  $f \in C_b(X)$  and  $t \geq 0$ , we have  $\mathfrak{P}_t f \in C_b(X)$ .

*Time continuity.* The trajectories  $u_t(\omega)$ ,  $\omega \in \Omega$ , are continuous in time.<sup>7</sup>

**Theorem 1.3.10.** *Let  $(u_t, \mathbb{P}_u)$  be a Markov process, let  $f : X \rightarrow \mathbb{R}$  be a bounded measurable function, and let  $\tau$  be a stopping time. Then, for any almost surely finite  $\mathcal{F}_\tau$ -measurable random variable  $\sigma : \Omega \rightarrow \mathcal{T}_+$  and any measure  $\lambda$  on  $(X, \mathcal{B}(X))$ , we have*

$$\mathbb{E}_\lambda(\mathbb{I}_{\{\tau < \infty\}} f(u_{\tau+\sigma}) | \mathcal{F}_\tau) = \mathbb{I}_{\{\tau < \infty\}} (\mathfrak{P}_\sigma f)(u_\tau) \quad \mathbb{P}_\lambda\text{-almost surely.} \quad (1.53)$$

Relation (1.53) is called the *strong Markov property* of the Markov family  $(u_t, \mathbb{P}_u)$ .

*Proof of Theorem 1.3.10.* We confine ourselves to the case in which  $\mathcal{T}_+ = \mathbb{Z}_+$ . The proof in the case of continuous-time Markov processes can be carried out with the help of approximation of  $\tau$  by stopping times with range in an increasing sequence of discrete subsets of  $\mathbb{R}_+$ ; e.g., see Section III.3 in [RY99].

*Step 1.* It suffices to show that

$$\mathbb{E}_\lambda(\mathbb{I}_{\{\tau < \infty\}} f(u_{\tau+m}) | \mathcal{F}_\tau) = \mathbb{I}_{\{\tau < \infty\}} (\mathfrak{P}_m f)(u_\tau), \quad (1.54)$$

where  $m \geq 0$  is an arbitrary integer, and the equality holds  $\mathbb{P}_\lambda$ -almost surely. Indeed, if (1.54) is established, then we can write

$$\begin{aligned} \mathbb{E}_\lambda\{\mathbb{I}_{\{\tau < \infty\}} f(u_{\tau+\sigma}) | \mathcal{F}_\tau\} &= \sum_{m=0}^{\infty} \mathbb{E}_\lambda\{\mathbb{I}_{\{\tau < \infty\}} \mathbb{I}_{\{\sigma=m\}} f(u_{\tau+m}) | \mathcal{F}_\tau\} \\ &= \mathbb{I}_{\{\tau < \infty\}} \sum_{m=0}^{\infty} \mathbb{I}_{\{\sigma=m\}} (\mathfrak{P}_m f)(u_\tau) \\ &= \mathbb{I}_{\{\tau < \infty\}} (\mathfrak{P}_\sigma f)(u_\tau). \end{aligned}$$

<sup>6</sup>As usual, the minimum over an empty set is equal to  $+\infty$ .

<sup>7</sup>This condition is imposed only if  $\mathcal{T}_+ = \mathbb{R}_+$ .

*Step 2.* We now prove (1.54). To this end, it suffices to show that

$$\mathbb{E}_\lambda(\mathbb{I}_{\Gamma \cap \{\tau < \infty\}} f(u_{\tau+m})) = \mathbb{E}_\lambda(\mathbb{I}_{\Gamma \cap \{\tau < \infty\}} (\mathfrak{P}_m f)(u_\tau)), \quad (1.55)$$

where  $\Gamma \in \mathcal{F}_\tau$  is an arbitrary subset. Let us note that

$$\mathbb{I}_{\Gamma \cap \{\tau < \infty\}} = \sum_{n=0}^{\infty} \mathbb{I}_{\Gamma \cap \{\tau = n\}}. \quad (1.56)$$

Since  $\Gamma_n := \Gamma \cap \{\tau = n\}$  belongs to  $\mathcal{F}_n$ , the Markov property (1.48) implies that

$$\begin{aligned} \mathbb{E}_\lambda(\mathbb{I}_{\Gamma_n} f(u_{\tau+m})) &= \mathbb{E}_\lambda(\mathbb{I}_{\Gamma_n} f(u_{n+m})) \\ &= \mathbb{E}_\lambda(\mathbb{I}_{\Gamma_n} \mathbb{E}\{f(u_{n+m}) \mid \mathcal{F}_n\}) \\ &= \mathbb{E}_\lambda(\mathbb{I}_{\Gamma_n} (\mathfrak{P}_m f)(u_n)). \end{aligned}$$

Combining this relation with (1.56), we obtain (1.55).  $\square$

**Corollary 1.3.11.** *Let  $\tau$  be a stopping time such that  $\tau \leq t$  almost surely. Then, for any  $\Gamma \in \mathcal{B}(X)$  and  $u \in X$ , we have*

$$P_t(u, \Gamma) = \mathbb{E}_u P_{t-\tau}(u_\tau, \Gamma). \quad (1.57)$$

More generally, if  $f: X \rightarrow \mathbb{R}$  be a bounded measurable function, then

$$\mathfrak{P}_t f(u) = \mathbb{E}_u (\mathfrak{P}_{t-\tau} f)(u_\tau) \quad \text{for any } u \in X. \quad (1.58)$$

*Proof.* To prove (1.57), let us set  $\sigma = t - \tau$  and apply (1.53) to the function  $f = \mathbb{I}_\Gamma$ . This results in the relation

$$\begin{aligned} P_t(u, \Gamma) &= \mathbb{E}_u \mathbb{I}_\Gamma(u_t) = \mathbb{E}_u \mathbb{E}_u(\mathbb{I}_\Gamma(u_{\tau+\sigma}) \mid \mathcal{F}_\tau) \\ &= \mathbb{E}_u (\mathfrak{P}_\sigma \mathbb{I}_\Gamma)(u_\tau) = \mathbb{E}_u P_\sigma(u_\tau, \Gamma). \end{aligned}$$

Relation (1.58) is a straightforward consequence of (1.57).  $\square$

In conclusion, we discuss yet another version of the strong Markov property. Recall that  $C(\mathcal{T}_+, X)$  stands for the the space of continuous functions<sup>8</sup> from  $\mathcal{T}_+$  to  $X$  with the topology of uniform convergence on bounded subsets of  $\mathcal{T}_+$ ; cf. (1.9). In view of Exercise 1.1.6,  $C(\mathcal{T}_+, X)$  is a Polish space, and we endow it with its Borel  $\sigma$ -algebra.

*Exercise 1.3.12.* Let  $f: C(\mathcal{T}_+, X) \rightarrow \mathbb{R}$  be a measurable function that is either bounded or non-negative. Show that, for any  $u \in X$  and any stopping time  $\tau$ , we have

$$\mathbb{E}_u (\mathbb{I}_{\{\tau < \infty\}} f(u_{\tau+\cdot}) \mid \mathcal{F}_\tau) = \mathbb{I}_{\{\tau < \infty\}} (\mathbb{E}_v f(u_\cdot)) \Big|_{v=u_\tau} \quad \mathbb{P}_u\text{-almost surely.} \quad (1.59)$$

<sup>8</sup>In the case  $\mathcal{T}_+ = \mathbb{Z}_+$ , we obtain the space of all functions from  $\mathbb{Z}_+$  to  $X$ .

### 1.3.2 Random dynamical systems

In Example 1.3.6, we constructed a Markov chain starting from a dynamical system depending on a random parameter. That construction can be extended to more general situations, enabling one to associate a Markov process with some PDE subject to random perturbations. A key point here is the concept of a random dynamical system, which is introduced below. The relation between random dynamical systems and Markov processes is described in the next subsection.

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space, let  $\mathcal{T}$  be either  $\mathbb{R}$  or  $\mathbb{Z}$ , and let  $\theta = \{\theta_t : \Omega \rightarrow \Omega, t \in \mathcal{T}\}$  be a group of measurable mappings. We shall say that  $\theta$  is *measure-preserving* if  $(\theta_t)_* \mathbb{P} = \mathbb{P}$  for all  $t \in \mathcal{T}$ ; that is,

$$\mathbb{P}(\theta_t(\Gamma)) = \mathbb{P}(\Gamma) \quad \text{for any } t \in \mathcal{T}, \Gamma \in \mathcal{F}. \quad (1.60)$$

*Example 1.3.13.* Let  $U$  be a Polish space and let  $\mu$  be a probability measure on  $U$ . Define a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  by the following rules.

- $\Omega$  is the set of functions  $\omega : \mathcal{T} \rightarrow U$ ; we shall write  $\omega_t$  for the value of  $\omega$  at the point  $t$ .
- $\mathcal{F}$  is the minimal  $\sigma$ -algebra generated by the cylindrical sets

$$\Gamma = \{\omega \in \Omega : \omega_{t_i} \in \Gamma_i \text{ for } i = 1, \dots, n\}, \quad (1.61)$$

where  $t_1, \dots, t_n \in \mathcal{T}$  and  $\Gamma_1, \dots, \Gamma_n \in \mathcal{B}(U)$ .

- $\mathbb{P}$  is the unique probability measure on  $(\Omega, \mathcal{F})$  such that

$$\mathbb{P}(\Gamma) = \prod_{i=1}^n \mu(\Gamma_i)$$

for any cylindrical set of the form (1.61).

With a slight abuse of notation, we shall denote by the same symbol the completion of the probability space defined above.

Let  $\theta = \{\theta_t : \Omega \rightarrow \Omega\}$  be the family of shifts on  $\Omega$ :

$$(\theta_t \omega)_s = \omega_{t+s} \quad \text{for } s \in \mathcal{T}.$$

Then  $\theta$  is a group of measure-preserving transformations on  $\Omega$ . We refer the reader to the book [Str93] for more details on this construction.

Below in this section, given a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , we shall always assume that a group of measure-preserving transformations  $\theta = \{\theta_t, t \in \mathcal{T}\}$  acts on it. Let  $\mathcal{T}_+ = \{t \in \mathcal{T} : t \geq 0\}$  and let  $X$  be a Polish space endowed with its Borel  $\sigma$ -algebra  $\mathcal{B}(X)$ . Consider a family  $\Phi = \{\varphi_t : \Omega \times X \rightarrow X, t \in \mathcal{T}_+\}$  of measurable mappings.

**Definition 1.3.14.** We shall say that  $\Phi$  is a (continuous) *random dynamical system* (or an *RDS*) in  $X$  over  $\theta$  if the following properties are satisfied.

- (i) *Continuity.* For any  $\omega \in \Omega$  and  $t \in \mathcal{T}_+$ , the mapping  $\varphi_t^\omega : X \rightarrow X$  is continuous.
- (ii) *Cocycle property.* For any  $t, s \in \mathcal{T}_+$  and  $\omega \in \Omega$ , we have

$$\varphi_0^\omega = \text{Id}_X, \quad \varphi_{t+s}^\omega = \varphi_t^{\theta^s \omega} \circ \varphi_s^\omega. \quad (1.62)$$

If  $\mathcal{T} = \mathbb{R}$ , we assume in addition that the trajectories  $\varphi_t^\omega(u)$  are continuous in  $t \in \mathcal{T}_+$  for any  $\omega \in \Omega$  and  $u \in X$ . If  $\mathcal{T} = \mathbb{Z}$ , we shall sometimes say  $\Phi$  is a *discrete-time RDS*, to emphasise the difference between the two cases.

In what follows, we often omit  $\omega$  from notation and write  $\varphi_t u$  instead of  $\varphi_t^\omega(u)$ . The following example gives a canonical way for constructing an RDS associated with a random perturbation of a deterministic mapping.

*Example 1.3.15.* In the setting of Example 1.3.6, let us construct a discrete-time RDS in  $H$  whose trajectories have the same distribution as  $\{v_k, k \in \mathbb{Z}_+\}$ .

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be the complete probability space defined in Example 1.3.13, where  $\mathcal{T} = \mathbb{Z}$ ,  $U = H$ , and  $\mu$  is the law of  $\eta_k$ . Let  $\theta = \{\theta_k : \Omega \rightarrow \Omega\}$  be the corresponding group of shifts on  $\Omega$ . We now define measurable mappings  $\varphi_k : \Omega \times H \rightarrow H$  by the formulas

$$\varphi_1^\omega u = S(u) + \omega_1, \quad \varphi_k^\omega u = S(\varphi_{k-1}^\omega u) + \omega_k, \quad k \geq 1, \quad (1.63)$$

where  $\omega = (\omega_j, j \in \mathbb{Z})$ . It is easy to verify that  $\Phi = \{\varphi_k, k \in \mathbb{Z}_+\}$  is an RDS over  $\theta$ . Moreover, for any  $u \in H$  the distribution of the corresponding trajectory  $\{u_k\}$  under the law  $\mathbb{P}$  coincides with that of the sequence  $\{v_k\}$  defined by (1.52) with  $v = u$ .

An important class of RDS is formed by those possessing an additional property of independence of the past and the future. To each such RDS there corresponds a Markov processes. We now turn to a description of those RDS and a construction of associated Markov processes.

### 1.3.3 Markov RDS

As before, let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space on which acts a group  $\theta = \{\theta_t, t \in \mathcal{T}\}$  of measure-preserving transformations, let  $X$  be a Polish space, and let  $\Phi = \{\varphi_t^\omega, t \in \mathcal{T}_+\}$  be an RDS in  $X$  over  $\theta$ . For any  $p, q \in \mathcal{T}$  with  $p < q$ , we denote by  $\mathcal{F}_{[p,q]} \subset \mathcal{F}$  the sub- $\sigma$ -algebra generated by the subsets of  $\mathcal{F}$  of zero measure and the  $X$ -valued random variables  $\varphi_t^{\theta^s \omega} u$ , where

$$u \in X, \quad t, s \in \mathcal{T}, \quad p \leq s < q, \quad 0 < t \leq q - s. \quad (1.64)$$

The definition implies that, for any  $t, s \in \mathcal{T}$  and  $u \in X$ , the random variable  $\varphi_t^{\theta^s \omega} u$  is  $\mathcal{F}_{[s, s+t]}$ -measurable. We also define the  $\sigma$ -algebras

$$\begin{aligned} \mathcal{F}_{[-\infty, q]} &= \sigma(\mathcal{F}_{[p,q]} : p \in \mathcal{T}, p < q), \\ \mathcal{F}_{[p, +\infty]} &= \sigma(\mathcal{F}_{[p,q]} : q \in \mathcal{T}, q > p), \\ \mathcal{F}_{[-\infty, +\infty]} &= \sigma(\mathcal{F}_{[p,q]} : p, q \in \mathcal{T}, p < q). \end{aligned}$$

*Exercise 1.3.16.* Describe the  $\sigma$ -algebra  $\mathcal{F}_{[p,q]}$  for the RDS of Example 1.3.15.

*Hint:*  $\mathcal{F}_{[p,q]}$  coincides with the  $\sigma$ -algebra of the sets of the form

$$\{(\omega_j, j \in \mathbb{Z}) : (\omega_{p+1}, \dots, \omega_q) \in \Gamma\}, \quad \Gamma \in \underbrace{\mathcal{B}(H) \otimes \dots \otimes \mathcal{B}(H)}_{q-p \text{ times}}. \quad (1.65)$$

generated by  $\omega_k$ ,  $p+1 \leq k \leq q$  and the subsets of zero measure.

*Exercise 1.3.17.* Show that for any finite or infinite numbers  $p < q$ , we have

$$\theta_t^{-1}(\mathcal{F}_{[p,q]}) = \mathcal{F}_{[p+t, q+t]}, \quad t \in \mathcal{T}.$$

We now introduce an important concept of Markov RDS. Let us set

$$\mathcal{F}^- = \mathcal{F}_{[-\infty, 0]}, \quad \mathcal{F}^+ = \mathcal{F}_{[0, +\infty]}.$$

The  $\sigma$ -algebras  $\mathcal{F}^-$  and  $\mathcal{F}^+$  are called the *past* and the *future* of  $\Phi$ , respectively.

**Definition 1.3.18.** The RDS  $\Phi$  is said to be *Markov* if its past and future are independent.

*Exercise 1.3.19.* Show that the discrete-time RDS defined in Example 1.3.15 is Markov.

*Exercise 1.3.20.* Let  $\mathcal{F}_t^- = \mathcal{F}_{[-\infty, t]}$  and  $\mathcal{F}_t^+ = \mathcal{F}_{[t, +\infty]}$ . Show that, in the case of Markov RDS, the  $\sigma$ -algebras  $\mathcal{F}_t^-$  and  $\mathcal{F}_t^+$  are independent for any  $t \in \mathcal{T}$ .

In what follows, we assume that the RDS  $\Phi$  under study is Markov. We write  $\mathcal{F}_t$  instead of  $\mathcal{F}_t^-$  and call  $\{\mathcal{F}_t, t \in \mathcal{T}\}$  the *filtration generated by  $\Phi$* . Note that the very definition of  $\mathcal{F}_t$  implies that the filtered probability space  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  satisfies the usual hypotheses.

An important property of Markov RDS is that the distribution of the random variable  $\varphi_t u$ , where  $u$  is an  $X$ -valued random variable, depends only on the distribution of the initial state  $u$ , provided that the latter is  $\mathcal{F}_0$ -measurable. In other words, a Markov RDS defines an evolution in the space of probability measures on the phase space  $X$ . To prove this assertion, we shall need the Markov property; cf. (1.45).

**Proposition 1.3.21.** *Let  $f : X \rightarrow \mathbb{R}$  be a bounded measurable function. Then, for any  $\mathcal{F}_0$ -measurable random variable  $u : \Omega \rightarrow X$  and any  $s, t \in \mathcal{T}_+$ , we have*

$$\mathbb{E}\{f(\varphi_{s+t}u) \mid \mathcal{F}_s\} = \{\mathbb{E}f(\varphi_t v)\}_{v=\varphi_s^\omega u} \quad \mathbb{P}\text{-almost surely}. \quad (1.66)$$

*Proof.* Let us consider a function  $g : \Omega \times X \rightarrow \mathbb{R}$  defined as

$$g(\omega, v) = f(\varphi_t^\theta \omega v), \quad \omega \in \Omega, \quad v \in X.$$

The definition of the  $\sigma$ -algebras  $\mathcal{F}_{[p,q]}$  implies that, for any  $v \in X$ , the function  $g(\omega, v)$  is  $\mathcal{F}_s^+$ -measurable and that  $\varphi_s^\omega u$  is  $\mathcal{F}_s$ -measurable. Since  $\mathcal{F}_s$  and  $\mathcal{F}_s^+$



are independent, it follows that (for instance, see Problem 9 in Section 10.1 of [Dud02])

$$\mathbb{E}\{g(\omega, \varphi_s^\omega u) \mid \mathcal{F}_s\} = \{\mathbb{E}g(\omega, v)\}_{v=\varphi_s^\omega u} \quad \text{almost surely.} \quad (1.67)$$

On the other hand, in view of the cocycle property, we have

$$g(\omega, \varphi_s^\omega u) = f(\varphi_t^{\theta_s \omega} \circ \varphi_s^\omega u) = f(\varphi_{s+t}^\omega u). \quad (1.68)$$

Combining (1.67) and (1.68), we obtain (1.66).  $\square$

We now introduce the *transition function* for  $\Phi$ ; cf. (1.46). For any  $v \in X$  and  $t \in \mathcal{T}_+$ , we denote by  $P_t(v, \cdot)$  the law of  $\varphi_t v$ :

$$P_t(v, \Gamma) = \mathbb{P}\{\varphi_t v \in \Gamma\}, \quad v \in X, \quad \Gamma \in \mathcal{B}(X). \quad (1.69)$$

The following result is the analogue for Markov RDS of the property described in Exercise 1.3.4.

**Corollary 1.3.22.** *Let  $u : \Omega \rightarrow X$  be an  $\mathcal{F}_0$ -measurable random variable and let  $\mu = \mathcal{D}(u)$ . Then the distribution  $\mu_t$  of  $\varphi_t^\omega u$  is given by the formula*

$$\mu_t(\Gamma) = \int_X P_t(v, \Gamma) \mu(dv). \quad (1.70)$$

*In particular, the measure  $\mu_t$  depends only on  $\mu$  (and not on the random variable  $u$ ).*

*Proof.* In view of (1.66) with  $f = \mathbb{I}_\Gamma$ , for any  $\Gamma \in \mathcal{B}(X)$ , we have

$$\mathbb{E} \mathbb{I}_\Gamma(\varphi_t u) = \mathbb{E} \{ \mathbb{E}(\mathbb{I}_\Gamma(\varphi_t u) \mid \mathcal{F}_0) \} = \mathbb{E} \{ (\mathbb{E} \mathbb{I}_\Gamma(\varphi_t v)) \}_{v=u}.$$

It remains to note that  $\mathbb{E} \mathbb{I}_\Gamma(\varphi_t v) = \mathbb{P}\{\varphi_t v \in \Gamma\} = P_t(v, \Gamma)$ .  $\square$

Proposition 1.3.21 implies that the transition function 1.69 satisfies the Kolmogorov–Chapman relation. So it is the transition function of some Markov process; see Section III.1 in [RY99]. Using the Markov RDS  $\Phi$ , it is possible to construct a Markov process with this transition function in a canonical way. We now present this construction.

Let us denote by  $\Omega'$  the product space  $X \times \Omega$  endowed with the product  $\sigma$ -algebra  $\mathcal{B}(X) \otimes \mathcal{F}$ , and introduce the filtration  $\mathcal{F}'_t = \mathcal{B}(X) \otimes \mathcal{F}_t$ . Let us define the process  $u_t(\omega') = \varphi_t^\omega u$ , where  $\omega' = (u, \omega) \in \Omega'$ , and the family of probability measures  $\mathbb{P}_u = \delta_u \otimes \mathbb{P}$ , where  $\delta_u \in \mathcal{P}(X)$  is the Dirac measure concentrated at  $u$ . It is straightforward to see that  $(u_t, \mathbb{P}_u)$  is a Markov family whose transition function coincides with (1.69). We denote by  $\mathfrak{P}_t$  and  $\mathfrak{P}_t^*$  the corresponding Markov semigroups. The continuity of  $\varphi_t u$  with respect to  $u$  implies that  $(u_t, \mathbb{P}_u)$  possesses the Feller property; cf. Exercise 1.3.23 below. In what follows, we shall often drop the prime from notation and write  $\omega, \Omega, \mathcal{F}, \mathcal{F}_t$  instead of  $\omega', \Omega', \mathcal{F}', \mathcal{F}'_t$ . This will not lead to a confusion.

The following exercise implies, in particular, that the strong Markov property is true for Markov processes corresponding to continuous Markov RDS. In particular, relations (1.57) – (1.59) hold for them.

*Exercise 1.3.23.* Show that if  $\Phi$  is a continuous Markov RDS, then the corresponding Markov process possesses the Feller property. Combine this fact with Theorem 1.3.10 to conclude that the strong Markov property holds. *Hint:* Use Lebesgue's theorem on dominated convergence.

We conclude this subsection by a simple exercise describing the Markov semigroups in the case of deterministic dynamical systems.

*Exercise 1.3.24.* Let  $\Phi = \{\varphi_k^\omega : X \rightarrow X\}$  be an RDS such that

$$\varphi_1^\omega u = \psi(u) \quad \text{for any } \omega \in \Omega, u \in X,$$

where  $\psi : X \rightarrow X$  is a continuous operator. Describe the Markov semigroups  $\mathfrak{P}_k$  and  $\mathfrak{P}_k^*$  in terms of  $\psi$ .

### 1.3.4 Invariant and stationary measures

To every RDS  $\Phi = \{\varphi_t^\omega, t \in \mathcal{T}_+\}$  there corresponds a semigroup of measurable mappings in an extended phase space. Namely, we consider the product space  $(\Omega \times X, \mathcal{F} \otimes \mathcal{B}(X))$  and define a family of mappings  $\Theta = \{\Theta_t, t \in \mathcal{T}_+\}$  by the relation

$$\Theta_t(\omega, u) = (\theta_t \omega, \varphi_t^\omega u), \quad (\omega, u) \in \Omega \times X.$$

*Exercise 1.3.25.* Show that the family  $\Theta$  is a semigroup of measurable transformations.

We now introduce the concept of an invariant measure for  $\Phi$ . Let  $\mathcal{P}(\Omega \times X, \mathbb{P})$  be the set of probability measures on  $(\Omega \times X, \mathcal{F} \otimes \mathcal{B}(X))$  whose projection to  $\Omega$  coincides with  $\mathbb{P}$ .

**Definition 1.3.26.** A measure  $\mathfrak{M} \in \mathcal{P}(\Omega \times X, \mathbb{P})$  is said to be *invariant* for  $\Phi$  if  $(\Theta_t)_* \mathfrak{M} = \mathfrak{M}$  for any  $t \in \mathcal{T}_+$ . In other words,  $\Theta$  is a semigroup of measure-preserving transformations of the space  $(\Omega \times X, \mathcal{F} \otimes \mathcal{B}(X), \mathfrak{M})$  (cf. (1.60)).

In what follows, we shall need an alternative description of invariant measures for  $\Phi$ . Recall that a family  $\{\mu_\omega, \omega \in \Omega\} \subset \mathcal{P}(X)$  is called a *random probability measure* if for any  $\Gamma \in \mathcal{B}(X)$  the function  $\omega \mapsto \mu_\omega(\Gamma)$  from  $\Omega$  to  $\mathbb{R}$  is  $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$ -measurable. It is well known that any measure  $\mathfrak{M} \in \mathcal{P}(\Omega \times X, \mathbb{P})$  admits a *disintegration*

$$\mathfrak{M}(d\omega, du) = \mu_\omega(du) \mathbb{P}(d\omega), \quad (1.71)$$

where  $\{\mu_\omega\}$  is a random probability measure on  $X$ . Relation (1.71) means that if  $f : \Omega \times X \rightarrow \mathbb{R}$  is a bounded measurable function, then

$$\iint_{\Omega \times X} f(\omega, u) \mathfrak{M}(d\omega, du) = \int_{\Omega} \left\{ \int_X f(\omega, u) \mu_\omega(du) \right\} \mathbb{P}(d\omega).$$

Moreover, representation (1.71) is unique up to a set of zero measure in the sense that if  $\{\mu'_\omega, \omega \in \Omega\}$  is another random probability measure for which (1.71)

holds, then  $\mu_\omega = \mu'_\omega$  for almost every  $\omega \in \Omega$ . We refer the reader to Section 10.2 of [Dud02] for the proof of these results. Note that if  $\mathfrak{M}$  is represented in the form (1.71), then the projection of  $\mathfrak{M}$  to  $X$  coincides with the mean value of  $\mu_\omega$ . That is,

$$(\Pi_X)_* \mathfrak{M} = \mathbb{E} \mu. := \int_{\Omega} \mu_\omega \mathbb{P}(d\omega).$$

**Proposition 1.3.27.** *A measure  $\mathfrak{M} \in \mathcal{P}(\Omega \times X, \mathbb{P})$  is invariant for  $\Phi$  if and only if for any  $t \in \mathcal{T}_+$  its disintegration  $\{\mu_\omega\}$  satisfies the relation*

$$(\varphi_t^\omega)_* \mu_\omega = \mu_{\theta_t \omega} \quad \text{for almost all } \omega \in \Omega. \quad (1.72)$$

*Proof.* Suppose that  $\mathfrak{M}$  is an invariant measure for  $\Phi$ . Then for any bounded measurable functions  $f : X \rightarrow \mathbb{R}$  and  $g : \Omega \rightarrow \mathbb{R}$  we have

$$\int_{\Omega} \int_X f(\varphi_t^\omega u) g(\theta_t \omega) \mu_\omega(du) \mathbb{P}(d\omega) = \int_{\Omega} \int_X f(u) g(\omega) \mu_\omega(du) \mathbb{P}(d\omega). \quad (1.73)$$

Replacing  $g(\omega)$  by  $g(\theta_{-t}\omega)$  and using the invariance of  $\mathbb{P}$  with respect to  $\theta$ , we see that

$$\int_{\Omega} \left\{ \int_X f(\varphi_t^\omega u) \mu_\omega(du) \right\} g(\omega) \mathbb{P}(d\omega) = \int_{\Omega} \left\{ \int_X f(u) \mu_{\theta_t \omega}(du) \right\} g(\omega) \mathbb{P}(d\omega).$$

Since this relation is true for any function  $g$ , we conclude that

$$\int_X f(\varphi_t^\omega u) \mu_\omega(du) = \int_X f(u) \mu_{\theta_t \omega}(du) \quad \text{for almost all } \omega \in \Omega. \quad (1.74)$$

Now note that the Borel  $\sigma$ -algebra on a Polish space is countably generated (see Corollary 7.1.3 in the Appendix). Therefore (1.74) implies that (1.72) holds for any  $t \in \mathcal{T}_+$ .

Conversely, if (1.72) holds, then reversing the above arguments, we arrive at (1.73) for any bounded measurable functions  $f$  and  $g$ . Application of the monotone class technique (see Theorem 7.1.1 in the Appendix) shows that  $\mathfrak{M}$  is invariant for  $\Phi$ .  $\square$

We conclude this subsection with the concept of a stationary measure for a Markov RDS. Recall that the Markov semigroups for such an RDS were introduced in the foregoing subsection.

**Definition 1.3.28.** A measure  $\mu \in \mathcal{P}(X)$  is said to be *stationary* for an RDS  $\Phi$  if it is stationary for the corresponding Markov process, i.e.,  $\mathfrak{P}_t^* \mu = \mu$  for any  $t \in \mathcal{T}_+$ .

Note that, in contrast to invariant measures, stationary measures are defined only for Markov RDS.

*Exercise 1.3.29.* Let  $\mathcal{F}_t$  be the filtration generated by  $\Phi$  and let  $u : \Omega \rightarrow X$  be an  $\mathcal{F}_0$ -measurable random variable such that  $\mathcal{D}(u) = \mu$ . Show that if  $\mu$  is a stationary measure for  $\Phi$ , then

$$\mathcal{D}(\varphi_t u) = \mu \quad \text{for all } t \in \mathcal{T}_+.$$

Let  $\mathcal{A} \subset X$  be a closed subset. We shall say that  $\mathcal{A}$  is *absorbing* for the RDS  $\Phi$  if for any  $u \in X$  there is an almost surely finite random time  $T_u \in \mathcal{T}_+$  such that, with probability 1, we have

$$\varphi_t^\omega u \in \mathcal{A} \quad \text{for any } t \geq T_u.$$

**Lemma 1.3.30.** *If  $\mathcal{A}$  is an absorbing set for the RDS  $\Phi$  and  $\mu$  is a stationary measure for  $\Phi$ , then  $\text{supp } \mu \subset \mathcal{A}$ .*

*Proof.* We need to show that if a function  $f \in C_b(X)$  vanishes on  $\mathcal{A}$ , then  $(f, \mu) = 0$ . To this end, note that

$$(f, \mu) = (f, \mathfrak{P}_t^* \mu) = (\mathfrak{P}_t f, \mu) = \int_H \mathfrak{P}_t f(u) \mu(du).$$

Since  $\mathcal{A}$  is an absorbing set, we have  $f(\varphi_t u) = 0$  for  $t \geq T_u$ , whence, by Lebesgue's theorem on dominated convergence, we conclude that

$$\mathfrak{P}_t f(u) = \mathbb{E} f(\varphi_t u) \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

Hence,  $(f, \mu) = 0$  for any  $f \in C_b(X)$  vanishing on  $\mathcal{A}$ . This completes the proof of the lemma.  $\square$

We conclude this chapter by a result showing that a global stability of a Markov RDS implies the uniqueness of a stationary measure. For simplicity, we confine ourselves to the case of a separable Banach space  $X$  with a norm  $\|\cdot\|$ .

**Proposition 1.3.31.** *Let  $\Phi = \{\varphi_t, t \in \mathcal{T}_+\}$  be a Markov RDS in  $X$  such that*

$$\|\varphi_t^\omega u - \varphi_t^\omega v\| \leq \psi_{u,v} \alpha(t) \|u - v\|, \quad t \in \mathcal{T}_+, \quad \omega \in \Omega, \quad u, v \in X, \quad (1.75)$$

where  $R > 0$  is arbitrary,  $\psi_{u,v}$  is an almost surely finite random variable measurable with respect to its arguments, and  $\alpha(t)$  is a function going to zero as  $t \rightarrow +\infty$ . Then  $\Phi$  has at most one stationary measure. Moreover, if a stationary measure  $\mu$  exists, then for any  $R \geq 1$ ,  $u \in X$ , and  $t \in \mathcal{T}_+$ , we have

$$\|P_t(u, \cdot) - \mu\|_L^* \leq 2\mu(B_X(R)^c) + \int_{B_X(R)} \mathbb{E} \left( 2 \wedge (\psi_{u,v} \alpha(t) \|u - v\|) \right) \mu(dv). \quad (1.76)$$

*Proof.* Let  $\mu, \nu \in \mathcal{P}(X)$  be two stationary distributions. Then for any bounded measurable function  $f : X \rightarrow \mathbb{R}$  we have

$$(f, \mu) - (f, \nu) = (\mathfrak{P}_t f, \mu) - (\mathfrak{P}_t f, \nu) = \iint_{X \times X} \mathbb{E} (f(\varphi_t u) - f(\varphi_t v)) \mu(du) \nu(dv).$$

Assuming that  $f \in L_b(X)$  with  $\|f\|_L \leq 1$  and using inequality (1.76), for any  $R > 0$  we derive

$$|(f, \mu) - (f, \nu)| \leq 2\mu(B_X^c(R)) + 2\nu(B_X^c(R)) + 2 \iint_{X \times X} \mathbb{E} (1 \wedge R \psi_{u,v} \alpha(t)) \mu(du) \mu(dv).$$

Passing to the limit as  $t \rightarrow +\infty$  and then  $R \rightarrow +\infty$ , we see that  $(f, \mu) = (f, \nu)$  for any  $f \in L_b(X)$ . This implies that  $\mu = \nu$ .

We now prove (1.76). To this end, we essentially repeat the above argument. Namely, take an arbitrary function  $f \in C_b(X)$  and write

$$(f, P_t(u, \cdot)) - (f, \mu) = (\mathfrak{P}_t f, \delta_u) - (\mathfrak{P}_t f, \mu) = \int_X \mathbb{E} (f(\varphi_t u) - f(\varphi_t v)) \mu(dv).$$

Using (1.75) to bound the right-hand side of this relation and taking the supremum over all  $f \in L_b(X)$  with  $\|f\|_L \leq 1$ , we arrive at the required result.  $\square$

## Notes and comments

The basic facts about Sobolev spaces used in this monograph are very well known, and their proof is widely available in the literature; see the books of Lions, Magenes [LM72] and Taylor [Tay97]. A comprehensive treatment of the theory of Sobolev spaces can be found, for instance, in the books [Ada75, Maz85] by Adams and Maz'ja. Classes of measurable functions with range in a Banach or Hilbert space and the concept of Bochner's integral on an abstract measurable space are studied in Chapter V of [Yos95], and a rich source of their applications in the theory of nonlinear PDE's is the book [Lio69].

An excellent book about the measure theory on metric spaces is the one by Dudley [Dud02]. A lot of useful information can also be found in Bogachev's two-volume book [Bog07]. The idea of coupling goes back to Doebelin [Doe38, Doe40], but its systematic application in the theory of stochastic processes started much later; see the papers [Har55, Pit74] by Harris and Pitman. Lemma 1.2.24 was often used by Dobrushin [Dob68, Dob74] in his celebrated work on Gibbs systems.

The general theory of random dynamical systems and Markov processes is treated in many textbooks and monographs; e.g., see the books of Revuz [Rev84], Kifer [Kif86], Karatzas, Shreve [KS91], and L. Arnold [Arn98]. Our presentation of this subject essentially follows one of these references.



## Chapter 2

# Two-dimensional Navier–Stokes equations

In this chapter we present some well-known results on the 2D Navier–Stokes equations. We begin with the deterministic case. Galerkin approximations are used to prove existence, uniqueness, and regularity of solutions. Some important estimates due to Foias and Prodi are also established. We next consider the Navier–Stokes system perturbed by a random external force and prove well-posedness of the Cauchy problem and some *a priori* estimates for solutions. Combining these results with the classical Bogolyubov–Krylov argument, it is then shown that the randomly forced Navier–Stokes equations have at least one stationary measure.

### 2.1 Cauchy problem for the deterministic system

#### 2.1.1 Equations and boundary conditions

The 2D Navier–Stokes (NS) equations have the form

$$\begin{cases} \dot{u} + \langle u, \nabla \rangle u - \nu \Delta u + \nabla p = f(t, x), \\ \operatorname{div} u = 0. \end{cases} \quad (2.1)$$

Here  $u = (u^1, u^2)$  and  $p$  are unknown velocity field and pressure,  $\nu > 0$  is the kinematic viscosity,  $f$  is the density of an external force, and  $\langle u, \nabla \rangle$  stands for the differential operator  $u^1 \partial_1 + u^2 \partial_2$ . Equations (2.1) are considered either in a bounded domain  $Q \subset \mathbb{R}^2$  with a smooth boundary  $\partial Q$  or on the standard torus  $\mathbb{T}^2 = \mathbb{R}^2 / 2\pi\mathbb{Z}^2$ . The latter means that the space variable  $x = (x_1, x_2)$  belongs to  $\mathbb{R}^2$ , and the functions  $u$ ,  $p$ , and  $f$  are assumed to be  $2\pi$ -periodic with respect to  $x_i$ ,  $i = 1, 2$ . All the results and proofs remain valid without any change

for non-standard tori  $\mathbb{T}^2/a(\mathbb{Z} \oplus b\mathbb{Z})$ , where  $a, b > 0$ . In the case of a bounded domain, Eqs. (2.1) are supplemented with the no-slip boundary condition:

$$u|_{\partial Q} = 0. \quad (2.2)$$

The Cauchy problem for the Navier–Stokes system consists in finding a pair of functions  $(u, p)$  of appropriate regularity that satisfy (2.1) and (2.2) in a sense to be specified, as well as the initial condition

$$u(0, x) = u_0(x), \quad (2.3)$$

where  $u_0$  is a given function. In the case of torus, the boundary condition (2.2) should be omitted.

In what follows, to simplify the presentation, results of this chapter usually are proved for the torus. If the corresponding proof in the case of a bounded domain is different, or the result is not valid, we mention this explicitly.

### 2.1.2 Leray decomposition

Let  $H^m(\mathbb{T}^2; \mathbb{R}^2)$  be the space of vector fields on  $\mathbb{T}^2$  whose components belong to the Sobolev space of order  $m$  and let

$$H_\sigma^m = \{u \in H^m(\mathbb{T}^2; \mathbb{R}^2) : \operatorname{div} u = 0 \text{ on } \mathbb{T}^2\},$$

where the divergence is taken in the sense of distributions. It is clear that  $H_\sigma^m$  is a closed subspace of  $H^m(\mathbb{T}^2; \mathbb{R}^2)$ .

The following result can be easily established with the help of Fourier expansions of square-integrable functions.

*Exercise 2.1.1.* Show that  $H_\sigma^m$  coincides with the closure in  $H^m(\mathbb{T}^2; \mathbb{R}^2)$  of the space

$$\mathcal{V} = \{u \in C^\infty(\mathbb{T}^2; \mathbb{R}^2) : \operatorname{div} u = 0 \text{ on } \mathbb{T}^2\}.$$

Let us recall that  $\dot{H}^m = \dot{H}^m(\mathbb{T}^2)$  stands for the space of functions in  $H^m(\mathbb{T}^2)$  with zero mean value (see (1.2)), and let  $\nabla H^{m+1}$  be the space of functions  $u \in H^m(\mathbb{T}^2; \mathbb{R}^2)$  that are representable in the form  $u = \nabla p$  for some  $p \in H^{m+1}$ . Since  $\|\nabla p\|_k$  is a norm on  $\dot{H}^{m+1}$ , we see that  $\nabla H^{m+1}$  is a closed subspace in  $H^m(\mathbb{T}^2; \mathbb{R}^2)$ . The following result due to Helmholtz is a common tool in the theory of Navier–Stokes equations since the work of Leray.

**Theorem 2.1.2.** *For any  $m \in \mathbb{R}$ , the space  $H^m(\mathbb{T}^2; \mathbb{R}^2)$  admits the direct decomposition*

$$H^m(\mathbb{T}^2; \mathbb{R}^2) = H_\sigma^m \dot{+} \nabla H^{m+1}. \quad (2.4)$$

Moreover, the sum is orthogonal for  $m = 0$ , and decompositions (2.4) corresponding to two different values  $m_1$  and  $m_2$  give the same representation for any function belonging to the intersection  $H^{m_1}(\mathbb{T}^2; \mathbb{R}^2) \cap H^{m_2}(\mathbb{T}^2; \mathbb{R}^2)$ .



*Proof.* We first note that the intersection of  $H_\sigma^m$  and  $\nabla H^{m+1}$  consists of the zero function. Let us expand a function  $u \in H^m(\mathbb{T}^2; \mathbb{R}^2)$  into the Fourier series:

$$u(x) = \sum_{s \in \mathbb{Z}^2} u_s e^{isx}, \quad u_s \in \mathbb{C}^2.$$

Denoting  $s^\perp = (-s_2, s_1)$ , we write

$$u(x) = \sum_{s \in \mathbb{Z}^2} \frac{\langle u_s, s^\perp \rangle s^\perp}{|s|^2} e^{isx} + \sum_{s \in \mathbb{Z}^2} \frac{\langle u_s, s \rangle s}{|s|^2} e^{isx}. \quad (2.5)$$

To prove (2.4), it remains to note that the first sum on the right-hand side is an element of  $H_\sigma^m$ , while the second is the gradient of the function

$$-i \sum_{s \in \mathbb{Z}^2} \frac{\langle u_s, s \rangle}{|s|^2} e^{isx},$$

which belongs to  $H^{m+1}$ . The fact that (2.4) is an orthogonal sum for  $m = 0$  is easy to show with the help of integration by parts, and compatibility of decompositions (2.4) corresponding to different values of  $m$  follows from (2.5).  $\square$

In what follows, we denote by  $\Pi : H^m(\mathbb{T}^2; \mathbb{R}^2) \rightarrow H_\sigma^m$  the projection associated with decomposition (2.4). It is called the *Leray projection*. Even though the definition of  $\Pi$  depends on  $m$ , Theorem 2.1.2 implies that the function  $\Pi u$  depends only on  $u$ , and not on the space  $H^m(\mathbb{T}^2; \mathbb{R}^2)$  in which the operator  $\Pi$  is considered. Therefore we do not indicate the dependence of the projection operators on  $m$ .

An analogue of the Leray–Helmholtz (2.4) decomposition is true for any bounded Lipschitz domain. Let us formulate the corresponding result in the case  $m = 0$ . We set

$$L_\sigma^2(Q; \mathbb{R}^2) = \{u \in L^2(Q; \mathbb{R}^2) : \operatorname{div} u = 0 \text{ in } Q, \langle u, \mathbf{n} \rangle = 0 \text{ on } \partial Q\},$$

where  $\mathbf{n}$  denotes the outward unit normal to  $\partial Q$ . Furthermore, define  $\nabla H^1$  as the space of functions  $u \in L^2(Q; \mathbb{R}^2)$  that are representable in the form  $u = \nabla p$  for some  $p \in H^1(Q)$ . A proof of the following theorem can be found in [Tem79, Chapter 1].

**Theorem 2.1.3.** *The subspaces  $L_\sigma^2$  and  $\nabla H^1$  are closed in  $L^2(Q; \mathbb{R}^2)$ , and we have the orthogonal decomposition*

$$L^2(Q; \mathbb{R}^2) = L_\sigma^2 \oplus \nabla H^1. \quad (2.6)$$

As before, we denote by  $\Pi : L^2(Q; \mathbb{R}^2) \rightarrow L_\sigma^2$  the Leray projection. It can be shown that  $\Pi$  is continuous in  $H^1(Q; \mathbb{R}^2)$  and admits a continuous extension to an operator from  $H^{-1}(Q; \mathbb{R}^2)$  to the dual space of  $L_\sigma^2 \cap H_0^1(Q; \mathbb{R}^2)$ ; see Section I.1 in [Tem79]. We refer the reader to Chapter 17 of the book [Tay97] for a detailed presentation of the corresponding results in the case when  $Q$  is a compact manifold.

### 2.1.3 Properties of some multilinear maps

In this subsection, we establish some elementary estimates of linear and non-linear functionals relevant to the Navier–Stokes equations. We shall need two auxiliary lemmas.

**Lemma 2.1.4.** *Let  $X_1, \dots, X_r$  and  $Y$  be Banach spaces and let  $F$  be an  $r$ -linear map from the direct product  $X_1 \times \dots \times X_r$  to  $Y$  such that*

$$\|F(u_1, \dots, u_r)\|_Y \leq C \|u_1\|_{X_1} \dots \|u_r\|_{X_r} \quad \text{for all } u_j \in X_j, j = 1, \dots, r. \quad (2.7)$$

*Then  $F$  is continuous. Moreover, if  $V_i \subset X_i$ ,  $i = 1, \dots, r$ , are dense vector spaces,  $F$  is defined for  $u_j \in V_j$ , and inequality (2.7) holds for  $u_j \in V_j$ , then  $F$  extends to an  $r$ -linear continuous map from  $X_1 \times \dots \times X_r$  to  $Y$ .*

The proof of this lemma is straightforward. To formulate the second result, we introduce some notations. For any sufficiently regular vector fields  $u$  and  $v$  on the torus  $\mathbb{T}^2$ , we set

$$Lu = -\Pi\Delta u, \quad B(u, v) = \Pi(\langle u, \nabla \rangle v). \quad (2.8)$$

For the periodic boundary conditions,  $Lu$  equals  $-\Delta u$ , but this is not the case for bounded domains and Dirichlet boundary conditions. It is easy to check with the help of Fourier expansion that  $L$  is a self-adjoint operator in  $L^2_\sigma$  with the domain  $D(L) = H^2_\sigma$ . It is called *Stokes operator*.

Define the space

$$\mathcal{H} = \{u \in L^2(0, T; H^1_\sigma) : \dot{u} \in L^2(0, T; H^{-1}_\sigma)\}$$

and endow it with the norm

$$\|u\|_{\mathcal{H}} = \left( \int_0^T (\|u(t)\|_1^2 + \|\dot{u}(t)\|_{-1}^2) dt \right)^{1/2},$$

where the time derivative is understood in the sense of distributions. Note that  $\mathcal{H}$  is a Hilbert space. A proof of the following result can be found in [LM72] (see Theorem 3.1 in Chapter 1) and [Lio69] (see Theorem 5.1 in Chapter 1).

**Lemma 2.1.5.** *The space  $\mathcal{H}$  is continuously embedded in  $C(0, T; L^2_\sigma)$  and compactly embedded in  $L^2(0, T; H^m_\sigma)$  for any  $m < 1$ . In particular,  $u(t)$  is a well defined vector in  $L^2_\sigma$  for any  $u \in \mathcal{H}$  and  $t \in [0, T]$ .*

We now turn to some estimates that will be needed in the sequel. The first result essentially is an integration-by-parts formula.

**Proposition 2.1.6.** *For any  $u, v \in \mathcal{H}$  and  $0 \leq t \leq T$ , we have*

$$\int_0^T \langle Lu(t), v(t) \rangle dt = \int_0^T \langle \nabla u(t), \nabla v(t) \rangle dt, \quad (2.9)$$

$$\int_0^t \langle \dot{u}, u \rangle ds = \frac{1}{2} (|u(t)|_2^2 - |u(0)|_2^2). \quad (2.10)$$

*Proof.* In view of Lemma 2.1.4, it suffices to establish (2.9) and (2.10) for smooth functions  $u$  and  $v$ . In this case, both relations are obvious.  $\square$

Let us note that the function  $t \mapsto \langle \dot{u}, u \rangle$  belongs to  $L_1(0, T)$  for any  $u \in \mathcal{H}$ . It follows from (2.10) that  $|u(t)|^2$  is an absolutely continuous function, and

$$\frac{d}{dt} |u(t)|^2 = 2\langle \dot{u}, u \rangle \quad \text{for any } u \in \mathcal{H}.$$

**Proposition 2.1.7.** *For any divergence-free smooth vector fields  $u, v$ , and  $w$ , we have*

$$\langle B(u, v), v \rangle = 0, \quad (2.11)$$

$$\langle B(u, v), w \rangle = -\langle B(u, w), v \rangle, \quad (2.12)$$

$$|\langle B(u, v), w \rangle| \leq C \|u\|_{1/2} \|v\|_{1/2} \|w\|_1, \quad (2.13)$$

$$\|B(u, v)\|_{-1} \leq C \|u\|_{1/2} \|v\|_{1/2}, \quad (2.14)$$

$$\|B(u, v)\|_{-3} \leq C \|u\|_2 \|v\|_2, \quad (2.15)$$

where  $C > 0$  is a constant not depending on the functions.

*Proof.* Integrating by parts, we derive

$$\begin{aligned} \langle B(u, v), v \rangle &= \sum_{j,l=1}^2 \int_{\mathbb{T}^2} u^j (\partial_j v^l) v^l dx = \sum_{j,l=1}^2 \frac{1}{2} \int_{\mathbb{T}^2} u^j \partial_j |v|^2 dx \\ &= -\frac{1}{2} \int_{\mathbb{T}^2} (\operatorname{div} u) |v|^2 dx = 0. \end{aligned}$$

This proves (2.11). To establish (2.12), it suffices to use (2.11) with  $v$  replaced by  $v + w$ .

Let us prove (2.13). Applying consecutively relation (2.12), the Hölder inequality, and the continuous embedding  $H^{1/2} \subset L^4$  (see Property 1.1.2), we obtain

$$\begin{aligned} |\langle B(u, v), w \rangle| &= |\langle B(u, w), v \rangle| \leq C_1 \int_{\mathbb{T}^2} |u| |\nabla w| |v| dx \\ &\leq C_2 \|\nabla w\|_2 \|u\|_4 \|v\|_4 \leq C_3 \|w\|_1 \|u\|_{1/2} \|v\|_{1/2}. \end{aligned} \quad (2.16)$$

Inequality (2.14) follows from (2.13) by duality. Finally, to establish (2.15), it suffices to estimate the integral in (2.16) by the product  $\|\nabla w\|_\infty \|u\|_2 \|v\|_2$  and apply the Sobolev embedding  $H^2 \subset L^\infty$ .  $\square$

Combining Lemma 2.1.4 and Propositions 2.1.6 and 2.1.9, we obtain the following result.

**Corollary 2.1.8.** *The operator  $B$  defined initially on smooth vector fields extends to a continuous bilinear map from  $H_\sigma^{1/2} \times H_\sigma^{1/2}$  to  $H_\sigma^{-1}$  and from  $L_\sigma^2 \times L_\sigma^2$  to  $H_\sigma^{-3}$ .*

Consider now a trilinear map  $b$  defined on smooth non-autonomous vector fields on  $\mathbb{T}^2$  by the relation

$$b(u, v, w) = \langle B(u(t, \cdot), v(t, \cdot)), w(t, \cdot) \rangle.$$

**Proposition 2.1.9.** *The map  $b(u, v, w)$  is continuous from  $\mathcal{H} \times \mathcal{H} \times L^2(0, T; H_\sigma^1)$  to  $L_1(0, T)$ , and*

$$b(u, v, v) = 0 \quad \text{for any } u, v \in \mathcal{H}, \quad (2.17)$$

where the equality holds in the space  $L_1(0, T)$ . In particular, for any  $u, v \in \mathcal{H}$  the function  $B(u, v)$  belongs to  $L^2(0, T; H_\sigma^{-1})$ .

*Proof.* We first note that  $\mathcal{H} \subset L^4(0, T; H^{1/2})$ . Indeed, since  $\|u\|_{1/2}^4 \leq |u|_2^2 \|u\|_1^2$ , we have

$$\|u\|_{L^4(0, T; H^{1/2})}^4 = \int_0^T \|u(t)\|_{1/2}^4 dt \leq \left( \sup_{0 \leq t \leq T} |u(t)|_2^2 \right) \int_0^T \|u(t)\|_1^2 dt \leq \|u\|_{\mathcal{H}}^4.$$

Hence, using (2.14), for smooth functions  $u, v \in \mathcal{H}$  and  $w \in L^2(0, T; H_\sigma^1)$ , we obtain

$$\begin{aligned} \int_0^T |b(u, v, w)| ds &\leq C \int_0^T \|w\|_1 \|u\|_{1/2} \|v\|_{1/2} ds \\ &\leq \left( \int_0^T \|w\|_1^2 ds \right)^{1/2} \left( \int_0^T \|u\|_{1/2}^4 ds \right)^{1/4} \left( \int_0^T \|v\|_{1/2}^4 ds \right)^{1/4} \\ &\leq C \|w\|_{L^2(0, T; H^1)} \|u\|_{\mathcal{H}} \|v\|_{\mathcal{H}}. \end{aligned} \quad (2.18)$$

Lemma 2.1.4 implies now the continuity of  $b$ . By duality, it follows from inequality (2.18) that  $B(u, v)$  belongs to  $L^2(0, T; H_\sigma^{-1})$ . Finally, relation (2.17), which is true for smooth functions in view of (2.11), extends to the general case by continuity.  $\square$

### 2.1.4 Reduction to an abstract evolution equation

We now introduce the concept of a solution for the Navier–Stokes system. Let us fix a constant  $T > 0$  and functions  $f \in L^2(0, T; H^{-1})$  and  $u_0 \in L_\sigma^2(\mathbb{T}^2, \mathbb{R}^2)$ . We shall denote by  $\mathcal{D}'$  the space of  $\mathbb{R}^2$ -valued distributions on  $(0, T) \times \mathbb{T}^2$ .

**Definition 2.1.10.** A pair of functions  $(u, p)$  is called a *solution of Equation (2.1)* on  $(0, T) \times \mathbb{T}^2$  if  $u \in \mathcal{H}$ ,  $p \in L^2(0, T; L^2)$ , and (2.1) holds in  $\mathcal{D}'$ . If, in addition, relation (2.3) is also satisfied, then  $(u, p)$  is called a *solution of the Cauchy problem (2.1), (2.3)*.

Let us note that (2.1) is not a system of evolution equations in the sense that it does not contain the time derivative of the unknown function  $p$ . However, it is possible to exclude the pressure from the problem in question and to obtain a nonlocal nonlinear PDE which can be regarded as an evolution equation in a Hilbert space.

To this end, let us apply formally the Leray projection  $\Pi$  to the first equation in (2.1). Since  $\Pi(\nabla p) = 0$ ,  $u(t) \in L^2_\sigma$  for any  $t$ , and  $\Pi\dot{u} = \dot{u}$ , using notation (2.8) we obtain

$$\dot{u} + \nu Lu + B(u) = \Pi f(t), \quad (2.19)$$

where  $B(u) = B(u, u)$ . Let us introduce the concept of a solution for (2.19).

**Definition 2.1.11.** A function  $u \in \mathcal{H}$  is called a *solution of Equation (2.19)* on  $(0, T) \times \mathbb{T}^2$  if (2.19) holds in  $\mathcal{D}'$ . If, in addition, the initial condition (2.3) is also satisfied, then  $u$  is called a *solution of the Cauchy problem (2.19), (2.3)*.

Note that the functions  $\Delta u$  and  $\langle u, \nabla \rangle u$  are elements of  $L^2(0, T; H^{-1})$ , and therefore their Leray projections are well defined and belong to  $L^2(0, T; H^{-1}_\sigma)$ . Thus, all the terms in (2.2) belong to  $\mathcal{D}'$ , and Definition 2.1.11 makes sense. The following result shows that the Cauchy problems for (2.1) and (2.19) are equivalent.

**Theorem 2.1.12.** *Let a pair  $(u, p) \in \mathcal{H} \times L^2(0, T; L^2)$  be a solution of (2.1) on  $(0, T) \times \mathbb{T}^2$ . Then  $u \in \mathcal{H}$  is a solution of (2.19). Conversely, if  $u \in \mathcal{H}$  is a solution of (2.19), then there is  $p \in L^2(0, T; L^2)$  such that  $(u, p)$  is a solution of (2.1).*

*Proof.* Let  $(u, p)$  be a solution for (2.1). Then integrating the first equation in (2.1) with respect to time, we obtain

$$u(t) = u_0 + \int_0^t (\nu \Delta u - \langle u, \nabla \rangle u - \nabla p + f) ds, \quad 0 \leq t \leq T, \quad (2.20)$$

where the equality holds in  $H^{-1}$ . Applying to (2.20) the Leray projection  $\Pi$ , which is continuous in  $H^{-1}$ , and using the fact that  $\Pi$  commutes with the integration in time, we obtain

$$u(t) = u_0 + \int_0^t (-\nu Lu - B(u) + \Pi f) ds, \quad 0 \leq t \leq T. \quad (2.21)$$

This implies that (2.19) holds in  $\mathcal{D}'$ . Conversely, let  $u \in \mathcal{H}$  be a solution of (2.19). Then  $\dot{u} \in L^2(0, T; H^{-1}_\sigma)$ . Consider the equation

$$\nabla p(t) = g(t) := -\dot{u} + \nu \Delta u - \langle u, \nabla \rangle u + f, \quad (2.22)$$

whose right-hand side  $g$  belongs to  $L^2(0, T; H^{-1})$ . If we show that

$$g(t) \in \nabla L^2 \quad \text{for almost all } t \in [0, T], \quad (2.23)$$

then we can resolve (2.23) with respect to  $p$  and find a solution  $p \in L^2(0, T; L^2)$ . In this case, the construction will imply that  $(u, p)$  satisfies (2.1) in the sense of distributions.

To prove (2.23), note that  $u$  satisfies (2.21). It follows that

$$\Pi \left( \int_0^t g(s) ds \right) = \int_0^t \Pi g(s) ds = 0 \quad \text{for } 0 \leq t \leq T.$$

whence we conclude that  $\Pi g(t) = 0$  for almost all  $t \in [0, T]$ . This implies the required relation (2.23).  $\square$

An analogue of Theorem 2.1.12 is true for the Navier–Stokes system in a bounded domain. However, its proof is more complicated, and we refer the reader to the book [Tem79] for the details.

### 2.1.5 Existence and uniqueness of solution

In this subsection, we prove that problem (2.19), (2.3) possesses a unique solution  $u \in \mathcal{H}$  for any functions  $u_0 \in L^2_\sigma$  and  $f \in L^2(0, T; H^{-1})$ . In view of Theorem 2.1.12, this will imply that problem (2.1), (2.3) is also well posed in an appropriate functional space.

**Theorem 2.1.13.** *Let  $u_0 \in L^2_\sigma$  and  $f \in L^2(0, T; H^{-1})$ . Then problem (2.19), (2.3) has a unique solution  $u \in \mathcal{H}$ , which satisfies the inequality*

$$|u(t)|_2^2 + \nu \int_0^t |\nabla u(s)|_2^2 ds \leq |u_0|_2^2 + \nu^{-1} \int_0^t \|f(s)\|_{-1}^2 ds, \quad 0 \leq t \leq T. \quad (2.24)$$

*Proof.* To simplify the presentation, we shall assume that  $\nu = 1$ .

*Uniqueness.* Let  $u_1, u_2 \in \mathcal{H}$  be two solutions equal to  $u_0$  at  $t = 0$ . Then the difference  $u = u_2 - u_1$  satisfies the equation

$$\dot{u} + Lu + B(u_2, u) + B(u, u_1) = 0. \quad (2.25)$$

Multiplying this relation in  $H$  by  $u(s)$  and integrating in time, we derive

$$\int_0^t (\langle \dot{u}, u \rangle + \langle Lu, u \rangle + \langle B(u, u_1), u \rangle) ds = 0. \quad (2.26)$$

Since  $\langle Lu, u \rangle = |\nabla u|_2^2$  and  $u(0) = 0$ , combining (2.26) with (2.10), (2.13) and using the second inequality in (1.8), we obtain

$$\begin{aligned} \frac{1}{2} |u(t)|_2^2 + \int_0^t |\nabla u(s)|_2^2 ds &= - \int_0^t \langle B(u, u_1), u \rangle ds \\ &\leq C_1 \int_0^t |\nabla u_1|_2 |\nabla u|_2 |u|_2 ds \\ &\leq \frac{1}{2} \int_0^t |\nabla u(s)|_2^2 ds + C_2 \int_0^t |\nabla u_1|_2^2 |u|_2^2 ds. \end{aligned}$$

It follows that

$$|u(t)|_2^2 + \int_0^t |\nabla u(s)|_2^2 ds \leq 2C_2 \int_0^t |\nabla u_1(s)|_2^2 |u(s)|_2^2 ds.$$

Since  $t \mapsto |\nabla u(t)|_2^2$  is an integrable function, Gronwall's lemma implies that  $u \equiv 0$ .

*Existence.* The proof of existence is based on a well-known general approach called *Galerkin method*. Roughly speaking, we construct solutions for finite-dimensional approximations of the equation and then show that they have a limit point which satisfies the original equation. The proof is divided into three steps.

*Step 1: A priori estimate.* Let  $u(t, x)$  be a *smooth* solution of (2.19). We consider the function  $u \mapsto \frac{1}{2} |u|_2^2$  and calculate its derivative along trajectories of (2.19):

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |u(t)|_2^2 &= \langle u(t), \dot{u}(t) \rangle = \langle u, -Lu - B(u) + \Pi f(t) \rangle \\ &= -|\nabla u|_2^2 + \langle u, f \rangle \leq -|\nabla u|_2^2 + |\nabla u|_2 \|f\|_{-1} \\ &\leq -\frac{1}{2} |\nabla u|_2^2 + \frac{1}{2} \|f\|_{-1}^2, \end{aligned}$$

where we used (2.11). Integrating in time, we obtain inequality (2.24) with  $\nu = 1$ . It follows that

$$\|u\|_{L^\infty(0,T;H)} + \|u\|_{L^2(0,T;H^1)} \leq |u_0|_2 + \|f\|_{L^2(0,T;H^{-1})} =: C_1(u_0, f). \quad (2.27)$$

Furthermore, using (2.14) and the interpolation inequality, we obtain

$$\|B(u)\|_{-1} \leq C \|u\|_{1/2}^2 \leq C |u| |\nabla u|_2,$$

whence it follows that

$$\|B(u)\|_{L^2(0,T;H^{-1})} \leq C_2(u_0, f).$$

Expressing  $\dot{u}$  in terms of  $u$  from (2.19) and recalling (2.27), we get

$$\|\dot{u}\|_{L^2(0,T;H^{-1})} \leq C_3(u_0, f).$$

We have thus shown that

$$\|u\|_{\mathcal{H}} \leq C(u_0, f). \quad (2.28)$$

*Step 2: Galerkin approximations.* Let us introduce the standard trigonometric basis in  $L_\sigma^2$ . Denote by  $Z_0^2$  the set of non-zero integer vectors  $s = (s_1, s_2)$ . For  $s \in Z_0^2$ , let

$$e_s = \begin{cases} c_s s^\perp \sin\langle s, x \rangle, & s \in Z_+^2, \\ c_s s^\perp \cos\langle s, x \rangle, & s \in Z_-^2, \end{cases} \quad (2.29)$$

where  $c_s = (\sqrt{2\pi}|s|)^{-1}$ ,  $s^\perp = (-s_2, s_1)$ ,  $Z_+^2$  stands for the set of vectors  $s \in Z_0^2$  such that either  $s_1 > 0$  or  $s_1 = 0$  and  $s_2 > 0$ , and  $Z_-^2$  is the complement of  $Z_+^2$  in  $Z_0^2$ . Given an integer  $N > 0$ , we write

$$H_{(N)} = \text{span}(\{e_s, |s| \leq N\} \cup \{(1, 0), (0, 1)\}) \quad (2.30)$$

and denote by

$$\mathbf{P}_N : L^2(\mathbb{T}^2; \mathbb{R}^2) \rightarrow L^2(\mathbb{T}^2; \mathbb{R}^2)$$

the orthogonal projection onto  $H_{(N)}$ . It is clear that

$$H_{(N)} \subset C^\infty \cap L^2_\sigma, \quad \dim H_{(N)} < \infty,$$

and that the operator  $L$  maps the subspace  $H_{(N)}$  into itself. Let us apply  $\mathbf{P}_N$  to (2.19):

$$\mathbf{P}_N \dot{u} + \mathbf{P}_N Lu + \mathbf{P}_N B(u) = \mathbf{P}_N f.$$

A curve  $u(t) \in H_{(N)}$  satisfies this relation if and only if

$$\dot{u} + Lu + \mathbf{P}_N B(u) = \mathbf{P}_N f. \quad (2.31)$$

This is an ODE in  $H_{(N)}$  defined by a smooth vector field, with a square-integrable right-hand side. Let us supplement it with the initial condition

$$u(0) = \mathbf{P}_N u_0. \quad (2.32)$$

Problem (2.31), (2.32) has a unique solution  $u = u_N$  defined on a time interval  $[0, T_N)$ , where either  $T_N = T$ , or  $T_N < T$  and  $u_N(t)$  blows up as  $t \rightarrow T_N^-$ . We claim that  $T_N = T$  and

$$u_N \in C(0, T; H_{(N)}), \quad \dot{u}_N \in L^2(0, T; H_{(N)}). \quad (2.33)$$

To prove these claims, we shall derive an a priori estimate for  $u_N$ .

Let us calculate the derivative of the functional  $\frac{1}{2}|u|^2$  along trajectories of (2.31). We have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |u(t)|^2 &= \langle u, \dot{u} \rangle = \langle u, -Lu - \mathbf{P}_N B(u) + \mathbf{P}_N f \rangle \\ &= \langle u, -Lu - B(u) + f(t) \rangle. \end{aligned}$$

Repeating the argument used in Step 1, we see that  $u = u_N$  satisfy inequalities (2.27) and (2.28) with  $T = T_N$  uniformly in  $N \geq 1$ . It follows that  $T_N = T$  for all  $N$ . Thus, the Galerkin approximations  $\{u_N\}$  form a bounded sequence in  $\mathcal{H}$ .

*Step 3: Passage to the limit.* We now wish to pass the limit in Eq. (2.31) as  $N \rightarrow \infty$ . Since  $\mathcal{H}$  is a Hilbert space, and a closed ball in a Hilbert space is weakly compact, there is a sequence  $N_j \rightarrow \infty$  such that  $\{u_{N_j}\}$  converges weakly to an element  $u \in \mathcal{H}$  as  $j \rightarrow \infty$ . Since the linear operators

$$\begin{aligned} \mathcal{H} &\rightarrow L_2(0, T; H_\sigma^{-1}), & u &\mapsto \dot{u}, \\ \mathcal{H} &\rightarrow L_2(0, T; H_\sigma^{-1}), & u &\mapsto Lu, \end{aligned}$$

are continuous, we have

$$\dot{u}_{N_j} \rightharpoonup \dot{u} \quad \text{in} \quad L_2(0, T; H_\sigma^{-1}), \quad (2.34)$$

$$L u_{N_j} \rightharpoonup Lu \quad \text{in} \quad L_2(0, T; H_\sigma^{-1}). \quad (2.35)$$



Consider now the nonlinear term  $\{B(u_{N_j})\}$ . By Lemma 2.1.5,

$$u_{N_j} \rightarrow u \quad \text{in } L^2(0, T; H_\sigma^{1/2}). \quad (2.36)$$

Noting that, by (2.14), the bilinear form  $u \mapsto B(u)$  is continuous from the space  $L^2(0, T; H_\sigma^{1/2})$  to  $L^1(0, T; H_\sigma^{-1})$ , we see that

$$B(u_{N_j}) \rightarrow B(u) \quad \text{in } L^1(0, T; H_\sigma^{-1}). \quad (2.37)$$

Finally, due to (2.36),

$$u_{N_j}(0) \rightarrow u(0) \quad \text{in } L_\sigma^2. \quad (2.38)$$

Let us take any integer  $m \geq 1$  and apply the projection  $P_m$  to Eq. (2.31) with  $N = N_j \geq m$ . Since  $P_m \circ P_{N_j} = P_m$ , we obtain

$$P_m \dot{u}_{N_j} + P_m L u_{N_j} + P_m B(u_{N_j}) = P_m f.$$

Letting  $N_j \rightarrow \infty$  and using (2.34) – (2.37), we derive

$$P_m \dot{u} + P_m L u + P_m B(u) = P_m f,$$

where the equality holds in the space  $L^1(0, T; H_\sigma^{-1})$ . Since  $m$  is arbitrary, we conclude that

$$\dot{u} + L u + B(u) = f.$$

Furthermore, due to (2.38),

$$u(0) = \lim_{j \rightarrow \infty} u_{N_j}(0) = \lim_{j \rightarrow \infty} P_{N_j} u_0 = u_0.$$

We have thus proved the existence of a solution. It remains to establish inequality (2.24).

To this end, note that inequality (2.24) remains valid for  $u = u_N$ . Since  $\{u_{N_j}\}$  converges weakly to  $u \in \mathcal{H}$ , we see that

$$\|u\|_{L^2(0, t; H^1)} \leq \liminf_{j \rightarrow \infty} \|u_{N_j}\|_{L^2(0, t; H^1)} \quad \text{for any } t \leq T.$$

Combining this with convergence (2.36), we conclude that one can pass to the limit in (2.24) with  $u = u_{N_j}$  as  $j \rightarrow \infty$ . This results in inequality (2.24) for the constructed solution  $u(t, x)$ . The proof of Theorem 2.1.13 is complete.  $\square$

*Exercise 2.1.14.* Show that any solution  $u(t, x)$  of the Navier–Stokes system (2.19) satisfies the relation

$$|u(t)|_2^2 + 2\nu \int_0^t |\nabla u(s)|_2^2 ds = |u(0)|_2^2 + 2 \int_0^t \langle f(s), u(s) \rangle ds, \quad (2.39)$$

which is called the *energy balance*. *Hint:* Relation (2.39) can be obtained formally by taking the scalar product of Eq. (2.19) with  $u$  and integrating in time. To justify this calculation, use relation (2.10).

*Remark 2.1.15.* Our proof shows that any sequence of Galerkin approximations  $u_{N_j}$  contains a subsequence that converges to a solution of the Navier–Stokes system. By the uniqueness, the limiting function must coincide with the solution  $u(t, x)$  constructed in Theorem 2.1.13. Hence, the whole sequence  $\{u_N\}$  converges to  $u$  weakly in  $\mathcal{H}$  and, by Lemma 2.1.5, strongly in  $L^2(0, T; H_\sigma^r)$  with  $r < 1$ :

$$u_N \rightharpoonup u \quad \text{in } \mathcal{H}, \quad u_N \rightarrow u \quad \text{in } L^2(0, T; H_\sigma^r). \quad (2.40)$$

Theorem 2.1.13 on existence and uniqueness of a solution is true for 2D Navier–Stokes equations in a bounded domain with the Dirichlet boundary condition, and the proof of remains essentially the same; see Chapter III in [Tem79].

We conclude this subsection by a remark on solutions with zero mean value. Recall that the mean value  $\langle u \rangle$  of a function  $u \in H^m(\mathbb{T}^2)$  with  $m \geq 0$  is defined in (1.2). In the case of an arbitrary  $m \in \mathbb{R}$ , it can be defined as the zero-order coefficient in the Fourier expansion of  $u$ . Let  $u(t, x)$  be a solution of (2.19). Integrating in time Eq. (2.19) and taking the mean value of both sides, we obtain

$$\langle u(t) \rangle + \int_0^t \langle -\Delta u + B(u) \rangle ds = \langle u(0) \rangle + \int_0^t \langle f(s) \rangle ds.$$

Now note that the integrand on the left-hand side vanishes, and therefore

$$\langle u(t) \rangle = \langle u(0) \rangle + \int_0^t \langle f(s) \rangle ds.$$

So if the mean value of  $f(t)$  is zero for almost all  $t \in [0, T]$ , then  $\langle u(t) \rangle$  is constant. In what follows, we shall study problem (2.19), (2.3) for the case when the mean values of  $f(t)$  and  $u_0$  vanish. What has been said implies that in this case  $\langle u(t) \rangle = 0$  for all  $t \in [0, T]$ .

Let us introduce the function spaces

$$H = \{u \in L_\sigma^2(\mathbb{T}^2, \mathbb{R}^2) : \langle u \rangle = 0\}, \quad V = H^1 \cap H, \quad V^k = H^k \cap H \quad (2.41)$$

and denote by  $V^*$  the dual space of  $V$  with respect to the scalar product in  $L^2$ . Note that  $V^*$  can be regarded as the quotient space  $H^{-1}/W$ , where  $W$  is the space of functionals  $f \in H^{-1}$  vanishing on  $\nabla H^2$ . It follows from Poincaré’s inequality that the usual Sobolev norm on  $V^m$  is equivalent to  $\langle L^m u, u \rangle^{1/2}$ . Slightly abusing the notation, we sometimes denote the latter norm by  $\|\cdot\|_m$ . Finally, for a bounded domain  $Q$ , we write  $H = L_\sigma^2$  and  $V = H_0^1(Q; \mathbb{R}^2) \cap L_\sigma^2$  and denote by  $V^*$  the dual of  $V$ .

When dealing with stochastic Navier–Stokes equations, we shall very often apply Itô’s formula to processes that take values in one space and possess stochastic differentials in a larger space. In this context, the concept of a *Gelfand triple* will play an important role; see Section 7.6. Note that, the spaces  $V \subset H \subset V^*$  form a Gelfand triple, as do the spaces  $V^{k+1} \subset V^k \subset V^{k-1}$  for any integer  $k \geq 1$ .

### 2.1.6 Regularity of solutions

Navier–Stokes system is a parabolic-type equation, and in the 2D case it has good smoothing properties. The proof of this fact is particularly simple for the problem on the torus due to the following lemma.

**Lemma 2.1.16.** *If  $u \in H \cap H^2$ , then  $\langle B(u), \Delta u \rangle = 0$ .*

The proof of this result is based on the existence of a *stream function* for divergence-free vector fields.

*Exercise 2.1.17.* Show that for any vector field  $u \in H$  on the 2D torus there is a function  $\psi \in H^1(\mathbb{T}^2)$  such that  $u = \text{curl } \psi := (-\partial_2 \psi, \partial_1 \psi)$ . Moreover, the function  $\psi$  is unique up to an additive constant, and if  $u \in H^k(\mathbb{T}^2, \mathbb{R}^2)$ , then  $\psi \in H^{k+1}(\mathbb{T}^2)$ . *Hint:* Use the Fourier expansion.

*Proof of Lemma 2.1.16.* By Exercise 2.1.17, there is a function  $\psi \in H^3(\mathbb{T}^2)$  such that  $u = \text{curl } \psi$ . Furthermore, it follows from Theorem 2.1.2 that for any  $u \in H \cap H^2$  there is  $p \in H^1(\mathbb{T}^2)$  such that  $B(u) = \langle u, \nabla \rangle u - \nabla p$ . Thus, setting  $\text{curl } u = \partial_1 u^2 - \partial_2 u^1$  and integrating by parts, we obtain

$$\begin{aligned} \langle B(u), \Delta u \rangle &= \int_{\mathbb{T}^2} (\langle u, \nabla \rangle u - \nabla p) \text{curl}(\Delta \psi) \, dx \\ &= \int_{\mathbb{T}^2} \langle u, \nabla \rangle (\text{curl } u) \Delta \psi \, dx. \end{aligned} \quad (2.42)$$

where we used the relations

$$\text{curl}(\langle u, \nabla \rangle u) = \langle u, \nabla \rangle (\text{curl } u), \quad \text{curl}(\nabla p) = 0.$$

Now note that  $\text{curl } u = \text{curl}(\text{curl } \psi) = \Delta \psi$ . Substituting this into (2.42), integrating by parts, and using that  $\text{div } u = 0$ , we arrive at the required result.  $\square$

We can now state a first regularisation result for 2D Navier–Stokes equations. It says, roughly speaking, that the solution immediately becomes more regular than the initial function.

**Theorem 2.1.18.** *Let  $u_0 \in H$  and  $f \in L^2(0, T; H)$ . Then the solution  $u(t, x)$  of problem (2.19), (2.3) belongs to the space  $C(t_0, T; H^1) \cap L^2(t_0, T; H^2)$  for any  $t_0 > 0$ . Moreover, for  $0 \leq t \leq T$ , we have*

$$t \|u(t)\|_1^2 + \nu \int_0^t s \|u(s)\|_2^2 \, ds \leq |u_0|_2^2 + \nu^{-1} \int_0^t s |f(s)|_2^2 \, ds + \int_0^t \|f(s)\|_{-1}^2 \, ds. \quad (2.43)$$

*Proof.* We shall confine ourselves to a formal derivation of the a priori estimate (2.43) for smooth solutions. Its justification can be carried out with the help of Galerkin approximation; cf. proof of (2.24). Once (2.43) is established, we conclude immediately that  $u \in L^\infty(t_0, T; H^1) \cap L^2(t_0, T; H^2)$  for any  $t_0 > 0$ . Combining this with Eq. (2.19), we see that  $\dot{u} \in L^2(t_0, T; H)$ , whence it follows that  $u \in C(t_0, T; H^1)$ .

To prove (2.43), consider the functional  $\varphi(t, u) = t \langle Lu, u \rangle$ . Differentiating it with respect to time and using Eq. (2.19) and Lemma 2.1.16, we derive

$$\begin{aligned} \partial_t \varphi(t, u) &= \|u\|_1^2 + 2t \langle Lu, \dot{u} \rangle \\ &= \|u\|_1^2 + 2t \langle Lu, -\nu Lu - B(u) + \Pi f \rangle \\ &= \|u\|_1^2 - 2\nu t |Lu|_2^2 + 2t \langle Lu, f \rangle. \end{aligned}$$

Using the inequality  $2t \langle Lu, f \rangle \leq \nu t |Lu|_2^2 + \nu^{-1} t |f|_2^2$ , we obtain

$$\partial_t (t \|u\|_1^2) \leq \|u\|_1^2 - \nu t \|u\|_2^2 + \nu^{-1} t |f|_2^2.$$

Integrating from 0 to  $t$ , we get

$$t \|u(t)\|_1^2 + \nu \int_0^t s \|u(s)\|_2^2 ds \leq \nu^{-1} \int_0^t s |f(s)|_2^2 ds + \int_0^t \|u(s)\|_1^2 ds.$$

Using (2.24), we see that  $u$  satisfies (2.43).  $\square$

An analogue of Theorem 2.1.18 is true for the Navier–Stokes system in a bounded domain. However, in this case the proof becomes more complicated, because Lemma 2.1.16 is no longer true, and the term  $\langle B(u), Lu \rangle$  does not vanish. The corresponding argument is similar to that used below to prove the higher regularity of solutions. For simplicity, we assume again that  $\nu = 1$ .

**Theorem 2.1.19.** *Let  $u_0 \in H$ , and let  $f \in L^2(0, T; H^{m-1})$  for some integer  $m \geq 2$ . Then a solution  $u(t, x)$  of problem (2.19), (2.3) belongs to the space  $C(t_0, T; H^m) \cap L^2(t_0, T; H^{m+1})$  for any  $t_0 > 0$ . Moreover, there is a constant  $C_m > 0$  such that the following inequality holds for  $0 \leq t \leq T$ :*

$$\begin{aligned} t^m \|u(t)\|_m^2 + \int_0^t s^m \|u\|_{m+1}^2 ds &\leq \int_0^t s^m \|f(s)\|_{m-1}^2 ds \\ &+ C_m (|u_0|_2^2 + |u_0|_2^{4m+2} + \|f\|_{L^2(0, T; H)}^2 + \|f\|_{L^2(0, T; H)}^{4m+2}). \end{aligned} \quad (2.44)$$

*Proof.* As in the proof of the previous theorem, we shall confine ourselves to the formal derivation of (2.44). We argue by induction. Note that, in view of Theorem 2.1.18, inequality (2.44) holds for  $m = 1$ . We now assume that  $m \geq 2$  and set  $\varphi_m(t, u) = t^m \|u\|_m^2 = t^m \langle L^m u, u \rangle$ . Then, for a smooth solution  $u$ , we have

$$\begin{aligned} \partial_t \varphi_m(t, u) &= m t^{m-1} \|u\|_m^2 + 2t^m \langle L^m u, \dot{u} \rangle \\ &= m t^{m-1} \|u\|_m^2 - 2\nu t^m \|u\|_{m+1}^2 - 2t^m \langle L^m u, B(u) - f \rangle. \end{aligned} \quad (2.45)$$

To estimate the nonlinear term, we need the following proposition, whose proof is given at the end of this subsection.

**Lemma 2.1.20.** *For any integer  $m \geq 2$  there is a constant  $C_m > 0$  such that*

$$|\langle L^m u, B(u) \rangle| \leq C_m \|u\|_{\frac{4m-1}{m+1}}^{\frac{4m-1}{2m}} \|u\|_1^{\frac{m+1}{2m}} |u|_2^{1/2} \quad \text{for } u \in H \cap H^{m+1}. \quad (2.46)$$

Substituting (2.46) into (2.45) and carrying out some simple transformations, we derive

$$\partial_t(t^m \|u\|_m^2) + t^m \|u\|_{m+1}^2 \leq m t^{m-1} \|u\|_m^2 + C_1 t^m \|u\|_1^{2(m+1)} |u|_2^{2m} + t^m \|f\|_{m-1}^2. \quad (2.47)$$

By Theorems 2.1.13 and 2.1.18 with  $\nu = 1$ , we have

$$t \|u\|_1^2 |u|_2^2 \leq C (|u_0|_2^2 + \|f\|_{L^2(0,T,H)}^2)^2.$$

Substituting this into (2.47) and integrating in time, we obtain

$$\begin{aligned} t^m \|u\|_m^2 + \int_0^t s^m \|u(s)\|_{m+1}^2 ds &\leq m \int_0^t s^{m-1} \|u(s)\|_m^2 ds + \int_0^t s^m \|f\|_{m-1}^2 ds \\ &\quad + C_2 (|u_0|_2^2 + \|f\|_{L^2(0,T,H)}^2)^{2m} \int_0^t \|u(s)\|_1^2 ds. \end{aligned} \quad (2.48)$$

Using now inequality (2.24) and the induction hypothesis to estimate the first and third integrals on the right-hand side of (2.48), we arrive at the required estimate (2.44).  $\square$

*Proof of Lemma 2.1.20.* We first note that

$$\langle B(u), L^m u \rangle = \sum_{|\alpha|=m} C_\alpha \langle D^\alpha B(u), D^\alpha u \rangle, \quad (2.49)$$

where  $C_\alpha$  are some constants. Each term under the sum can be written as

$$\langle D^\alpha B(u), D^\alpha u \rangle = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \langle B(D^{\alpha-\beta} u, D^\beta u), D^\alpha u \rangle.$$

Since  $\langle B(u), D^\alpha u \rangle, D^\alpha u = 0$ , the term with  $\beta = \alpha$  vanishes, and using (2.13), we get

$$\begin{aligned} |\langle D^\alpha B(u), D^\alpha u \rangle| &\leq C \sum_{\beta < \alpha} \|D^{\alpha-\beta} u\|_{1/2} \|D^\beta u\|_1 \|D^\alpha u\|_{1/2} \\ &\leq C \sum_{\beta < \alpha} \|u\|_{1/2+m-|\beta|} \|u\|_{1+|\beta|} \|u\|_{m+\frac{1}{2}}, \end{aligned} \quad (2.50)$$

where  $m = |\alpha|$  and  $0 \leq |\beta| < m$ . Note that the numbers  $\frac{1}{2} + m - |\beta|$ ,  $1 + |\beta|$  and  $m + \frac{1}{2}$  lie between 1 and  $m + \frac{1}{2}$ . By the interpolation inequality (1.7), for any  $a \in [1, m + \frac{1}{2}]$ , we have

$$\|u\|_a \leq |u|_2^{1-\frac{a}{m+\frac{1}{2}}} \|u\|_{m+\frac{1}{2}}^{\frac{a}{m+\frac{1}{2}}} := X, \quad \|u\|_a \leq \|u\|_1^{1-\frac{a-1}{m}} \|u\|_{m+\frac{1}{2}}^{\frac{a-1}{m}} := Y.$$

Take any term in the sum in the right-hand side of (2.50). Estimating each of its factors by  $X^\theta Y^{1-\theta}$  with a suitable  $\theta$ , we obtain

$$|\langle D^\alpha B(u), D^\alpha u \rangle| \leq C \|u\|_{m+\frac{1}{2}}^{\frac{4m-1}{2m}} \|u\|_1^{\frac{m+1}{2m}} |u|_2^{1/2}.$$

Substituting this into (2.49), we arrive at (2.46).  $\square$

Theorem 2.1.19 implies, in particular, that if the right-hand side is infinitely smooth, then so is the solution of (2.19) for  $t > 0$ . It can also be shown that the space-time analyticity of  $f$  implies a similar property for the solution. We refer the reader to [FT89, DG95] for an exact formulation and proof of the corresponding result.

An analogue of Theorem 2.1.19 is true for bounded domains with smooth boundary. In this case, to prove the regularity of  $u$  we have to assume that the right-hand side is smooth both in space and time, and the initial function  $u_0$  satisfies some compatibility condition. These results are not needed for studying ergodic properties of the Navier–Stokes flow, and therefore we do not give any further details, referring the reader to [Tay97] for a systematic study of regularity of solutions to nonlinear PDE’s.

### 2.1.7 Navier–Stokes process

Let us consider the Navier–Stokes system (2.19) in which  $f \in L^2_{\text{loc}}(\mathbb{R}_+, H^{-1})$  and  $\nu = 1$ . By Theorem 2.1.13, for any  $\tau \in \mathbb{R}$  and any  $u_0 \in H$  Eq. (2.19) has a unique solution  $u \in C(\mathbb{R}_\tau, H) \cap L^2_{\text{loc}}(\mathbb{R}_\tau, V)$  such that  $u(\tau) = u_0$ , where  $\mathbb{R}_\tau = [\tau, \infty)$ . We introduce the resolving operator  $S_{t,\tau} : H \rightarrow H$  by the relation  $S_{t,\tau}(u_0) = u(t)$ . It is straightforward to verify that the family  $\{S_{t,\tau}, t \geq \tau \geq 0\}$  form a *process*:

$$S_{\tau,\tau} = \text{Id}_H, \quad S_{t,\tau} = S_{t,s} \circ S_{s,\tau} \quad \text{for any } t \geq s \geq \tau \geq 0,$$

where  $\text{Id}_H$  stands for the identity mapping in  $H$ . We shall call  $S_{t,\tau}$  the *Navier–Stokes process* or, for short, the *NS process*. Let us set  $S_t = S_{t,0}$ . The following proposition establishes a dissipativity property of the NS process.

**Proposition 2.1.21.** (i) *Let  $u_0 \in H$  and  $f \in L^2_{\text{loc}}(\mathbb{R}_+, H^{-1})$ . Then*

$$|S_t(u_0)|_2^2 \leq e^{-\alpha_1 t} |u_0|_2^2 + \int_0^t e^{-\alpha_1(t-s)} \|f(s)\|_{-1}^2 ds, \quad (2.51)$$

where  $\alpha_1 > 0$  stands for the first positive eigenvalue<sup>1</sup> of the Laplacian.

(ii) *Let  $u_0 \in V$  and  $f \in L^2_{\text{loc}}(\mathbb{R}_+, H)$ . Then*

$$\|S_t(u_0)\|_1^2 \leq e^{-\alpha_1 t} \|u_0\|_1^2 + \int_0^t e^{-\alpha_1(t-s)} |f(s)|_2^2 ds. \quad (2.52)$$

*Proof.* In view of (2.24) with  $\nu = 1$ , we have

$$|S_t(u_0)|_2^2 + \int_0^t |\nabla S_s(u_0)|_2^2 ds \leq |u_0|_2^2 + \int_0^t \|f(s)\|_{-1}^2 ds.$$

Using Poincaré’s inequality  $|\nabla u|_2^2 \geq \alpha_1 |u|_2^2$  to minorise the integrand in the left-hand side and applying the Gronwall inequality, we arrive at (2.51). The proof of (2.52) can be carried out by a similar argument, using Lemma 2.1.16; cf. the proof of Theorem 2.1.18.  $\square$

<sup>1</sup>In the case of the standard torus  $\mathbb{T}^2 = \mathbb{R}^2/2\pi\mathbb{Z}^2$ , we have  $\alpha_1 = 1$ .

*Exercise 2.1.22.* Prove inequality (2.52).

*Exercise 2.1.23.* Show that any solution of the Navier–Stokes system satisfies the inequality

$$|u(0)|_2^2 \leq C (t^{-1} \vee 1) \int_0^t (\|S_s(u(0))\|_1^2 + \|f(s)\|_{-1}^2) ds, \quad (2.53)$$

where  $C > 0$  is a constant not depending on  $t$  and  $u_0$ . *Hint:* Use (2.51) to estimate from above the left-hand side of (2.39).

Note that inequality (2.51) remains true for the Navier–Stokes system in a bounded domain, and the proof is literally the same as in the case of a torus. On the other hand, inequality (2.52) is no longer valid, because Lemma 2.1.16 does not hold in general. Still, the NS process is bounded from  $H$  to  $V$ , and some explicit bounds for  $\|S_t(u)\|_1$  can be obtained by repeating the arguments used in the proof of Theorem 2.1.18.

**Corollary 2.1.24.** *Suppose that  $f \equiv 0$ . Then the function  $u = 0$  is the only fixed point of the NS semigroup, and it is globally exponentially stable in  $H$ . Moreover, in the case of a torus, it is also globally exponentially stable in  $V$ .*

We now turn to continuity properties of the NS process. We shall write  $S_t(u_0, f)$  to indicate the dependence of the resolving operator on the right-hand side of the NS equation. The following result shows that  $S_t$  is uniformly Lipschitz continuous on bounded subsets of  $H$ .

**Proposition 2.1.25.** (i) *There exists a constant  $C > 0$  such that, for any functions  $u_{01}, u_{02} \in H$  and  $f_1, f_2 \in L_{\text{loc}}^2(\mathbb{R}_+, H^{-1})$ , we have*

$$\begin{aligned} |S_t(u_{01}, f_1) - S_t(u_{02}, f_2)|_2^2 &\leq |u_{01} - u_{02}|_2^2 \exp\left(C \int_0^t \|S_r(u_{01}, f_1)\|_1^2 dr\right) \\ &\quad + \int_0^t \exp\left(C \int_s^t \|S_r(u_{01}, f_1)\|_1^2 dr\right) \|f_1(s) - f_2(s)\|_{-1}^2 ds. \end{aligned} \quad (2.54)$$

(ii) *There exists a constant  $C > 0$  such that, for  $0 < t \leq 1$  and any functions  $u_{01}, u_{02} \in H$  and  $f_1, f_2 \in L_{\text{loc}}^2(\mathbb{R}_+, H)$ , we have*

$$\begin{aligned} \|S_t(u_{01}, f_1) - S_t(u_{02}, f_2)\|_1^2 &\leq C \int_0^t |f_1 - f_2|_2^2 ds \\ &\quad + A(t) t^{-3} \left( |u_{01} - u_{02}|_2^2 + \int_0^t \|f_1 - f_2\|_{-1}^2 ds \right), \end{aligned} \quad (2.55)$$

where we set

$$A(t) = \exp\left(C \int_0^t (\|S_r(u_{01}, f_1)\|_1^2 + \|S_r(u_{02}, f_2)\|_1^2 + |f_1|_2^2 + |f_2|_2^2) ds\right).$$

The proof of this result is rather standard and uses a well-known idea. Namely, one should take the difference between the equations corresponding to two solutions, multiply (in the sense of scalar product in  $H$ ) the resulting relation by an appropriate function, estimate the nonlinear terms with the help of Sobolev embedding and interpolation inequalities, and finally apply Gronwall's lemma. The realisation of this scheme is somewhat technical, and to keep the presentation on an elementary level, we postpone the proof of Proposition 2.1.25 until Section 2.6. The reader not interested in those technical details may safely skip it.

An analogue of Proposition 2.1.25 holds for the Navier–Stokes system in a bounded domain. Namely, in that case, inequality (2.54) and its proof remain true without any changes. As for (2.55), its proof used inequality (2.43), which does not hold for a general bounded domain. One can show, however, that the quantity  $\|S_t(u_{01}, f_1) - S_t(u_{02}, f_2)\|_{1/2}^2$  is bounded by the right-hand side of (2.55) in which  $t^{-3}$  is replaced by  $t^{-5/2}$ .

*Exercise 2.1.26.* Prove the above claims for the Navier–Stokes system in a bounded domain with the Dirichlet boundary condition.

The following exercise establishes the continuity of the resolving operator for the Navier–Stokes system in different norms. This will be important in Section 4.3.

*Exercise 2.1.27. (i)* Prove that, for any functions  $u_0 \in H$ ,  $h \in L^2(0, T; H)$ , and  $g \in L^\infty(0, T; V)$ , Eq. (2.19) with the right-hand side  $f = h + \partial_t g$  has a unique solution  $u \in \mathcal{X}_T := C(0, T; H) \cap L^2(0, T; V)$  issued from  $u_0$ .

*(ii)* Let  $u_1, u_2 \in \mathcal{X}_T$  be two solutions of Eq. (2.19) with  $f = \partial_t g_i$ ,  $i = 1, 2$ , where  $g_i \in L^\infty(0, T; V)$ . Then for any  $R > 0$  there is a constant  $C_R > 0$  such that if  $\|u_i(0)\|_2 \leq R$  and  $\|g_i\|_{L^\infty(0, T; V)} \leq R$ , then

$$\|u_1 - u_2\|_{\mathcal{X}_T} \leq C_R (\|u_1(0) - u_2(0)\|_2 + \|g_1 - g_2\|_{L^\infty(0, T; V)}).$$

### 2.1.8 Foias–Prodi estimates

The Foias–Prodi inequalities enable one to establish the exponential convergence of two solutions for the Navier–Stokes system, on condition that one has a good control over their low Fourier modes. We shall need two versions of these estimates that correspond to discrete- and continuous-time perturbations.

Let us fix an integer  $N \geq 1$  and denote by  $H_{(N)}$  the vector span<sup>2</sup> of the basis functions  $e_s$  with  $|s| \leq N$ ; see (2.29). We write  $P_N$  for the orthogonal projection in  $L^2$  onto  $H_{(N)}$ . Along with the Navier–Stokes system (2.19), let us consider the equation

$$\dot{u} + Lu + B(u) + \lambda P_N(u - u'(t)) = \Pi f(t), \quad (2.56)$$

<sup>2</sup>This space is slightly different from the one defined by (2.30), because we exclude the constant functions (1, 0) and (0, 1).



where  $u'(t)$  is a given function and  $\lambda > 0$  is a large parameter. Equation (2.56) is a Navier–Stokes-type system, and the arguments in the proof of Theorem 2.1.13 can be used to show that the Cauchy problem for (2.56) has a unique solution  $u \in C(\mathbb{R}_+, H) \cap L^2_{\text{loc}}(\mathbb{R}_+, V)$  for any  $u' \in L^2_{\text{loc}}(\mathbb{R}_+, H)$ .

**Theorem 2.1.28.** *Let  $f \in L^2_{\text{loc}}(\mathbb{R}_+, H^{-1})$  and let  $u' \in C(\mathbb{R}_+, H) \cap L^2_{\text{loc}}(\mathbb{R}_+, V)$  be a solution of the Navier–Stokes system (2.19) such that*

$$\int_s^t \|u'(r)\|_1^2 dr \leq \rho + K(t-s) \quad \text{for } s \leq t \leq s+T, \quad (2.57)$$

where  $s, T, \rho$ , and  $K$  are non-negative constants. Then for any  $M > 0$  there are constants  $N \geq 1$  and  $\lambda > 0$  depending on  $M$  and  $\rho$  such that for any solution  $u \in C(\mathbb{R}_+, H) \cap L^2_{\text{loc}}(\mathbb{R}_+, V)$  of (2.56) we have

$$|u(t) - u'(t)|_2 \leq e^{-M(t-s) + C\rho} |u(s) - u'(s)|_2 \quad \text{for } s \leq t \leq s+T, \quad (2.58)$$

where  $C > 0$  is an absolute constant.

*Proof.* We confine ourselves to the formal derivation of (2.58). The difference  $w = u - u'$  satisfies the equation

$$\dot{w} + Lw + B(u) - B(u') + \lambda \mathbf{P}_N w = 0.$$

Taking the scalar product in  $L^2$  of this equation with  $2w$  and using (2.11) and (2.13), we derive

$$\begin{aligned} \partial_t |w|_2^2 + 2|\nabla w|_2^2 + 2\lambda |\mathbf{P}_N w|_2^2 &\leq C_1 \|w\|_1 |w|_2 \|u'\|_1 \\ &\leq |\nabla w|_2^2 + C_2 \|u'\|_1^2 |w|_2^2. \end{aligned} \quad (2.59)$$

Now note that, by Poincaré’s inequality, we have

$$|\nabla w|_2^2 \geq N^2 |(I - \mathbf{P}_N)w|_2^2.$$

Substituting this into (2.59), we obtain

$$\partial_t |w|_2^2 + (\lambda_1 - C_2 \|u'\|_1^2) |w|_2^2 \leq 0,$$

where  $\lambda_1 = \min\{N^2, 2\lambda\}$ . Application of the Gronwall inequality results in

$$\begin{aligned} |w(t)|_2^2 &\leq |w(s)|_2^2 \exp\left(-\lambda_1(t-s) + C_2 \int_s^t \|u'(r)\|_1^2 dr\right) \\ &\leq |w(s)|_2^2 \exp(C_2 \rho - (\lambda_1 - C_2 K)(t-s)). \end{aligned}$$

Choosing  $\lambda$  and  $N$  so large that  $\lambda_1 \geq C_2 K + 2M$ , we arrive at (2.58) with  $C = C_2/2$ .  $\square$

*Remark 2.1.29.* The above proof does not use the fact that the space variables belong to the torus, and the result remains true for a bounded domain. In this case,  $\mathbf{P}_N$  denotes the orthogonal projection in  $L^2$  to the subspace spanned by the first  $N$  eigenfunctions of Stokes operator  $L = -\Pi\Delta$ . Furthermore, the regularity in time of the function  $f$  was not used either. In particular, if the right-hand side has the form  $f = h + \partial_t\zeta$ , where  $h \in L^2_{\text{loc}}(\mathbb{R}_+, H)$  and  $\zeta \in C(\mathbb{R}_+, H)$ , then the conclusion of Theorem 2.1.28 remains valid; cf. Exercise 2.1.27 for the existence of a solution. This observation will be important for the case in which the right-hand side is the sum of a deterministic function and the time derivative of a Brownian motion.

We now turn to the discrete version of the Foias–Prodi estimate. To simplify the presentation, we shall assume that the right-hand side  $f$  is independent of time and belongs to the space  $H$ . In this case, the resolving operator  $S_{t,\tau}$  of the Navier–Stokes equation (2.19) depends only on the difference  $t - \tau$ , and we shall write  $S_{t-\tau}$ . Let us fix a constant  $T > 0$  and consider four sequences  $u_k, u'_k, \zeta_k, \zeta'_k \in H$  satisfying the relations

$$u_k = S(u_{k-1}) + \zeta_k, \quad u'_k = S(u'_{k-1}) + \zeta'_k, \quad k \geq 1, \quad (2.60)$$

where  $S = S_T$ .

**Theorem 2.1.30.** *Suppose that*

$$\mathbf{P}_N u_k = \mathbf{P}_N u'_k, \quad (I - \mathbf{P}_N)\zeta_k = (I - \mathbf{P}_N)\zeta'_k \quad \text{for } l+1 \leq k \leq m, \quad (2.61)$$

where  $m > l \geq 0$  are some integers. Then there is an absolute constant  $C > 0$  such that

$$|u_k - u'_k|_2 \leq (CN^{-1})^{k-l} \exp\left(C(k-l)(\langle \|\mathbf{u}\|_1^2 \rangle_l^k + \langle \|\mathbf{u}'\|_1^2 \rangle_l^k + 1)\right) |u_l - u'_l|_2, \quad (2.62)$$

where  $l \leq k \leq m$ , and for a sequence  $\{v_j\} \subset H$  we set

$$\langle \|\mathbf{v}\|_1^2 \rangle_l^k = \frac{1}{k-l} \sum_{j=l}^{k-1} \int_0^T \|S_t(v_j)\|_1^2 dt.$$

*Proof.* Let us set  $\Delta_k = |u_k - u'_k|_2$ . It follows from (2.55), (2.61), and the Poincaré inequality that

$$\begin{aligned} \Delta_k &= |(I - \mathbf{P}_N)(u_k - u'_k)|_2 = |(I - \mathbf{P}_N)(S(u_{k-1}) - S(u'_{k-1}))|_2 \\ &\leq C_1 N^{-1} \|S(u_{k-1}) - S(u'_{k-1})\|_1 \\ &\leq C_2 N^{-1} \Delta_{k-1} \exp\left(C_2 \int_0^T (\|S_t(u_{k-1})\|_1^2 + \|S_t(u'_{k-1})\|_1^2 + |f|_2^2) dt\right). \end{aligned}$$

Iteration of this inequality results in (2.62).  $\square$

*Remark 2.1.31.* As in the case of continuous-time perturbations, the conclusion of Theorem 2.1.30 remains true for a bounded domain. The only difference is that the constant  $N^{-1}$  in (2.62) should be replaced by  $\alpha_N^{-1}$ , where  $\alpha_j$  stands for the  $j^{\text{th}}$  eigenvalue of the Stokes operator  $L$ .

### 2.1.9 Some hydrodynamical terminology

As we mentioned in the Introduction, the behaviour of incompressible fluid occupying a domain  $Q \subset \mathbb{R}^d$  (or the  $d$ -dimensional torus  $\mathbb{T}^d$ ), with  $d = 2$  or  $3$ , is described by the  $d + 1$  equations (2.1), supplemented by suitable boundary conditions if necessary. In those equations,  $u(t, x)$ ,  $p(t, x)$ , and  $\nu$  denote, respectively, the velocity, pressure, and the kinematic viscosity of the fluid. The quantity  $\frac{1}{2}|u(t)|_2^2 = \frac{1}{2} \int_Q |u(t, x)|^2 dx$  is called the *energy of the fluid* (at time  $t$ ).

Let us assume that  $f = 0$  and that the fluid satisfies the no-slip condition, i.e.,  $u = 0$  on  $\partial Q$ . Multiplying the first equation in (2.1) by  $u(t)$  in  $L_2$ , we obtain

$$\frac{1}{2} \frac{d}{dt} |u(t)|_2^2 + \nu \int_Q |\nabla u|^2 dx = 0.$$

Accordingly, the quantity  $\varepsilon = \nu \int_Q |\nabla u|^2 dx$  is called the *rate of dissipation of energy*.

The *Reynolds number* of the flow is defined as

$$R = \frac{\langle \text{characteristic scale for } x \rangle \cdot \langle \text{characteristic scale for } u \rangle}{\nu}.$$

The terms in the numerator are ambiguous, and for the purposes of this book, we understand them as follows:

$$\begin{aligned} \langle \text{characteristic scale for } x \rangle &= \text{diameter of } Q \text{ (or the period of torus),} \\ \langle \text{characteristic scale for } u \rangle &= \frac{1}{\sqrt{2}} |u|_2 =: \sqrt{E}. \end{aligned}$$

If  $u$  depends on a random parameter, then we modify this definition:

$$\langle \text{characteristic scale for } u \rangle = (\mathbb{E} E)^{1/2}.$$

In the 2D case, we denote by  $v = \text{curl } u = \partial_1 u^2 - \partial_2 u^1$  the *vorticity* of the flow, and call  $\Omega = \frac{1}{2} \int v^2 dx$  the *enstrophy*.

*Exercise 2.1.32.* Show that  $\Omega = \frac{1}{2} \int_Q |\nabla u|^2 dx$ .

Applying the operator  $\text{curl}$  to the first equation in (2.1) and using the relation  $\text{curl}(\langle u, \nabla \rangle u) = \langle u, \nabla \rangle v$ , we derive

$$\dot{v} - \nu \Delta v + \langle u, \nabla \rangle v = \text{curl } f. \quad (2.63)$$

Let us assume that  $f = 0$  and consider the case of a torus. Multiplying (2.63) by  $v$  in  $L_2$ , we obtain

$$\frac{1}{2} \frac{d}{dt} |v|_2^2 + \nu \int_{\mathbb{T}^2} |\nabla v|^2 dx = 0.$$

We deduce that  $\nu \int_{\mathbb{T}^2} |\nabla v|^2 dx$  is the *rate of dissipation of enstrophy*.

*Exercise 2.1.33.* Show that, in the case of a torus, we have  $\int_{\mathbb{T}^2} |\nabla v|^2 dx = |\Delta u|_2^2$ .

## 2.2 Stochastic Navier–Stokes equations

We now turn to the Navier–Stokes system with a random right-hand side. More precisely, we shall consider Eqs. (2.1) in which  $f$  has the form

$$f(t, x) = h(x) + \eta(t, x), \quad (2.64)$$

where  $h$  is a deterministic function and  $\eta$  is a stochastic process. For the latter, we study the following three cases that give rise to a Markov process in the phase space: random kick force, spatially regular white noise, and the time derivative of a compound Poisson process. Each of these three types of forces has its own advantages: the first and third ones are more realistic in the sense that they allow non-Gaussian perturbations, while the second can be considered as a limiting model for a large class of stationary processes whose time integral does not necessarily satisfy the condition of independence of increments; cf. Theorem 2.2.2 below and Théorème 4.3 in [Rio00].

Let us describe the above-mentioned models in more detail. We shall say that a stochastic process  $\eta$  is a *random kick force on  $\mathbb{T}^2$*  if it has the form (see Figure 2.1)

$$\eta(t, x) = \sum_{k=1}^{\infty} \eta_k(x) \delta(t - kT), \quad (2.65)$$

where  $T > 0$  is a constant and  $\{\eta_k\}$  is a sequence of i.i.d. random variables in  $H$ .

We shall say that a stochastic process  $\eta$  is a *spatially regular white noise* if

$$\eta(t, x) = \frac{\partial}{\partial t} \zeta(t, x), \quad \zeta(t, x) = \sum_{j=1}^{\infty} b_j \beta_j(t) e_j(x), \quad t \geq 0, \quad (2.66)$$

where  $\{e_j\}$  is an orthonormal basis in  $H$ ,  $b_j \geq 0$  are some constants such that

$$\mathfrak{B} := \sum_{j=1}^{\infty} b_j^2 < \infty, \quad (2.67)$$

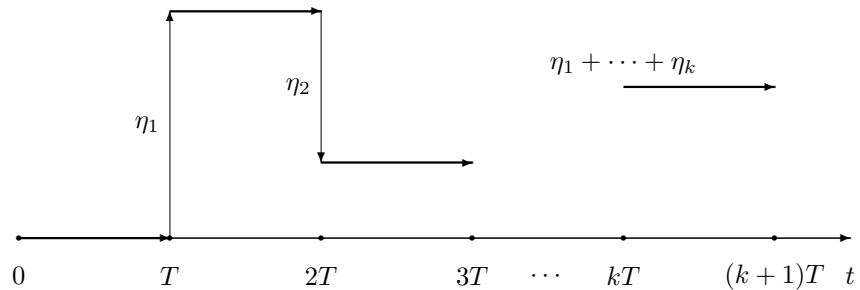


Figure 2.1: Time integral of a random kick force

and  $\{\beta_j\}$  is a sequence of independent standard Brownian motions. The following result is a simple consequence of Doob’s moment inequality (7.56) and Levy’s equivalence theorem<sup>3</sup> (see Theorem 9.7.1 in [Dud02]).

*Exercise 2.2.1.* Show that for any  $T > 0$  the series in (2.66) converges almost surely in the space  $C(0, T; H)$ . *Hint:* One can apply the argument used in the proof of Theorem 2.2.2 below to prove the convergence of the series in probability; see the solution of Exercise 2.4.17.

Finally, we shall say that  $\eta$  is the *time derivative of a compound Poisson process* if (see Figure 2.2)

$$\eta(t, x) = \sum_{k=1}^{\infty} \eta_k(x) \delta(t - t_k). \quad (2.68)$$

Here  $\{\eta_k\}$  are i.i.d. random variables in  $H$ ,  $\{t_k\}$  is a sequence independent of  $\{\eta_k\}$  such that<sup>4</sup>  $\tau_k = t_k - t_{k-1}$  are i.i.d. random variables with an exponential distribution; see Example 1.3.10 in [App04].

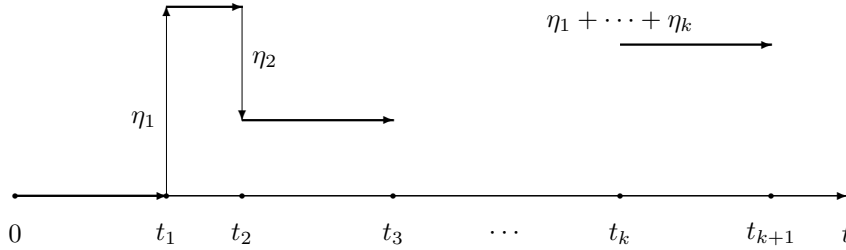


Figure 2.2: Compound Poisson process

We shall mostly study the first two types of noise, confining ourselves to a brief discussion of the results for the third class.

The concept of a solution for the Navier–Stokes system with random kicks and spatially regular white noise is defined in the next two subsections. Here we recall a result due to Donsker which shows the *universality of white noise*. Namely, we shall prove that an appropriately normalised high-frequency random kick force is close to a spatially regular white noise.

Let  $t_0 > 0$  be a constant, let  $X$  be a separable Banach space, and let  $D(0, t_0; X)$  be the space of functions that are right-continuous and have left-hand limits at any point of the interval  $[0, t_0]$ . It is a well-known fact that the space  $D(0, t_0; X)$  is complete with respect to the modified Skorokhod metric, which is defined in the following way (see [Bil99, Section 12]). Let  $\Lambda$  be the space of all strictly increasing continuous functions  $\lambda : [0, t_0] \rightarrow [0, t_0]$  such that  $\lambda(0) = 0$  and  $\lambda(t_0) = t_0$ . If  $\zeta_1, \zeta_2 \in D(0, t_0; X)$ , then the modified Skorokhod

<sup>3</sup>In [Dud02], this result is proved for real-valued random variables. However, analysing the proof, it is easily seen that if a series of independent random variables with range in a separable Banach space converges in probability, then it converges also almost surely.

<sup>4</sup>By definition, we set  $t_0 = 0$ .

distance is defined as the least upper bound of those  $\delta > 0$  for which there is  $\lambda \in \Lambda$  such that

$$\sup_{0 \leq s < t \leq t_0} \ln \left| \frac{\lambda(t) - \lambda(s)}{t - s} \right| \leq \delta, \quad \sup_{0 \leq t \leq t_0} \|\zeta_1 - \zeta_2 \circ \lambda\|_X \leq \delta.$$

In the case  $X = \mathbb{R}$ , we shall write  $D(0, t_0)$  instead of  $D(0, t_0; \mathbb{R})$ .

Let us consider a family of random kick forces depending on a parameter  $\varepsilon > 0$  and defined by

$$\eta_\varepsilon(t, x) = \sqrt{\varepsilon} \sum_{k=1}^{\infty} \eta_k(x) \delta(t - k\varepsilon). \quad (2.69)$$

Assume that  $\{\eta_k\}$  is a sequence of i.i.d. random variables of the form

$$\eta_k(x) = \sum_{j=1}^{\infty} b_j \xi_{jk} e_j(x), \quad (2.70)$$

where  $\{e_j\}$  and  $\{b_j\}$  are the same as in (2.66), and  $\{\xi_{jk}\}$  is a family of independent scalar random variables whose laws do not depend on  $k$  and satisfy the relations

$$\mathbb{E} \xi_{jk} = 0, \quad \mathbb{E} \xi_{jk}^2 = 1.$$

Let us define the time integrals of processes (2.69) and (2.66):

$$\begin{aligned} \zeta_\varepsilon(t, x) &= \sqrt{\varepsilon} \sum_{k=1}^{\infty} \eta_k(x) \theta(t - k\varepsilon) = \sqrt{\varepsilon} \sum_{j=1}^{\infty} b_j e_j(x) \sum_{k=1}^{\infty} \xi_{jk} \theta(t - k\varepsilon), \\ \zeta(t, x) &= \sum_{j=1}^{\infty} b_j \beta_j(t) e_j(x). \end{aligned} \quad (2.71)$$

Here  $\theta$  is the Heaviside function, that is,  $\theta(t) = 1$  for  $t \geq 0$  and  $\theta(t) = 0$  for  $t < 0$ . It is clear that, for any  $t_0 > 0$ , almost every trajectory of  $\zeta_\varepsilon$  and  $\zeta$  belongs to  $D(0, t_0; H)$ .

**Theorem 2.2.2.** *For any  $t_0 > 0$ , the family  $\mathcal{D}(\zeta_\varepsilon)$  converges to  $\mathcal{D}(\zeta)$  weakly in the space  $\mathcal{P}(D(0, t_0; H))$ .*

*Proof.* Let us write  $\zeta_\varepsilon$  as

$$\zeta_\varepsilon(t, x) = \sum_{j=1}^{\infty} b_j \beta_j^\varepsilon(t) e_j(x), \quad \beta_j^\varepsilon(t) = \sqrt{\varepsilon} \sum_{k=1}^{\infty} \xi_{jk} \theta(t - k\varepsilon).$$

Donsker's theorem implies that, for any  $t_0 > 0$ ,

$$\mathcal{D}(\beta_j^\varepsilon) \rightarrow \mathcal{D}(\beta_j) \quad \text{in } \mathcal{P}(D(0, t_0)) \text{ as } \varepsilon \rightarrow 0^+;$$

see Section 14 of [Bil99]. It follows that, for each  $n \geq 1$ ,

$$\mathcal{D}(\zeta_\varepsilon^n) \rightarrow \mathcal{D}(\zeta^n) \quad \text{in } \mathcal{P}(D(0, t_0; H)) \text{ as } \varepsilon \rightarrow 0^+, \quad (2.72)$$

where we set

$$\zeta_\varepsilon^n(t, x) = \sum_{j=1}^n b_j \beta_j^\varepsilon(t) e_j(x),$$

and  $\zeta^n(t, x)$  is defined in a similar way. In view of Doob's moment inequality (7.56), we have

$$\mathbb{E} \sup_{0 \leq t \leq t_0} |\zeta^n(t) - \zeta(t)|_2^2 \leq 4 \mathbb{E} |\zeta^n(t_0) - \zeta(t_0)|_2^2 \leq 4 \sum_{j=n+1}^{\infty} b_j^2 \rightarrow 0 \quad (2.73)$$

as  $n \rightarrow \infty$ , and a similar estimate holds for  $\zeta_\varepsilon^n$  and  $\zeta_\varepsilon$ . Combining (2.72) and (2.73) and using Exercise 1.2.19, we arrive at the required result.  $\square$

In Chapter 4, we shall need another version of Theorem 2.2.2 that concerns convergence to  $\zeta$  in the weak topology of the space of probability measures on  $C(0, t_0; H)$ . Namely, let us denote by  $\tilde{\zeta}_\varepsilon$  a continuous process that coincides with  $\zeta_\varepsilon$  at the points  $t_k^\varepsilon = k\varepsilon$  and is linear between any two consecutive points of that form. In this case, it is easy to check that

$$\sup_{t_{k-1}^\varepsilon \leq t \leq t_k^\varepsilon} |\zeta_\varepsilon(t) - \tilde{\zeta}_\varepsilon(t)|_2 \leq \sqrt{\varepsilon} |\eta_k|_2, \quad k \geq 1,$$

whence it follows that

$$\sup_{0 \leq t \leq t_0} |\zeta_\varepsilon(t) - \tilde{\zeta}_\varepsilon(t)|_2 \leq \sqrt{\varepsilon} \max\{|\eta_k|_2, 1 \leq k \leq [t_0/\varepsilon] + 1\}. \quad (2.74)$$

*Exercise 2.2.3.* Assuming that  $\mathbb{E}|\eta_1|_2^q < \infty$  for some  $q > 2$ , show that

$$\mathbb{P}\left\{ \sup_{0 \leq t \leq t_0} |\zeta_\varepsilon(t) - \tilde{\zeta}_\varepsilon(t)|_2 > \delta \right\} \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0, \quad (2.75)$$

where  $t_0 > 0$  and  $\delta > 0$  are arbitrary constants.

Combining this result with Theorem 2.2.2, it is not difficult to show that  $\mathcal{D}(\tilde{\zeta}_\varepsilon)$  converges to  $\mathcal{D}(\zeta)$  weakly in the space  $\mathcal{P}(D(0, t_0; H))$ . The following exercise shows that the convergence holds in fact for a stronger topology without any additional assumption on  $\eta_k$ .

*Exercise 2.2.4.* Prove that, for any  $t_0 > 0$ , the family  $\mathcal{D}(\tilde{\zeta}_\varepsilon)$  converges to  $\mathcal{D}(\zeta)$  weakly in the space  $\mathcal{P}(C(0, t_0; H))$ . *Hint:* Repeat the scheme used in the proof of Theorem 2.2.2, applying Donsker's theorem given in Section 8 of [Bil99].

## 2.3 Navier–Stokes equations perturbed by a random kick force

This section is devoted to a systematic study of the Navier–Stokes system (2.1) with a right-hand side given by (2.64), (2.65). We shall always assume that  $h \in H$  is a deterministic function and  $\{\eta_k\}$  is a sequence of  $H$ -valued i.i.d. random

variables defined on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . After projecting to the space  $H$ , the problem takes the form (cf. (2.19))

$$\dot{u} + \nu Lu + B(u) = h + \sum_{k=1}^{\infty} \eta_k \delta(t - kT). \quad (2.76)$$

As before, we shall consider the case of a torus and make remarks concerning the problem in a bounded domain. Accordingly, the problem will be studied in the functional spaces (2.41).

### 2.3.1 Existence and uniqueness of solution

We begin with a definition of the concept of a solution for (2.76). Let us denote by  $S_t : H \rightarrow H$  the resolving operator for Eq. (2.19) with  $f(t, x) = h(x)$ :

$$\dot{u} + \nu Lu + B(u) = h. \quad (2.77)$$

In other words,  $S_t$  takes  $u_0$  to  $u(t)$ , where  $u \in C(\mathbb{R}_+, H) \cap L^2_{\text{loc}}(\mathbb{R}_+, V)$  stands for the solution of (2.77) satisfying  $u(0) = u_0$ .

A *filtration* for (2.76) is defined as any non-decreasing family of  $\sigma$ -algebras  $\{\mathcal{G}_t, t \geq 0\}$  of the space  $(\Omega, \mathcal{F})$  such that  $\mathcal{G}_t = \mathcal{G}_{(k-1)T}$  for  $(k-1)T \leq t < kT$  and  $\eta_k$  is  $\mathcal{G}_{kT}$ -measurable and independent of  $\mathcal{G}_{(k-1)T}$  for any integer  $k \geq 1$ . Recall that a random process  $v(t)$ ,  $t \geq 0$  is said to be *adapted to  $\mathcal{G}_t$*  if the random variable  $v(t)$  is  $\mathcal{G}_t$ -measurable for any  $t \geq 0$ . Let us define the interval  $J_k = [(k-1)T, kT)$  and the space

$$\mathcal{H}(J_k) = \{u \in L^2(J_k, V) : \dot{u} \in L^2(J_k, V^*)\},$$

where the space  $V$  is defined in (2.41) and  $V^*$  denotes its dual. In view of Lemma 2.1.5, every element of  $\mathcal{H}(J_k)$  extends to a continuous curve  $\bar{J}_k \rightarrow H$ .

**Definition 2.3.1.** A random process  $u(t)$ ,  $t \geq 0$ , is called a *solution of Eq. (2.76)* if it is adapted to the filtration  $\mathcal{G}_t$  for (2.76), and almost every trajectory of  $u$  satisfies the following conditions for any integer  $k \geq 1$ :

- the restriction of  $u$  to  $J_k$  belongs to  $\mathcal{H}(J_k)$  and satisfies (2.77) (in particular,  $u : J_k \rightarrow H$  is a continuous curve which has a limit at the right endpoint of  $J_k$ );
- we have the relation

$$u(t_k^+) - u(t_k^-) = \eta_k, \quad (2.78)$$

where  $u(t_k^-)$  and  $u(t_k^+)$  denote the left- and right-hand limits of  $u$  at the point  $t_k := kT$ .

That is, on the interval  $(t_{k-1}, t_k)$  the function  $u$  is a solution of the “free” Navier–Stokes system (see (2.77)), and at any point  $t_k$  it has an instantaneous increment equal to the  $k^{\text{th}}$  kick  $\eta_k$  (see Figure 2.3). We shall always normalise



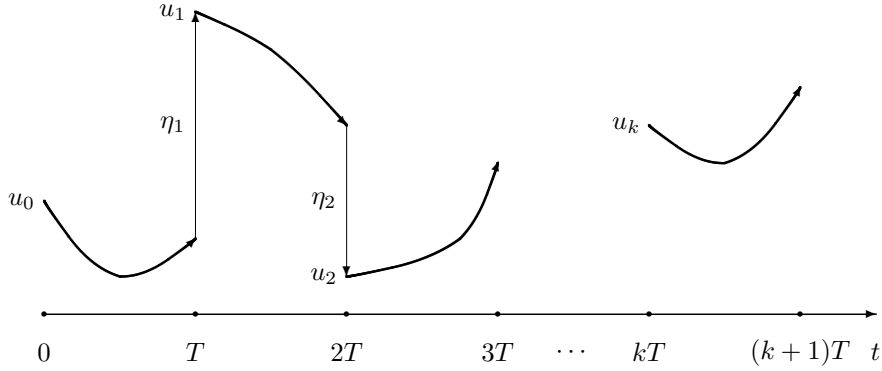


Figure 2.3: Evolution defined by Eq. (2.76)

the solutions of (2.76) by the condition of right-continuity at the points  $t_k$ ,  $k \geq 1$ . That is, denoting  $u_k = u(kT)$ ,  $k = 0, 1, \dots$ , we can write

$$u_k = S_T(u_{k-1}) + \eta_k, \quad k \geq 1, \quad (2.79)$$

and for  $t = kT + \tau$  with  $0 \leq \tau < T$ , we have  $u(t) = S_\tau(u_k)$ .

The following theorem establishes the existence and uniqueness of a solution for the Cauchy problem for (2.76). Its proof is a straightforward consequence of Theorem 2.1.13 and is left to the reader as an exercise.

**Theorem 2.3.2.** *Let  $u_0$  be an  $H$ -valued  $\mathcal{G}_0$ -measurable random variable. Then Eq. (2.76) has a solution satisfying the initial condition (2.3) for all  $\omega \in \Omega$ . Moreover, the solution is unique in the sense that if  $\tilde{u}$  is another random process with the above properties, then*

$$\mathbb{P}\{u(t) = \tilde{u}(t) \text{ for all } t \geq 0\} = 1.$$

### 2.3.2 Markov chain and RDS

Theorem 2.3.2 enables one to define a discrete-time process and an RDS in  $H$  which are associated with solutions of the Navier–Stokes system, evaluated at times  $kT$ ,  $k \in \mathbb{Z}_+$ . Indeed, the sequence  $\{u_k, k \geq 0\}$ , where  $u_k = u(kT)$ , satisfies Eq. (2.79). By Example 1.3.6, the latter defines a Markov family of random processes<sup>5</sup> on the extended probability space  $H \times \Omega$ , and by Example 1.3.15, it defines an RDS  $\Phi = \{\varphi_k^\omega\}$  in  $H$ , which is Markov in view of Exercise 1.3.19.

Recall that a measure  $\mu \in \mathcal{P}(H)$  is said to be *stationary* for (2.79) if  $\mathfrak{P}_1^* \mu = \mu$ , where  $\mathfrak{P}_k^*$  denotes the Markov semigroup associated with (2.79); see Section 1.3.1. A trajectory  $\{u_k, k \geq 0\}$  of (2.79) such that the law of  $u_k$  coincides with a stationary measure is called a *stationary solution*.

<sup>5</sup>Equation (2.79) itself defines a Markov chain (starting from  $u_0$ ) on the original probability space.

### 2.3.3 Additional results: higher Sobolev norms and time averages

Our next goal is to establish some estimates for solutions of (2.76). We begin with moments of solutions, evaluated at deterministic or random times proportional to  $T$ . To simplify the presentation, we assume that  $\nu = 1$ ; all the results remain valid for any  $\nu > 0$  if we suitably adjust the constants.

**Proposition 2.3.3.** (i) *In addition to the hypotheses of Theorem 2.3.2, assume that*

$$\mathbb{E}|u_0|_2^m < \infty, \quad K_m := \mathbb{E}|\eta_1|_2^m < \infty \quad (2.80)$$

for an integer  $m \geq 1$ . Then there are positive constants  $C_m = C_m(T)$  and  $q < 1$  such that

$$\mathbb{E}|u(kT)|_2^m \leq q^{kT} \mathbb{E}|u_0|_2^m + C_m(K_m + 1), \quad k \geq 1. \quad (2.81)$$

(ii) *If  $K_{2m} < \infty$  and  $\sigma$  is a stopping time with range in  $T\mathbb{Z}_+ \cup \{\infty\}$  satisfying the condition  $E := \sum_l (\mathbb{P}\{\sigma = lT\})^{1/2} < \infty$ , then we have*

$$\mathbb{E}(\mathbb{I}_{\{\sigma < \infty\}}|u(\sigma)|_2^m) \leq \mathbb{E}(q^\sigma|u_0|_2^m) + C_mE(K_{2m}^{1/2} + 1), \quad (2.82)$$

where we set  $q^\infty = 0$ .

*Proof.* Recall that  $u_k = u(kT)$ . Inequality (2.51) implies that

$$|S_T(v)|_2 \leq e^{-\alpha_1 T/2}|v|_2 + C_1|h|_2 \quad \text{for } v \in H.$$

Furthermore, it follows from (2.79) that

$$|u_k|_2 = |S_T(u_{k-1}) + \eta_k|_2 \leq |S_T(u_{k-1})|_2 + |\eta_k|_2.$$

Combining the last two relations, we obtain

$$|u_k|_2^m \leq q^T|u_{k-1}|_2^m + C_2(|\eta_k|_2^m + |h|_2^m), \quad (2.83)$$

where  $q < 1$  is a constant depending only on the first eigenvalue  $\alpha_1$ , and  $C_2 > 0$  depends on  $m$ . Iterating this inequality, we get

$$|u_k|_2^m \leq q^{kT}|u_0|_2^m + C_2 \sum_{l=1}^k q^{(k-l)T} (|\eta_l|_2^m + |h|_2^m). \quad (2.84)$$

Taking the mean value, we arrive at (2.81).

We now prove (2.82). Taking  $k = (\sigma/T) \wedge n$  in (2.84), where  $n \geq 1$  is an integer, and multiplying the resulting relation by  $\mathbb{I}_{\{\sigma < \infty\}}$ , we get

$$\begin{aligned} \mathbb{I}_{\{\sigma < \infty\}}|u_{(\sigma/T) \wedge n}|_2^m &\leq \mathbb{I}_{\{\sigma < \infty\}}q^{\sigma \wedge (nT)}|u_0|_2^m \\ &\quad + C_2\mathbb{I}_{\{1 \leq \sigma < \infty\}} \sum_{l=1}^{(\sigma/T) \wedge n} q^{\sigma \wedge (nT) - lT} (|\eta_l|_2^m + |h|_2^m) \\ &\leq q^{\sigma \wedge (nT)}|u_0|_2^m + C_2 \sum_{s=1}^{\infty} \sum_{l=1}^{s \wedge n} (q^T)^{s \wedge n - l} \mathbb{I}_{\{\sigma = sT\}} (|\eta_l|_2^m + |h|_2^m). \end{aligned}$$

Taking the expectation and using Fatou's lemma to pass to the limit as  $n \rightarrow \infty$  in the left-hand side of the above inequality, we derive

$$\mathbb{E}(\mathbb{I}_{\{\sigma < \infty\}} |u_{\sigma/T}|^m) \leq \mathbb{E}(q^\sigma |u_0|_2^m) + C_2 \mathcal{E}_m, \quad (2.85)$$

where

$$\mathcal{E}_m = \lim_{n \rightarrow \infty} \sum_{s=1}^{\infty} \sum_{l=1}^{s \wedge n} (q^T)^{s \wedge n - l} \mathbb{E}(\mathbb{I}_{\{\sigma = sT\}} (|\eta_l|_2^m + |h|_2^m)).$$

Applying the Cauchy–Schwarz inequality, we get

$$\begin{aligned} \mathcal{E}_m &\leq \lim_{n \rightarrow \infty} \sum_{s=1}^{\infty} \sum_{l=1}^{s \wedge n} (q^T)^{s \wedge n - l} (\mathbb{P}\{\sigma = sT\})^{\frac{1}{2}} ((\mathbb{E} |\eta_l|_2^{2m})^{\frac{1}{2}} + |h|_2^m) \\ &= (K_{2m}^{\frac{1}{2}} + 1) \lim_{n \rightarrow \infty} \sum_{s=1}^{\infty} \sum_{r=0}^{s \wedge n - 1} q^{Tr} (\mathbb{P}\{\sigma = sT\})^{\frac{1}{2}} \\ &\leq (K_{2m}^{\frac{1}{2}} + 1) \sum_{s=1}^{\infty} \sum_{r=0}^{s-1} q^{Tr} (\mathbb{P}\{\sigma = sT\})^{\frac{1}{2}}. \end{aligned}$$

Substituting this inequality into (2.85), we obtain (2.82).  $\square$

Proposition 2.3.3 implies that if all moments of the initial function  $u_0$  and those of the random perturbations  $\eta_k$  are finite, then so are moments of the solution. If we replace this hypothesis by the stronger condition of finiteness of a second exponential moment, then we can obtain some bounds of a similar quantity for solutions. More precisely, we have the following result.

**Proposition 2.3.4.** *In addition to the hypotheses of Theorem 2.3.2, assume that for some  $\varkappa_0 > 0$  we have*

$$\mathbb{E} \exp(\varkappa_0 |u_0|_2^2) < \infty, \quad Q := \mathbb{E} \exp(\varkappa_0 |\eta_1|_2^2) < \infty. \quad (2.86)$$

Then there are positive constants  $\varkappa$  and  $q < 1$  such that

$$\mathbb{E} \exp(\varkappa |u(kT)|_2^2) \leq C (\mathbb{E} \exp(\varkappa |u_0|_2^2))^{q^{kT}}, \quad k \geq 0, \quad (2.87)$$

where we set  $C = (Q \exp(\varkappa_0 |h|_2^2))^\gamma$  and  $\gamma = (1 - q^T)^{-1}$ .

*Proof.* It follows from inequality (2.83) with  $m = 2$  that

$$\begin{aligned} \mathbb{E} \exp(\varkappa |u_k|_2^2) &\leq \mathbb{E} \exp(\varkappa q^T |u_{k-1}|_2^2 + C_2 \varkappa (|\eta_k|_2^2 + |h|_2^2)) \\ &= \mathbb{E} \exp(\varkappa q^T |u_{k-1}|_2^2) \mathbb{E} \exp(C_2 \varkappa |\eta_k|_2^2) \exp(C_2 \varkappa |h|_2^2), \end{aligned}$$

where we used the independence of  $\eta_k$  and  $u_{k-1}$ . Choosing  $\varkappa = \varkappa_0 / C_2$  and applying Hölder's inequality, we derive

$$\mathbb{E} \exp(\varkappa |u_k|_2^2) \leq Q \exp(\varkappa_0 |h|_2^2) (\mathbb{E} \exp(\varkappa |u_{k-1}|_2^2))^{q^T}.$$

Iterating this inequality, we obtain (2.87).  $\square$

We now turn to some estimates for the mean value of higher Sobolev norms. Since the corresponding proofs are similar to those of Propositions 2.3.3 and 2.3.4, we shall confine ourselves to formulating the results.

**Proposition 2.3.5.** (i) *In addition to the hypotheses of Theorem 2.3.2, assume that*

$$\mathbb{E}|u_0|_2^m < \infty, \quad \mathbb{E}\|\eta_1\|_s^m < \infty \quad \text{for all } m \geq 1, \quad (2.88)$$

where  $s \geq 1$  is an integer. Then there are integers  $l_{sm} \geq m$  and a constant  $C_{sm}$  such that

$$\mathbb{E}\|u(kT)\|_s^m \leq C_{sm}(\mathbb{E}|u((k-1)T)|_2^{l_{sm}} + \mathbb{E}\|\eta_1\|_s^m + 1), \quad k \geq 1. \quad (2.89)$$

(ii) *In addition to the hypotheses of Theorem 2.3.2, assume that*

$$\mathbb{E}\exp(\varkappa_0|u_0|_2^2) < \infty, \quad \mathbb{E}\exp(\varkappa_0\|\eta_1\|_s^2) < \infty. \quad (2.90)$$

where  $s \geq 1$  is an integer. Then there are positive constants  $p_s$ ,  $\varkappa_s$ , and  $C_s$  such that

$$\mathbb{E}\exp(\varkappa_s\|u(kT)\|_s^{p_s}) \leq C_s\mathbb{E}\exp(\varkappa|u((k-1)T)|_2^2)\mathbb{E}\exp(\varkappa_0\|\eta_1\|_s^2), \quad (2.91)$$

where  $k \geq 1$  is an arbitrary integer and  $\varkappa$  is the constant in Proposition 2.3.4 (ii).

Let us note that, in view of Propositions 2.3.3 and 2.3.4, the right-hand sides of (2.89) and (2.91) are finite. Thus, to ensure that the moments of higher Sobolev norms of solutions are finite, it suffices to assume that so are the moments of the  $L^2$ -norm of the initial function. More precisely, we have the following result on uniform bounds for the mean value of higher Sobolev norms of solutions.

**Corollary 2.3.6.** *Under the hypotheses of parts (i) or (ii) of Proposition 2.3.5, we have respectively the following estimates for the solutions of Eq. (2.76) such that  $u(0) = u_0$ :*

$$\mathbb{E}\|u(kT)\|_s^m \leq C_{sm}(u_0), \quad k \geq 1, \quad (2.92)$$

$$\mathbb{E}\exp(\varkappa_s\|u(kT)\|_s^{p_s}) \leq C_{sm}(u_0), \quad k \geq 1, \quad (2.93)$$

where  $C_{sm}$  and  $C_s$  are some constants depending on the initial condition and the right-hand side of (2.76).

*Exercise 2.3.7.* Prove Proposition 2.3.5 and Corollary 2.3.6. *Hint:* Use Theorem 2.1.19 to establish (2.89) and (2.91); a proof can be found in the Appendix of [KS01b].

We now derive some estimates for the stopping time at which the time average of solutions becomes large. Given a function  $u(t)$  with range in  $H^1$  and integers  $l < k$ , we define

$$\langle \|u\|_1^2 \rangle_l^k = \frac{1}{k-l} \int_{lT}^{kT} \|u(t)\|_1^2 dt.$$

Let us fix a constant  $M > 0$  and consider the stopping time

$$\tau(M) = T \min\{k \geq 1 : \langle \|u\|_1^2 \rangle_0^{k+1} > M\}, \quad (2.94)$$

where we set  $\tau(M) = +\infty$  if the condition in the brackets is never satisfied.

**Proposition 2.3.8.** *Assume that the conditions of Proposition 2.3.4 are fulfilled. Then there are positive constants  $\delta = \delta(\varkappa_0, T)$  and  $C = C(T, h)$  such that*

$$\mathbb{P}\{\tau(M) = kT\} \leq R(CQ)^k e^{-\delta M(k+1)} \quad \text{for } k \geq 1, \quad (2.95)$$

where  $R = \mathbb{E} e^{\varkappa_0 |u_0|^2}$ . Furthermore, if  $\delta M > \ln(R + CQ)$ , then

$$\mathbb{P}\{\tau(M) = +\infty\} \geq \frac{R}{R + CQ} > 0. \quad (2.96)$$

*Proof. Step 1.* We first show that, for any integer  $k \geq 1$ ,

$$|u(t_k)|_2^2 + \int_0^{t_k} \|u(t)\|_1^2 dt \leq 2|u_0|_2^2 + C_1 \left( \sum_{l=1}^k |\eta_l|_2^2 + t_k \|h\|_{-1}^2 \right), \quad (2.97)$$

where  $t_k = kT$ , and  $C_1$  is a constant depending on  $T$ . Indeed, on any interval  $J_k = [t_{k-1}, t_k]$  the solution  $u(t, x)$  satisfies the deterministic Navier–Stokes equations (2.77) with  $\nu = 1$ . Thus, by (2.24), we have

$$|u(t_l^-)|_2^2 + \int_{t_{l-1}}^{t_l} \|u(t)\|_1^2 dt \leq |u(t_{l-1})|_2^2 + T \|h\|_{-1}^2.$$

Furthermore, relation (2.78) implies that

$$\begin{aligned} |u(t_l)|_2^2 &= |u(t_l^-)|_2^2 + |\eta_l|_2^2 + 2\langle \eta_l, u(t_l^-) \rangle \\ &\leq (1 + \mu) |u(t_l^-)|_2^2 + (1 + \mu^{-1}) |\eta_l|_2^2, \end{aligned}$$

where the constant  $\mu > 0$  will be chosen later. Combining these two inequalities, we derive

$$(1 + \mu)^{-1} |u(t_l)|_2^2 + \int_{t_{l-1}}^{t_l} \|u(t)\|_1^2 dt \leq |u(t_{l-1})|_2^2 + T \|h\|_{-1}^2 + \mu^{-1} |\eta_l|_2^2.$$

Taking the sum over  $l = 1, \dots, k$ , we obtain

$$\begin{aligned} (1 + \mu)^{-1} |u(t_k)|_2^2 + \int_0^{t_k} \|u(t)\|_1^2 dt \\ \leq |u_0|_2^2 + \frac{\mu}{1 + \mu} \sum_{l=1}^{k-1} |u(t_l)|_2^2 + t_k \|h\|_{-1}^2 + \mu^{-1} \sum_{l=1}^k |\eta_l|_2^2. \end{aligned}$$

It follows from inequality (2.53), applied to the intervals  $J_l$ ,  $1 \leq l \leq k-1$ , that

$$\sum_{l=1}^{k-1} |u(t_l)|_2^2 \leq C_2 \int_0^{t_k} \|u(t)\|_1^2 dt + C_2 t_k \|h\|_{-1}^2,$$

where  $C_2 > 0$  is a constant depending on  $T$ . Choosing the constant  $\mu \in (0, 1)$  so small that  $2\mu C_2 \leq 1 + \mu$ , we derive (2.97) from these two inequalities.

*Step 2.* Let us set

$$I_k = (k+1) \langle \|u\|_1^2 \rangle_0^{k+1}.$$

Inequality (2.97) implies that

$$I_k \leq 2|u_0|_2^2 + C_3 \left( \sum_{l=1}^k |\eta_l|_2^2 + t_{k+1} \|h\|_{-1}^2 \right).$$

Let the constant  $\delta > 0$  be so small that  $(C_3 \vee 2)\delta \leq \varkappa_0$ . Then it follows from this inequality and (2.86) that

$$\mathbb{E} e^{\delta I_k} \leq R(CQ)^k.$$

Chebyshev's inequality now implies that

$$\mathbb{P}\{I_k > M(k+1)\} \leq R(CQ)^k e^{-\delta M(k+1)}.$$

It remains to note that  $\{\tau(M) = k\} \subset \{I_k > M(k+1)\}$ .

*Step 3.* It remains to establish (2.96). Note that if  $\delta M > \ln(R + CQ)$ , then in view of (2.95) we have

$$\mathbb{P}\{\tau(M) < \infty\} = \sum_{k=1}^{\infty} \mathbb{P}\{\tau(M) = k\} \leq \sum_{k=1}^{\infty} R(CQ)^k e^{-\delta M(k+1)} \leq \frac{CQ}{R + CQ}.$$

This inequality readily implies the required result.  $\square$

## 2.4 Navier–Stokes equations perturbed by a spatially regular white noise

### 2.4.1 Existence and uniqueness of solution, and Markov process

Let us turn to the equation

$$\dot{u} + \nu Lu + B(u) = h + \eta, \quad \eta = \frac{\partial \zeta}{\partial t}, \quad \zeta(t, x) = \sum_{j=1}^{\infty} b_j \beta_j(t) e_j(x), \quad (2.98)$$

where  $h \in H$  is a deterministic function, and  $b_j$ ,  $\beta_j$  and  $e_j$  are the same as in (2.66). In particular, we assume that

$$\mathfrak{B} = \sum_{j=1}^{\infty} b_j^2 < \infty.$$

This ensures that  $\zeta(t, \cdot)$  is a continuous process with range in  $H$ . Since for almost every  $\omega \in \Omega$  the curve  $\zeta^\omega(t)$  is only  $H$ -continuous, i.e. has low smoothness, we cannot apply deterministic tools to construct solutions of (2.98); see Exercise 2.1.27. Using instead a stochastic approach, we shall show that, for a.e.  $\omega$  and for any initial function  $u(0) = u_0 \in H$ , Eq. (2.98) has a unique solution  $u(t, x)$ . Moreover, solutions can be written as  $u = U(u_0, \zeta^\omega(\cdot))$ , where  $U$  is a Borel mapping of its arguments.

Let us start with a concept of solution for (2.98). We shall always assume that the Brownian motions  $\{\beta_j\}$  are defined on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with a filtration  $\mathcal{G}_t$ ,  $t \geq 0$ , and that the  $\sigma$ -algebras  $\mathcal{G}_t$  are completed with respect to  $(\mathcal{F}, \mathbb{P})$ , that is,  $\mathcal{G}_t$  contains all the  $\mathbb{P}$ -null sets  $A \in \mathcal{F}$ . In this case, we shall say that the filtered probability space  $(\Omega, \mathcal{F}, \mathcal{G}_t, \mathbb{P})$  satisfies the *usual hypotheses*.

**Definition 2.4.1.** An  $H$ -valued random process  $u(t)$ ,  $t \geq 0$ , is called a *solution* for (2.98) if :

- (a) The process  $u(t)$  is adapted to the filtration  $\mathcal{G}_t$ , and its almost every trajectory belongs to the space<sup>6</sup>

$$\mathcal{X} := C(\mathbb{R}_+; H) \cap L_{\text{loc}}^2(\mathbb{R}_+, V).$$

- (b) Equation (2.98) holds in the sense that, with probability 1,

$$u(t) + \int_0^t (\nu Lu + B(u)) ds = u(0) + th + \zeta(t), \quad t \geq 0, \quad (2.99)$$

where the equality holds in the space  $H^{-1}$ .

It is easy to see that, with probability 1, the left- and right-hand sides of (2.99) belong to the space  $C(\mathbb{R}_+, H^{-1})$ , so relation (2.99) makes sense. The construction of a solution for (2.98) given below is based on a reduction to an equation with random coefficients. Namely, let us consider the stochastic Stokes equation

$$\dot{z} + \nu Lz = \eta(t, x) \quad t \geq 0. \quad (2.100)$$

The concept of a solution for (2.100) is similar to that for (2.98), and we do not repeat its definition. The following result uses the existence of an analytic

<sup>6</sup>We recall that  $C(\mathbb{R}_+; X)$  is a Fréchet space endowed with distance (1.9).

semigroup  $e^{-\nu tL}$  generated by the Stokes operator, and the concept of a stochastic convolution. We refer the reader to Chapter 1 of [Hen81] and Section 5.1 of [DZ92] for the corresponding definitions and results.<sup>7</sup>

**Proposition 2.4.2.** *Equation (2.100) has a solution  $z(\cdot) \in \mathcal{X}$  satisfying the initial condition  $z(0) = 0$ . Moreover, this solution is unique in the sense that if  $\tilde{z}(t)$  is another solution, then  $z \equiv \tilde{z}$  almost surely. Finally, the solution  $z$  can be represented as the stochastic convolution*

$$z(t) = \int_0^t e^{-\nu(t-s)L} d\zeta(s), \quad (2.101)$$

and for almost every  $\omega$  we have  $z^\omega = Z(\zeta^\omega)$ , where  $Z : C(\mathbb{R}_+; H) \rightarrow C(\mathbb{R}_+; V^*)$  is the continuous linear mapping defined by the right-hand side of (2.102).

*Proof.* The uniqueness of a solution can be established by the same argument as in the deterministic case (see the proof of Theorem 2.1.13), and therefore we shall confine ourselves to the proof of its existence and to representation (2.101).

Integrating by parts, we rewrite the stochastic integral (2.101) in the Wiener form (cf. Section 2.1 in [McK69])

$$\int_0^t e^{-\nu(t-s)L} d\zeta(s) = \zeta(t) - \nu \int_0^t L e^{-\nu(t-s)L} \zeta(s) ds. \quad (2.102)$$

In view of Exercise 2.2.1, we have  $\zeta \in C(0, T; H)$  almost surely. Since the operator  $L e^{-\nu tL}$  is continuous from  $H$  to  $V^*$  with a norm bounded by  $C(\nu t)^{-1/2}$ , it follows from (2.102) that  $z(t)$  is a random process in  $V^*$  that has continuous trajectories and is adapted to the filtration  $\mathcal{G}_t$ . The required assertions will be established if we prove that, with probability 1, the trajectories of  $z$  belongs to the space  $\mathcal{X}$  introduced in Definition 2.4.1 and satisfies the relation

$$z(t) + \nu \int_0^t L z(s) ds = \zeta(t), \quad t \geq 0. \quad (2.103)$$

The proof of this fact is rather simple in the case when the orthonormal basis  $\{e_j\}$  entering the definition of  $\zeta$  (see (2.98)) consists of the eigenfunctions of the Stokes operator  $L$ . To simplify the presentation, we shall consider only that particular case, leaving the general situation to the reader as an exercise.

Let us consider a sequence of processes  $z_n$  defined by

$$z_n(t) = \zeta^n(t) - \nu \int_0^t L e^{-\nu(t-s)L} \zeta^n(s) ds,$$

<sup>7</sup>The stochastic convolution is encountered only in this chapter. The reader not interested in that concept may regard the right-hand side of (2.102) as the definition of the stochastic convolution (the proof of Proposition 2.4.2 uses only representation (2.102)). Alternatively, the convolution  $z(t)$  can be defined by duality:

$$\langle \xi, z(t) \rangle = \int_0^t \langle e^{-\nu(t-s)L} \xi, d\zeta(s) \rangle = \sum_{j=1}^{\infty} b_j \int_0^t \xi_j(s) d\beta_j(s), \quad \xi \in H,$$

where  $\xi_j(s)$  is the  $j^{\text{th}}$  component of the vector  $e^{-\nu(t-s)L} \xi$ .



where  $\zeta^n(t) = \sum_{j=1}^n b_j \beta_j(t) e_j$ . Since the subspace  $H_n = \text{span}\{e_1, \dots, e_n\}$  is invariant under  $L$  and  $\|Le^{-cL}\|_{\mathcal{L}(H, V^*)} \leq c^{-1}$ , we see that  $z_n$  is an  $H_n$ -valued process with continuous trajectories such that

$$\sup_{0 \leq t \leq T} \mathbb{E} \|z_n(t) - z(t)\|_{V^*}^2 \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad (2.104)$$

where  $T > 0$  is an arbitrary constant. Moreover, it is easy to see that

$$z_n(t) + \nu \int_0^t Lz_n(s) ds = \zeta^n(t), \quad t \geq 0. \quad (2.105)$$

We claim that there is a subsequence  $z_{n_k}$  converging almost surely to a limit  $\bar{z}$  in the space  $\mathcal{X}_T = C(0, T; H) \cap L^2(0, T; V)$ . Indeed, for any integers  $m < n$  the  $H_n$ -valued process  $z_{mn} = z_n - z_m$  satisfies the stochastic differential equation

$$\dot{z}_{mn} + \nu Lz_{mn} = \sum_{j=m+1}^n b_j \dot{\beta}_j(t) e_j(x).$$

Applying Itô's formula to  $|z_{mn}(t)|_2^2$ , we derive

$$|z_{mn}(t)|_2^2 = \int_0^t (-2\nu |\nabla z_{mn}(s)|_2^2 + F_{mn}) ds + 2 \sum_{j=m+1}^n b_j \int_0^t \langle z_{mn}, e_j \rangle d\beta_j, \quad (2.106)$$

where  $F_{mn} = \sum_{j=m+1}^n b_j^2$ . Taking the mean value in (2.106), we see that

$$\sup_{0 \leq t \leq T} \mathbb{E} |z_{mn}(t)|_2^2 + 2\nu \mathbb{E} \int_0^T \|z_{mn}(s)\|_1^2 dt \leq C_4 T F_{mn}. \quad (2.107)$$

Furthermore, application of Doob's moment inequality (7.56) to the stochastic integral in (2.106) results in (cf. derivation of (2.73))

$$\mathbb{E} \sup_{0 \leq t \leq T} |z_{mn}(t)|_2^2 \leq F_{mn} T + C_6 \mathbb{E} \int_0^T \sum_{j=1}^n b_j^2 (z_{mn}(t), e_j)^2 dt. \quad (2.108)$$

Now note that  $F_{mn} \rightarrow 0$  as  $m, n \rightarrow \infty$ . Combining this with (2.107) and (2.108), we conclude that

$$\mathbb{E} \sup_{0 \leq t \leq T} |z_{mn}(t)|_2^2 + \mathbb{E} \int_0^T \|z_{mn}(s)\|_1^2 dt \rightarrow 0 \quad \text{as } m, n \rightarrow \infty. \quad (2.109)$$

Thus, the sequence  $\{z_n\}$  converges to a limit  $\bar{z}$  in the space  $L^2(\Omega, \mathcal{X}_T)$ . It follows that a subsequence  $\{z_{n_k}\}$  converges almost surely to  $\bar{z}$  in  $\mathcal{X}_T$ . Since

$$z_{n_k}(t) + \nu \int_0^t Lz_{n_k}(s) ds = \zeta^{n_k}(t), \quad 0 \leq t \leq T,$$

where the equality holds in  $H^{-1}$  for almost all  $\omega \in \Omega$ , then passing to the limit as  $k \rightarrow \infty$ , we arrive at relation (2.103) with  $z = \bar{z}$ . Thus, the function  $\bar{z}$  is a solution of (2.100) vanishing at  $t = 0$ . The fact that  $\bar{z}$  can be written in the form (2.101) can easily be established by passing to the limit in a similar relation for  $z_{n_k}$ .  $\square$

*Remark 2.4.3.* Denote

$$H_\zeta = \overline{\text{span}\{e_j : b_j \neq 0\}} \subset H, \quad (2.110)$$

and denote by  $m_\zeta$  the law of  $\zeta$ . This is a measure on  $C(\mathbb{R}_+; H)$  whose support equals  $C(\mathbb{R}_+; H_\zeta)$ . We write  $\mathcal{F}_\zeta$  for the  $m_\zeta$ -completion of the Borel  $\sigma$ -algebra on  $C(\mathbb{R}_+; H)$ . The process  $z_n(t)$  constructed in the proof has the form  $z_n = Z_n(\zeta)$ , where  $Z_n$  is a continuous operator from  $C(\mathbb{R}_+; H)$  to  $\mathcal{X}$ , and our proof shows that  $m_\zeta$ -almost surely the sequence  $\{Z_n\}$  converges to  $Z$ . In particular, it follows that

$$Z : (C(\mathbb{R}_+; H), \mathcal{F}_\zeta) \rightarrow (\mathcal{X}, \mathcal{B}(\mathcal{X}))$$

is a measurable mapping. It is straightforward to see that, for any  $T > 0$ , the restriction of  $Z(\zeta)$  to  $[0, T]$  depends only on  $\zeta_{[0, T]}$ . Finally, if

$$\mathfrak{B}_1 = \sum_{j=1}^{\infty} \alpha_j b_j^2 < \infty,$$

where  $\{\alpha_j\}$  are the eigenvalues of the Stokes operator, then  $\zeta \in C(\mathbb{R}_+; V)$  almost surely. The restriction of  $Z$  to  $C(\mathbb{R}_+; V)$  defines a continuous linear mapping from  $C(\mathbb{R}_+; V)$  to  $\mathcal{X}$ .

*Exercise 2.4.4.* Without assuming that  $\{e_j\}$  is an eigenbasis for  $L$ , prove that the stochastic convolution (2.101) defines a random process that has continuous trajectories and satisfies relation (2.103).

We now show how to reduce the stochastic Navier-Stokes system (2.98) to an equation with random coefficients. As before, we assume that the deterministic and random forces  $h$  and  $\eta$  satisfy the hypotheses formulated in the beginning of this subsection. From now on, we change slightly the definition of the space  $\mathcal{H}$  and denote

$$\mathcal{H} = \{u \in L^2(0, T; V) : \dot{u} \in L^2(0, T; V^*)\};$$

cf. Section 2.1.3. We seek a solution in the form  $u = z + v$ , where  $z(t)$  is the process constructed in Proposition 2.4.2. The function  $v$  must satisfy the equations

$$\dot{v} + \nu Lv + B(v + z) = h, \quad (2.111)$$

$$v(0) = u_0. \quad (2.112)$$

This is a Navier-Stokes type system with random coefficients entering through the stochastic process  $z$ .

**Proposition 2.4.5.** *Let  $\nu$  and  $T$  be some positive constants and let  $h \in H$  be a given function. Then for any  $u_0 \in H$  and  $z \in \mathcal{X}_T$  problem (2.111), (2.112) has a unique solution  $v \in \mathcal{H}$ . Moreover, the operator  $\mathcal{R} : H \times \mathcal{X}_T \rightarrow \mathcal{H}$  taking a pair  $(u_0, z)$  to the solution  $v \in \mathcal{H}$  is locally Lipschitz continuous. That is, for any  $R > 0$  there is a constant  $C_R > 0$  such that*

$$\|\mathcal{R}(u_{01}, z_1) - \mathcal{R}(u_{02}, z_2)\|_{\mathcal{H}} \leq C_R(|u_{01} - u_{02}|_2 + \|z_1 - z_2\|_{\mathcal{X}_T}), \quad (2.113)$$

where  $u_{0i} \in H$  and  $z_i \in \mathcal{X}_T$  are arbitrary functions whose norms do not exceed  $R$ .

In the case  $z \equiv 0$ , this result is a straightforward consequence of Theorem 2.1.13 and Proposition 2.1.25. The case of a non-zero function  $z$  can be handled with similar arguments, and therefore we omit the proof; cf. Exercise 2.1.27. We are now ready to establish the main theorem on existence and uniqueness of a solution for the stochastic Navier–Stokes equation (2.98).

**Theorem 2.4.6.** *For any  $\nu > 0$  and any  $\mathcal{G}_0$ -measurable random variable  $u_0(x)$  the Navier–Stokes system (2.98) has a solution  $u(t)$ ,  $t \geq 0$ , satisfying the initial condition*

$$u(0) = u_0 \quad \text{almost surely.} \quad (2.114)$$

Moreover, if  $\tilde{u}(t)$  is another solution of (2.98), (2.114), then  $u(t) \equiv \tilde{u}(t)$  almost surely. Furthermore, the solution  $u(t)$  possesses the following properties.

- (i) *Almost all trajectories of  $u(t)$  are continuous with range in  $H$  and locally square integrable with range in  $V$ .*
- (ii) *The process  $u(t)$  can be written in the form*

$$u(t) = u_0 + \int_0^t f(s) ds + \zeta(t), \quad t \geq 0, \quad (2.115)$$

where  $f(t)$  is a  $V^*$ -valued  $\mathcal{G}_t$ -progressively measurable process such that

$$\mathbb{P}\left\{\int_0^T \|f(t)\|_{V^*}^2 dt < \infty \text{ for any } T > 0\right\} = 1. \quad (2.116)$$

Let  $\mathcal{C}_T = C(0, T; H)$  and  $m_{\zeta, T} = \mathcal{D}(\zeta|_{[0, T]})$ . Then  $\text{supp } m_{\zeta, T} = C(0, T; H_\zeta)$ ; see (2.110). Let  $\mathcal{F}_{\zeta, T}$  be  $m_{\zeta, T}$ -the completion of  $\mathcal{B}(\mathcal{C}_T)$ . Remark 2.4.3 and Proposition 2.4.5 imply the following result.

**Proposition 2.4.7.** *There exists a measurable mapping*

$$U : (H \times \mathcal{C}_T, \mathcal{B}(H) \otimes \mathcal{F}_{\zeta, T}) \rightarrow (\mathcal{X}_T, \mathcal{B}(\mathcal{X}_T)),$$

which is locally Lipschitz continuous in  $u \in H$  for  $m_{\zeta, T}$ -almost every  $\omega \in \mathcal{C}_T$ , such that

$$u^\omega = U(u_0, \zeta^\omega) \quad \text{for } m_{\zeta, T}\text{-almost every } \omega \text{ and all } u_0 \in H. \quad (2.117)$$

Moreover, the restriction of  $U$  to  $H \times C(0, T; V)$  is locally Lipschitz continuous in both variables.

In particular, if we denote by  $U_t$  the restriction of  $U$  at time  $t \in [0, T]$ , then for any (random) initial function  $u_0^\omega$  which is independent of  $\zeta^\omega$  the law of  $u^\omega(t)$  can be written as the image of the product measure  $\mathcal{D}(u_0) \otimes m_{\zeta, t}$  under the mapping  $U_t$ :

$$\mathcal{D}(u^\omega(t)) = U_{t*}(\lambda \otimes m_{\zeta, t}), \quad \lambda = \mathcal{D}(u_0). \quad (2.118)$$

*Proof of Theorem 2.4.6.* We first prove the uniqueness. If  $u$  and  $\tilde{u}$  are two solutions, then for almost every  $\omega \in \Omega$  the difference  $w = u - \tilde{u}$  belongs to  $\mathcal{X}$ , vanishes at  $t = 0$ , and satisfies the equation

$$\dot{w} + \nu Lw + B(w, u) + B(\tilde{u}, w) = 0.$$

This equation implies that  $\dot{w} \in L^2(0, T; V^*)$  for any  $T > 0$ , and as in the case of deterministic equation (see the proof of Theorem 2.1.13), it follows that  $w \equiv 0$  almost surely.

To prove the existence, we denote by  $z(t)$  the solution of Eq. (2.100) vanishing at zero; see Proposition 2.4.2. Let  $\Omega_0 \subset \Omega$  be a set of full measure that consists of those  $\omega \in \Omega$  for which  $z \in \mathcal{X}$ . We define  $v \in \mathcal{X}$  as the solution of problem (2.111), (2.112) for  $\omega \in \Omega_0$  and set  $v = 0$  on the complement of  $\Omega_0$ . Then the random process  $u = v + z$  is a solution of problem (2.98), (2.114).

We now prove assertions (i) and (ii). The continuity of trajectories of  $u$  in  $H$  follows from a similar property for  $z$  and  $v$ . Relation (2.99) implies that  $u(t)$  can be written in the form (2.115) with  $f(t) = h - \nu Lu(t) - B(u(t))$ . Finally, it follows from Propositions 2.4.2 and 2.4.5 that (2.116) holds.  $\square$

Theorem 2.4.6 established above not only gives the existence and uniqueness of a solution for the stochastic Navier–Stokes equation (2.98), but also ensures that the solution meets Condition 7.7.4 from Section 7.7. Therefore the infinite-dimensional Itô formula proved in the Appendix applies to the constructed solution. This provides us with a convenient tool for deriving a priori estimates without resorting to finite-dimensional approximations of the Navier–Stokes system.

In conclusion, let us note that the family of solutions corresponding to all possible initial data form a Markov family; see Section 1.3.1. This fact is true for “any” PDE perturbed by the time derivative of a stochastic process with independent increments, provided that the corresponding Cauchy problem is well posed. We refer the reader to Chapter 7 of [Øks03] for the proof of the Markov property in the case of SDE’s (see also Section V.5 of [Kry02]) and to Section 9.2 of [DZ92] where this property is proved for various classes of SPDE’s. In the context of the Navier–Stokes equations perturbed by a spatially regular white noise, the Markov property can be established with the help of a similar argument, using relation (2.118) for the law of a solution. We also note that, as will be shown in Section 2.4.4, the family of solutions for (2.98) form a Markov RDS, and as was explained in Section 1.3.3, to each Markov RDS one can associate, in a canonical way, a Markov process. This provides a rigorous proof of the above-mentioned claim.

For the reader’s convenience, we now briefly describe the construction of a Markov process associated with the stochastic Navier–Stokes system. Let  $\tilde{\Omega} = H \times \Omega$ , where  $(\Omega, \mathcal{F}, \mathcal{G}_t, \mathbb{P})$  is the filtered probability space defined in the beginning of this section (see page 67), and let

$$\tilde{\mathcal{F}}_t = \mathcal{B}(H) \otimes \mathcal{G}_t, \quad \mathbb{P}_v = \delta_v \otimes \mathbb{P},$$

where  $v \in H$ . Writing  $\tilde{\omega} = (v, \omega)$  for points of  $\tilde{\Omega}$ , we now define the process  $t \mapsto u_t(\tilde{\omega}) \equiv u^\omega(t; v)$ , where  $u^\omega(t; v)$  denotes the solution of (2.98) constructed in Theorem 2.4.6 with a deterministic initial function  $u_0 = v$ . We thus obtain a family of Markov processes  $(u_t, \mathbb{P}_v)$ ,  $v \in H$ , associated with (2.98). All concepts and properties that are valid in the general setting of Markov processes apply in the case under study. In particular, a measure  $\mu \in \mathcal{P}(H)$  is said to be *stationary* for Eq. (2.98) if  $\mathfrak{P}_t^* \mu = \mu$  for all  $t \geq 0$ , where  $\mathfrak{P}_t^* : \mathcal{P}(H) \rightarrow \mathcal{P}(H)$  denotes the Markov semigroup for the family  $(u_t, \mathbb{P}_v)$ , and a solution  $u(t)$  of Eq. (2.98) is called *stationary* if its law  $\mathcal{D}(u(t))$  coincides with a stationary measure for all  $t \geq 0$ . Let us also note that  $\mathfrak{P}_t^* \lambda$  coincides with the right-hand side of (2.118).

### 2.4.2 Additional results: energy balance, higher Sobolev norms, and time averages

Our first result concerns the *energy balance* for solutions of the Navier–Stokes system (2.98). It is an analogue of relation (2.39) in the stochastic case.

**Proposition 2.4.8.** *Under the hypotheses of Theorem 2.4.6, assume that the initial function  $u_0$  has a finite second moment,  $\mathbb{E} |u_0|_2^2 < \infty$ . Then for any  $\nu > 0$  the following relation holds for a solution  $u(t)$  of problem (2.98), (2.114):*

$$\mathbb{E} |u(t)|_2^2 + 2\nu \mathbb{E} \int_0^t |\nabla u(s)|_2^2 ds = \mathbb{E} |u_0|_2^2 + \mathfrak{B}t + 2\mathbb{E} \int_0^t \langle u, h \rangle ds, \quad t \geq 0. \quad (2.119)$$

Moreover,

$$\mathbb{E} |u(t)|_2^2 \leq e^{-\nu\alpha_1 t} \mathbb{E} |u_0|_2^2 + (\nu\alpha_1)^{-1} \mathfrak{B} + (\nu\alpha_1)^{-2} |h|_2^2, \quad t \geq 0, \quad (2.120)$$

where  $\alpha_1 > 0$  stands for the first eigenvalue<sup>8</sup> of the Stokes operator  $L$ .

*Proof.* We wish to apply Theorem 7.7.5 to the functional  $F(u) = |u|_2^2$ . Since

$$\partial_u F(u; v) = 2\langle u, v \rangle, \quad \partial_u^2 F(u; v) = 2|v|_2^2,$$

relation (7.24) takes the form

$$|u(t \wedge \tau_n)|_2^2 = |u_0|_2^2 + 2 \int_0^{t \wedge \tau_n} (\langle u, h - \nu Lu \rangle + \mathfrak{B}) ds + 2 \sum_{j=1}^{\infty} b_j \int_0^{t \wedge \tau_n} \langle u, e_j \rangle d\beta_j(s), \quad (2.121)$$

where we set

$$\tau_n = \inf\{t \geq 0 : |u(t)|_2 > n\}.$$

Taking the mean value in (2.121) and using Doob's optional sampling theorem, we obtain

$$\mathbb{E} |u(t \wedge \tau_n)|_2^2 + 2\nu \mathbb{E} \int_0^{t \wedge \tau_n} |\nabla u(s)|_2^2 ds = \mathbb{E} |u_0|_2^2 + \mathfrak{B} \mathbb{E} (t \wedge \tau_n) + 2\mathbb{E} \int_0^{t \wedge \tau_n} \langle u, h \rangle ds. \quad (2.122)$$

---

<sup>8</sup>Recall that, in the case of periodic boundary conditions, we consider the problem in the spaces of functions with zero mean value, and therefore the first eigenvalue of  $L$  is positive.

Since the trajectories of  $u(t)$  are continuous  $H$ -valued functions of time, we have  $\tau_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Passing to the limit in (2.122) as  $n \rightarrow \infty$  and using the monotone convergence theorem, we arrive at (2.119).

We now note that

$$|\nabla u|_2^2 \geq \alpha_1 |u|_2^2, \quad 2|\langle u, h \rangle| \leq \nu \alpha_1 |u|_2^2 + (\nu \alpha_1)^{-1} |h|_2^2.$$

Substituting these estimates into (2.119) and using Gronwall's inequality, we arrive at (2.120).  $\square$

The technique used in the previous proposition applies to higher moments of the  $L^2$ -norm of solutions and gives some a priori estimates for  $\mathbb{E} |u(t)|_2^{2m}$  in terms of  $E |u_0|_2^{2m}$ , where  $m \geq 2$  is an arbitrary integer. We shall not give either a proof or a formulation of the corresponding results, referring the reader to the book [VF88]. Instead we now establish an estimate for the second exponential moment that will imply the above-mentioned results under a stronger assumption on the initial function.

**Proposition 2.4.9.** *Under the hypotheses of Theorem 2.4.6, there is a constant  $c > 0$  not depending on  $\nu$ ,  $h$ , and  $\{b_j\}$  such that if  $\varkappa > 0$  satisfies the condition*

$$\varkappa \sup_{j \geq 1} b_j^2 \leq c, \tag{2.123}$$

*then the following assertion holds: for any  $\mathcal{G}_0$ -measurable  $H$ -valued random variable  $u_0(x)$  such that*

$$\mathbb{E} \exp(\varkappa \nu |u_0|_2^2) < \infty \tag{2.124}$$

*the corresponding solution of (2.98), (2.114) satisfies the inequality*

$$\mathbb{E} \exp(\varkappa \nu |u(t)|_2^2) \leq e^{-\varkappa \nu^2 t} \mathbb{E} \exp(\varkappa \nu |u_0|_2^2) + K(\nu, \varkappa, \mathfrak{B}, h), \tag{2.125}$$

*where we set*

$$K(\nu, \varkappa, \mathfrak{B}, h) = \nu^{-1} R \exp(\varkappa R / \alpha_1), \quad R = C \nu^{-1} |h|_2^2 + \mathfrak{B} + \nu, \tag{2.126}$$

*and  $C > 0$  is an absolute constant.*

*Proof.* We wish to apply Theorem 7.7.5 to a functional  $F : H \rightarrow \mathbb{R}$  defined as  $F(u) = \exp(\varkappa \nu |u|_2^2)$ . It is straightforward to verify that

$$\begin{aligned} \partial_u F(u; v) &= 2\varkappa \nu F(u) \langle u, v \rangle, \\ \partial_u^2 F(u; v) &= 2\varkappa \nu F(u) (|v|_2^2 + 2\varkappa \nu \langle u, v \rangle^2). \end{aligned}$$

It follows that relation (7.24) can be rewritten as

$$F(u(t \wedge \tau_n)) = F(u_0) + \int_0^{t \wedge \tau_n} A(s) ds + M(t \wedge \tau_n), \tag{2.127}$$

where  $M(t)$  is a stochastic integral (cf. (2.121)), and

$$A(t) = 2\kappa\nu F(u(t)) \left( -\nu \langle u(t), Lu(t) \rangle + \langle u(t), h \rangle + \frac{1}{2} \mathfrak{B} + \kappa\nu \sum_{j=1}^{\infty} b_j^2 \langle u(t), e_j \rangle^2 \right).$$

Taking the mean value in (2.127) and using Doob's optional sampling theorem, we derive

$$\mathbb{E} F(u(t \wedge \tau_n)) = \mathbb{E} F(u_0) + \mathbb{E} \int_0^{t \wedge \tau_n} A(s) ds. \quad (2.128)$$

Now note that if (2.123) holds with  $c \leq (8\alpha_1)^{-1}$ , then

$$\begin{aligned} A(t) &\leq \kappa\nu F(u(t)) (-\nu \|u(t)\|_V^2 + C_1 \nu^{-1} |h|_2^2 + \mathfrak{B}) \\ &\leq -\kappa\nu^2 F(u(t)) + \kappa\nu^2 K(\nu, \kappa, \mathfrak{B}, h), \end{aligned}$$

where  $\|u\|_V^2 = \langle u, Lu \rangle$ . Substituting this inequality into (2.128) and using Fatou's lemma to pass to the limit as  $n \rightarrow \infty$ , we obtain

$$\mathbb{E} \exp(\kappa\nu |u(t)|_2^2) + \kappa\nu^2 \int_0^t \mathbb{E} \exp(\kappa\nu |u(s)|_2^2) ds \leq \kappa\nu^2 K(\nu, \kappa, \mathfrak{B}, h).$$

Application of Gronwall's inequality completes the proof of the proposition.  $\square$

The next result gives an estimate for the rate of growth of solutions as  $t \rightarrow \infty$ . Let us set

$$\mathcal{E}_u(t) = |u(t)|_2^2 + \nu \int_0^t |\nabla u(s)|_2^2 ds. \quad (2.129)$$

**Proposition 2.4.10.** *Assume that the hypotheses of Theorem 2.4.6 are fulfilled, and  $u_0$  is an  $H$ -valued random variable with a finite second moment. Then for any  $\rho > 0$  the solution  $u(t)$  of problem (2.98), (2.114) satisfies the inequality*

$$\mathbb{P} \left\{ \sup_{t \geq 0} (\mathcal{E}_u(t) - (\mathfrak{B} + 2(\alpha_1 \nu)^{-1} |h|_2^2) t) \geq |u_0|_2^2 + \rho \right\} \leq e^{-\gamma \nu \rho}, \quad (2.130)$$

where we set  $\gamma = \frac{1}{4} \alpha_1 (\sup_j b_j^2)^{-1}$ .

*Proof.* Let us recall that relation (2.121) was established in the proof of Proposition 2.4.8. In view of (2.119), we have

$$\mathbb{E} \sum_{j=1}^{\infty} b_j^2 \int_0^t \langle u(s), e_j \rangle^2 ds \leq \left( \sup_{j \geq 1} b_j^2 \right) \int_0^t |u(s)|_2^2 ds < \infty \quad \text{for any } t \geq 0.$$

So the stochastic integral

$$M_t = 2 \sum_{j=1}^{\infty} b_j \int_0^t \langle u(s), e_j \rangle d\beta_j(s)$$

defines a martingale (see Section 7.11), and relation (2.121) implies that, with probability 1, we have

$$\mathcal{E}_u(t) = |u_0|_2^2 + \mathfrak{B}t + \nu \int_0^t |\nabla u(s)|_2^2 ds + 2 \int_0^t \langle u(s), h \rangle ds + M_t, \quad t \geq 0. \quad (2.131)$$

To estimate the right-hand side of this relation, we apply the classical method of exponential supermartingales; e.g., see Problem 5 in Section 2.9 of [McK69] for the one-dimensional case. Namely, noting that the quadratic variation of  $M_t$  is equal to

$$\langle M \rangle_t = 4 \sum_{j=1}^{\infty} b_j^2 \int_0^t \langle u(s), e_j \rangle^2 ds$$

(see Section 7.11 and [Kry02]), we rewrite (2.131) in the form

$$\mathcal{E}_u(t) - \mathfrak{B}t = |u_0|_2^2 + \left( M_t - \frac{1}{2} \gamma \nu \langle M \rangle_t \right) + K_t, \quad (2.132)$$

where we set

$$K_t = 2 \int_0^t \langle u(s), h \rangle ds + \frac{1}{2} \gamma \nu \langle M \rangle_t - \nu \int_0^t |\nabla u(s)|_2^2 ds.$$

Since

$$\begin{aligned} \left| \int_0^t \langle u(s), h \rangle ds \right| &\leq \frac{\nu}{4} \int_0^t |\nabla u(s)|_2^2 ds + (\alpha_1 \nu)^{-1} |h|_2^2 t, \\ \langle M \rangle_t &\leq 4\alpha_1^{-1} \sup_{j \geq 1} b_j^2 \int_0^t |\nabla u(s)|_2^2 ds = \frac{1}{\gamma} \int_0^t |\nabla u(s)|_2^2 ds, \end{aligned}$$

we see that

$$K_t \leq 2(\alpha_1 \nu)^{-1} |h|_2^2 t.$$

Substituting this inequality into (2.132), we obtain

$$\mathcal{E}_u(t) - (\mathfrak{B} + 2(\alpha_1 \nu)^{-1} |h|_2^2) t \leq |u_0|_2^2 + \left( M_t - \frac{1}{2} \gamma \nu \langle M \rangle_t \right). \quad (2.133)$$

Consider the exponential function  $\exp(\gamma \nu M_t - \frac{1}{2} (\gamma \nu)^2 \langle M \rangle_t)$ . Applying the exponential supermartingale result given in Section 7.11, we see that

$$\begin{aligned} \mathbb{P} \left\{ \sup_{t \geq 0} (M_t - \frac{1}{2} \gamma \nu \langle M \rangle_t) \geq \rho \right\} &= \mathbb{P} \left\{ \sup_{t \geq 0} \exp(\gamma \nu M_t - \frac{1}{2} (\gamma \nu)^2 \langle M \rangle_t) \geq e^{\gamma \nu \rho} \right\} \\ &\leq e^{-\gamma \nu \rho}. \end{aligned}$$

Combining this with (2.133), we arrive at the required inequality (2.130).  $\square$

**Corollary 2.4.11.** *Under the hypotheses of Proposition 2.4.10, assume that the random variable  $u_0$  satisfies the condition*

$$\mathbb{E} |u_0|_2^{2m} < \infty, \quad (2.134)$$



where  $m \geq 1$  is an integer. Then there is a constant  $C_m > 0$  depending only on  $m$  such that for any  $T \geq 1$  we have

$$\mathbb{E} \sup_{0 \leq t \leq T} \mathcal{E}_u(t)^m \leq C_m (\mathbb{E} |u_0|_2^{2m} + T^m \mathfrak{B}^m + \nu^{-m} (\gamma^{-m} + \alpha_1^{-m} T^m |h|_2^{2m})), \quad (2.135)$$

where  $\gamma > 0$  is the constant defined in Proposition 2.4.10.

*Proof.* Inequality (2.130) implies that

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq T} \mathcal{E}_u(t) \geq |u_0|_2^2 + CT + \rho \right\} \leq e^{-\gamma \nu \rho}, \quad (2.136)$$

where we set  $C = \mathfrak{B} + 2(\alpha_1 \nu)^{-1} |h|_2^2$ . Now note that if  $\xi$  and  $\eta$  are non-negative random variables, then

$$\begin{aligned} \mathbb{E} \xi^m &\leq 2^{m-1} (\mathbb{E} (\xi - \eta)^m \mathbb{I}_{\{\xi > \eta\}} + \mathbb{E} \eta^m) \\ &= 2^{m-1} \int_0^\infty \mathbb{P} \{ \xi - \eta > \lambda^{1/m} \} d\lambda + 2^{m-1} \mathbb{E} \eta^m. \end{aligned}$$

Applying this inequality to  $\xi = \sup_t \mathcal{E}_u(t)$  and  $\eta = |u_0|_2^2 + CT$  and using (2.136), we derive

$$\mathbb{E} \sup_{0 \leq t \leq T} \mathcal{E}_u(t)^m \leq 2^{m-1} \int_0^\infty \exp(-\gamma \nu \lambda^{1/m}) d\lambda + 2^{m-1} \mathbb{E} (|u_0|_2^2 + CT)^m.$$

It remains to note that the right-hand side can be estimated by that of (2.135) with a constant  $C_m$  depending only on  $m$ .  $\square$

Let us emphasise that Theorem 2.4.6, Propositions 2.4.8 – 2.4.10, and Corollary 2.4.11 are valid both for periodic and Dirichlet boundary conditions, and the same proof work in these two cases. In the next result concerning higher Sobolev norms, it is essential that the problem is considered under periodic boundary conditions. To simplify the presentation, we shall assume that the sequence  $\{e_j\}$  entering the right-hand side of (2.98) coincides with the family of normalised eigenfunctions of the Stokes operator  $L$  and denote by  $\alpha_1 \leq \alpha_2 \leq \dots$  the corresponding eigenvalues. For any integer  $k \geq 0$ , we set

$$\mathcal{E}_u(k, t) = t^k \|u(t)\|_k^2 + \nu \int_0^t s^k \|u(s)\|_{k+1}^2 ds,$$

so that  $\mathcal{E}_u(0, t)$  coincides with the function  $\mathcal{E}_u(t)$  defined by (2.129).

**Proposition 2.4.12.** *Let  $k \geq 1$  be an integer, let  $h \in V^k$ , and let the sequence  $\{b_j\}$  satisfy the condition*

$$\mathfrak{B}_k := \sum_{j=1}^{\infty} \alpha_j^k b_j^2 < \infty. \quad (2.137)$$

Let  $u_0$  be a  $\mathcal{G}_0$ -measurable  $H$ -valued random variable satisfying (2.134) for any integer  $m \geq 1$ . Then for any  $m \geq 1$  and  $T \geq 1$  there is a  $C(k, m, T) > 0$  such that

$$\mathbb{E} \sup_{0 \leq t \leq T} \mathcal{E}_u(k, t)^m \leq C(k, m, T)(1 + \nu^{-m(7k+2)}(\mathbb{E}|u_0|_2^{4m(k+1)} + 1)). \quad (2.138)$$

In what follows, this result is only used in Exercise 2.5.8 (for proving similar bounds for stationary solutions) and in Section 3.5.3 (devoted to convergence of mean values of functionals on higher Sobolev space). Its proof is based on the standard idea of differentiating higher Sobolev norms along solutions. Accurate calculations, which involve Itô's formula for infinite-dimensional diffusions and some interpolation inequalities, are rather technical, and we present them in Section 2.6. The reader not interested in these details may safely skip them, together with the rest of this subsection, and jump directly to the final part of this section, which is devoted to the construction of a Markov RDS corresponding to Navier–Stokes equations with white noise; see page 81.

**Corollary 2.4.13.** *Under the conditions of Proposition 2.4.12, assume that inequality (2.124) holds for the initial function  $u_0$  with a constant  $\varkappa > 0$  satisfying (2.123). Then for any integer  $m \geq 1$  and any constants  $T \geq 1$  and  $t_0 \geq 1$  we have*

$$\begin{aligned} & \mathbb{E} \sup_{t_0 \leq t \leq t_0+T} \|u(t)\|_k^{2m} + \nu^m \mathbb{E} \left( \int_{t_0}^{t_0+T} \|u(s)\|_{k+1}^2 ds \right)^m \\ & \leq C (1 + \nu^{-m(9k+4)} (e^{-\varkappa \nu t_0} \mathbb{E} \exp(\varkappa \nu |u_0|_2^2) + K) + \nu^{-m(7k+2)}), \end{aligned} \quad (2.139)$$

where  $K = K(\nu, \varkappa, \mathfrak{B}, h)$  is defined by (2.126), and  $C > 0$  is a constant depending on  $k, m, T$ , and  $\varkappa$ .

*Proof.* Let us denote by  $P_{km}(t_0, T)$  the left-hand side of (2.139). Applying Proposition 2.4.12 in which the interval  $[0, T]$  is replaced by  $[t_0 - 1, t_0 + T]$ , we conclude that

$$P_{km}(t_0, T) \leq C_1 (1 + \nu^{-m(7k+2)} (\mathbb{E}|u(t_0 - 1)|_2^{4m(k+1)} + 1)). \quad (2.140)$$

Now note that

$$\nu^{-m(7k+2)} \mathbb{E}|u(t_0 - 1)|_2^{4m(k+1)} \leq C_2 \nu^{-m(9k+4)} \mathbb{E} \exp(\varkappa \nu |u(t_0 - 1)|_2^2),$$

where  $C_2$  depends only on  $k, m$ , and  $\varkappa$ . Using Proposition 2.4.9 to bound the quantity  $\mathbb{E} \exp(\varkappa \nu |u(t_0 - 1)|_2^2)$  and substituting the resulting estimate into (2.140), we arrive at (2.139).  $\square$

We now state an exercise summarising some further properties of solutions for the stochastic Navier–Stokes system on the torus. They can be established with the help of the methods developed above.

*Exercise 2.4.14.* Let us consider the stochastic Navier–Stokes system (2.11) in which  $h \in H$  is a deterministic function and  $\{e_j\}$  is the family of normalised eigenfunctions of the Stokes operator  $L$ .

(i) Assume that  $\mathfrak{B}_1 < \infty$  (see (2.137)) and the initial function  $u_0$  is such that

$$\mathbb{E} \exp(\varkappa \nu \|u_0\|_1^2) < \infty,$$

where  $\varkappa > 0$  satisfies inequality (2.123) with a sufficiently small  $c > 0$ . Show that

$$\mathbb{E} \exp(\varkappa \nu \|u(t)\|_1^2) \leq e^{-\varkappa \nu^2 t} \mathbb{E} \exp(\varkappa \nu \|u_0\|_1^2) + \nu^{-1} R_1 e^{\varkappa R_1 / \alpha_1}, \quad (2.141)$$

where  $R_1 = C\nu^{-1}|h|_2^2 + \mathfrak{B}_1 + \nu$ , and  $C > 0$  is a constant not depending on other parameters. *Hint:* Repeat the argument used in the proof of Proposition 2.4.9.

(ii) In addition to the above conditions, assume that

$$\sum_{j=1}^{\infty} e^{\rho \sqrt{\alpha_j}} (\langle h, e_j \rangle^2 + b_j^2) < \infty, \quad (2.142)$$

where  $\rho > 0$  is a constant. Prove that there are positive constants  $p$  and  $q$  such that, for any any integer  $m \geq 1$  and any  $t_0 \geq 1$  and  $T \geq 1$ , we have

$$\mathbb{E} \sup_{t_0 \leq t \leq t_0 + T} \left( \sum_{j=1}^{\infty} \exp(\sqrt{\alpha_j} r_\nu \nu^p) \langle u(t), e_j \rangle^2 \right)^m \leq C_m \nu^{-q}, \quad \mathbb{E} r_\nu^{-m} \leq C_m,$$

where  $C_m > 0$  is a constant not depending on  $\nu$ , and  $r_\nu$  is a random variable not depending on  $p$ ,  $q$ , and  $m$ . *Hint:* A proof can be found in [Shi02].

### 2.4.3 Universality of white noise forces

Theorem 2.2.2 and Exercise 2.2.4 show that a spatially regular white noise force is universal in the sense that an appropriately normalised high-frequency random kick force converges to it in distribution. In this section, we shall prove that this property is inherited by solutions of the corresponding equations. Namely, together with Eq. (2.98), let us consider the Navier–Stokes system perturbed by the high-frequency kick force (2.69):

$$\dot{u} + \nu Lu + B(u) = h + \eta_\varepsilon, \quad \eta_\varepsilon(t, x) = \sqrt{\varepsilon} \sum_{k=1}^{\infty} \eta_k(x) \delta(t - k\varepsilon), \quad (2.143)$$

where  $\{\eta_k\}$  is a sequence of i.i.d. random variables of the form (2.70) and  $0 < \varepsilon \leq 1$ . As was explained in Section 2.3, Eq. (2.143) is equivalent to the discrete-time system

$$u_k = S_\varepsilon(u_{k-1}) + \sqrt{\varepsilon} \eta_k, \quad k \geq 1, \quad (2.144)$$

where  $S_t$  denotes the time shift along trajectories of Eq. (2.143) with  $\eta_\varepsilon \equiv 0$ . With a slight abuse<sup>9</sup> of notation, we denote by  $P_t^\varepsilon(u, \cdot)$ ,  $t \in \varepsilon\mathbb{Z}_+$ , the transition function associated with (2.144) and by  $\mathfrak{P}_t(\varepsilon)$  and  $\mathfrak{P}_t^*(\varepsilon)$  the corresponding Markov operators. We extend the transition function  $P_t^\varepsilon(u, \cdot)$  (as well as the corresponding Markov semigroups) to the half-line  $\mathbb{R}_+$  by defining it as the law of the solution for (2.143) issued from  $u \in H$ .

**Theorem 2.4.15.** *In addition to the above hypotheses, assume that  $\mathfrak{B}_1 < \infty$ . Then for any  $T > 0$  and  $R > 0$  we have*

$$\sup_{0 \leq t \leq T} \sup_{\|v\|_2 \leq R} \|P_t^\varepsilon(v, \cdot) - P_t(v, \cdot)\|_L^* \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0^+. \quad (2.145)$$

*Proof.* We need to prove that if  $u(t; v)$  and  $u_\varepsilon(t; v)$  are solutions issued from  $v \in H$  for the Navier–Stokes system perturbed by a high-frequency kick force or a spatially regular white noise, then for any sequence  $\{\varepsilon_n > 0\}$  going to zero we have

$$\sup_{t, v, f} |\mathbb{E}(f(u_{\varepsilon_n}(t; v)) - f(u(t; v)))| \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad (2.146)$$

where the supremum is taken over  $t \in [0, T]$ ,  $v \in B_H(R)$ , and  $f \in L_b(H)$  with  $\|f\|_L \leq 1$ . This is done in two steps.

*Step 1.* We first note that it suffices to prove convergence (2.146) in which  $u_\varepsilon$  is replaced by the solution  $\tilde{u}_\varepsilon$  of Eq. (4.110) with  $\eta = \partial_t \tilde{\zeta}_\varepsilon$ , where  $\tilde{\zeta}_\varepsilon$  is the continuous process defined in Section 2.2. Indeed, it follows from the condition  $\mathfrak{B}_1 < \infty$  that, with probability 1, the random variables  $\eta_k$  belong to a bounded ball in  $V$ . Therefore the definition of  $\tilde{\zeta}_\varepsilon$  implies that (cf. (2.74))

$$\|\tilde{\zeta}_\varepsilon - \zeta_\varepsilon\|_{L^\infty(0, T; V)} \leq C_1 \sqrt{\varepsilon}.$$

Recalling Exercise 2.1.27, we see that, on the set  $\Gamma_r = \{\|\zeta_\varepsilon\|_{L^\infty(0, T; V)} \leq r\}$ , we have the estimate

$$\sup_{0 \leq t \leq T} \|u_\varepsilon(t; v) - u(t; v)\|_1 \leq C_2(r) \sqrt{\varepsilon}.$$

On the other hand, applying the Doob–Kolmogorov inequality (7.55) to  $\zeta_\varepsilon$ , we derive

$$\mathbb{P}(\Gamma_r^c) = \mathbb{P}\left\{ \sup_{0 \leq t \leq T} \|\zeta_\varepsilon(t)\|_1 > r \right\} \leq r^{-2} \mathbb{E} \|\zeta_\varepsilon(T)\|_1^2 \leq C_3(T) r^{-2}.$$

Combining these estimates, for any  $t \in [0, T]$ ,  $v \in B_H(R)$ , and  $f \in L_b(H)$  with  $\|f\|_L \leq 1$ , we obtain

$$\begin{aligned} |\mathbb{E}(f(u_{\varepsilon_n}(t, v)) - f(u(t, v)))| &\leq \mathbb{E}(\mathbb{I}_{\Gamma_r} \|u_\varepsilon(t; v) - u(t; v)\|_1) + 2\mathbb{P}(\Gamma_r^c) \\ &\leq C_2(r) \sqrt{\varepsilon} + 2C_3(T) r^{-2}. \end{aligned}$$

<sup>9</sup>In Section 2.3, the discrete time varies in  $\mathbb{Z}_+$  independently of the length  $T$  of the interval between two consecutive kicks. Since we wish to compare solutions for systems perturbed by white noise and random kicks, we have to take here the same scale of time.

The right-hand side of this inequality can be made arbitrarily small by choosing  $R \gg 1$  and  $\varepsilon \ll 1$ .

*Step 2.* We now prove (2.146) with  $u_{\varepsilon_n}$  replaced by  $\tilde{u}_{\varepsilon_n}$ . To this end, note that the expression in (2.146) depends only on the laws of  $u_{\varepsilon_n}$  and  $u$ . In other words, if we replace  $\tilde{\zeta}_{\varepsilon_n}$  and  $\zeta$  by some other random variables  $\xi_n$  and  $\xi$  valued in  $C(0, T; V)$  such that

$$\mathcal{D}(\xi_n) = \mathcal{D}(\tilde{\zeta}_{\varepsilon_n}), \quad \mathcal{D}(\xi) = \mathcal{D}(\zeta), \quad (2.147)$$

then the corresponding solutions  $w_n(t; v)$  and  $w(t; v)$  of Eq. (4.110) will satisfy the relation

$$\mathbb{E}(f(\tilde{u}_{\varepsilon_n}(t; v)) - f(u(t; v))) = \mathbb{E}(f(w_n(t; v)) - f(w(t; v))), \quad t \in [0, T].$$

In view of Exercise 2.2.4, we have  $\mathcal{D}(\tilde{\zeta}_{\varepsilon_n}) \rightarrow \mathcal{D}(\zeta)$  in the weak topology of the space  $\mathcal{P}(C(0, T; V))$ . Therefore, by Skorohod's embedding theorem (see Theorem 11.7.2 in [Dud02]), there are random variables  $\xi_n$  and  $\xi$  valued in  $C(0, T; V)$  such that (2.147) holds and, for any  $\delta > 0$ , we have

$$\mathbb{P}\{\|\xi_n - \xi\|_{C(0, T; V)} < \delta\} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

On the other hand, the Doob–Kolmogorov inequality (7.55) implies that

$$\mathbb{P}\{\|\xi\|_{C(0, T; V)} > r\} \rightarrow 0 \quad \text{as } r \rightarrow \infty.$$

Combining these two relations with the argument used in Step 1, we easily prove the required convergence. The proof of the theorem is complete.  $\square$

The following exercise shows that convergence of solutions remains true in the case when the initial function is random, and moreover, it is uniform with respect to measures almost entirely concentrated on bounded parts of the phase space.

*Exercise 2.4.16.* Under the hypotheses of Theorem 2.4.15, prove that, for any  $T > 0$  and any subset  $\Lambda \subset \mathcal{P}(H)$  satisfying the condition

$$\sup_{\lambda \in \Lambda} \int_H |u|_2^2 \lambda(du) < \infty,$$

we have

$$\sup_{0 \leq t \leq T} \sup_{\lambda \in \Lambda} \|\mathfrak{P}_t^*(\varepsilon)\lambda - \mathfrak{P}_t^*\lambda\|_L^* \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0^+. \quad (2.148)$$

*Hint:* Use the idea applied in Step 3 of the proof of Theorem 4.3.1.

#### 2.4.4 RDS associated with Navier–Stokes equations

We conclude this section with a construction of a Markov RDS associated with the Navier–Stokes system perturbed by a spatially regular white noise. For any

$b > 0$ , denote by  $(\Omega_0, \mathcal{A}_0, \mathbb{P}_b)$  the probability space associated with the two-sided centred Brownian motion with variance  $b^2$ . Namely,  $\Omega_0$  is the space of continuous functions  $\omega_t : \mathbb{R} \rightarrow \mathbb{R}$  vanishing at  $t = 0$  with the metric of uniform convergence on bounded intervals (see (1.9)),  $\mathcal{A}_0$  is the Borel  $\sigma$ -algebra, and  $\mathbb{P}_b$  is a Gaussian measure on  $(\Omega, \mathcal{A}_0)$  such that, for any  $t_0 < t_1 < \dots < t_n$  and any Borel subsets  $\Gamma_1, \dots, \Gamma_n \subset \mathbb{R}$ , we have

$$\mathbb{P}_b(\{\omega \in \Omega_0 : \omega_{t_1} - \omega_{t_0} \in \Gamma_1, \dots, \omega_{t_n} - \omega_{t_{n-1}} \in \Gamma_n\}) = \prod_{k=1}^n \mathcal{N}_{b(t_k - t_{k-1})}(\Gamma_k),$$

where  $\mathcal{N}_\sigma$  stands for the normal distribution with zero mean value and variance  $\sigma^2$ . Given a random process  $\zeta$  of the form (2.98), we denote

$$\Lambda = \{j \geq 1 : b_j \neq 0\}.$$

Define  $(\Omega, \mathcal{F}, \mathbb{P})$  as the completion of the direct product of the probability spaces  $(\Omega_0, \mathcal{A}_0, \mathbb{P}_{b_j})$  with  $j \in \Lambda$ . Elements of  $\Omega$  are continuous functions  $\omega : \mathbb{R} \rightarrow \mathbb{R}^\Lambda$ , where  $\mathbb{R}^\Lambda$  is endowed with the Tikhonov topology; see formula (1.9) and the discussion after it.

*Exercise 2.4.17.* Let  $\omega^{(j)}$  be the  $j$ th component of  $\omega \in \Omega$ . Show that the series

$$\tilde{\zeta}(t, x) = \sum_{j \in \Lambda} \omega_t^{(j)} e_j(x), \quad t \in \mathbb{R}, \quad (2.149)$$

defines an  $H$ -valued random process with almost surely continuous trajectories. Show also that the law of the restriction of  $\tilde{\zeta}$  to the positive half-line coincides with that of  $\zeta$ .

Let us define a family of shifts on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  by the relation

$$(\theta_t \omega)_s = \omega_{t+s} - \omega_t, \quad t, s \in \mathbb{R}.$$

*Exercise 2.4.18.* Show that  $\theta = \{\theta_t : \Omega \rightarrow \Omega, t \in \mathbb{R}\}$  is a group of measure-preserving transformations on  $(\Omega, \mathcal{F}, \mathbb{P})$ .

Everything is now ready for the construction of an RDS associated with the stochastic Navier–Stokes system (2.98). We first explain the construction for the simpler case of a more regular force. Namely, assume that  $\mathfrak{B}_1 < \infty$ . In this case, the trajectories of  $\tilde{\zeta}$  belong to  $C(\mathbb{R}; V)$  on a set  $\Omega_* \in \mathcal{F}$  of full measure, and solutions of (2.98) with  $\zeta = \tilde{\zeta}$  can be written as  $u = \tilde{\zeta} + v$ , where  $v$  is a solution of the problem

$$\dot{v} + \nu L(v + \tilde{\zeta}) + B(v + \tilde{\zeta}) = h, \quad v(0) = u_0.$$

It is clear that  $\Omega_*$  can be chosen to be invariant with respect to the group  $\{\theta_t\}$ . Let us define  $\varphi_t^\omega u_0 = \tilde{\zeta}^\omega(t) + v^\omega(t)$  for  $\omega \in \Omega_*$  and  $\varphi_t^\omega u_0 = u_0$  for  $\omega \notin \Omega_*$ . The co-cycle property is an obvious consequence of the uniqueness of solutions for the above problem, while the Markov property follows from the fact that  $\varphi_t^{\theta_s \omega} u_0$  depends only on the increments  $\{\omega_{s+r} - \omega_s, r \in [0, t]\}$ .

In the general case, the construction of an RDS associated with (2.98) is similar. However, we need to use the Ornstein–Uhlenbeck process because the regularity of  $\tilde{\zeta}$  is not sufficient. Let us consider the stochastic Stokes equation (2.100) in which  $\eta = \partial_t \tilde{\zeta}$ . Its solution issued from zero can be written as

$$\tilde{z}(t) = \int_0^t e^{-\nu(t-s)L} d\tilde{\zeta}(s), \quad t \geq 0. \quad (2.150)$$

Let  $\tilde{\Omega}$  be the set of those  $\omega \in \Omega$  for which the corresponding trajectory  $\tilde{z}^\omega$  belongs to the space  $\mathcal{X} = C(\mathbb{R}_+; H) \cap L_{\text{loc}}^2(\mathbb{R}_+; V)$ . Then  $\tilde{\Omega} \in \mathcal{F}$ , and by Proposition 2.4.2, we have  $\mathbb{P}(\tilde{\Omega}) = 1$ . It follows that  $\Omega_* = \bigcap_{n \in \mathbb{Z}} \theta_n(\tilde{\Omega})$  is a set of full measure. We now denote

$$\varphi_t^\omega u_0 = \begin{cases} \tilde{z}^\omega(t) + \tilde{v}^\omega(t), & \text{for } \omega \in \Omega_*, \\ u_0, & \text{for } \omega \notin \Omega_*, \end{cases} \quad (2.151)$$

where  $\tilde{v}^\omega$  stands for the solution of (2.111), (2.112) with  $z = \tilde{z}^\omega$ . The proof of the following result on the existence of an RDS associated with the Navier–Stokes system is given in Section 2.6.

**Theorem 2.4.19.** *The family  $\Phi = \{\varphi_t^\omega, t \geq 0, \omega \in \Omega\}$  is a continuous Markov RDS in  $H$  over  $\theta$ . Moreover, for any  $u_0 \in H$  the law of the random variable  $\{\varphi_t u_0, t \geq 0\}$  in the space  $C(\mathbb{R}_+; H)$  coincides with that of the solution for (2.98), (2.114).*

## 2.5 Existence of a stationary distribution

There is a general and simple method for constructing stationary distributions for Markov processes that satisfy a compactness condition. It was introduced by Bogoliouboff and Kryloff in [KB37] and enables one to prove the existence of a stationary measure for a large class of nonlinear stochastic PDE’s. In this section, we first use the method of Bogoliouboff–Kryloff to establish an abstract result and then apply it to the Markov process associated with the 2D Navier–Stokes system with random perturbation.

### 2.5.1 Bogolyubov–Krylov argument

Let  $(\Omega, \mathcal{F})$  be a measurable space with a filtration  $\{\mathcal{F}_t\}$ , let  $X$  be a Polish space, and let  $(u_t, \mathbb{P}_v)$ ,  $v \in X$ , be a family of  $X$ -valued Markov processes on  $(\Omega, \mathcal{F})$  adapted to  $\mathcal{F}_t$ . We assume that the corresponding transition function  $P_t(u, \cdot)$  associated with  $(u_t, \mathbb{P}_v)$  possesses the Feller property and that the trajectories  $u_t(\omega)$  are continuous in time for any  $\omega \in \Omega$ . Denote by  $\mathfrak{P}_t$  and  $\mathfrak{P}_t^*$  the Markov semigroups associated with  $P_t(u, \cdot)$ . Let us fix  $\lambda \in \mathcal{P}(X)$  and define a family of measures  $\bar{\lambda}_t$ ,  $t \in \mathcal{T}_+$ , be the relation

$$\bar{\lambda}_t(\Gamma) = \frac{1}{t} \int_0^t (\mathfrak{P}_s^* \lambda)(\Gamma) ds = \frac{1}{t} \int_0^t \int_X P_s(u, \Gamma) \lambda(du) ds, \quad \Gamma \in \mathcal{B}(X), \quad (2.152)$$

where the integral in  $s \in [0, t]$  should be replaced by the sum over  $s = 0, \dots, t-1$  if the time parameter is discrete. The Feller property implies that  $P_s(u, \Gamma)$  is a measurable function of  $(s, u) \in \mathcal{T}_+ \times X$ , and therefore the integrals in (2.152) are well defined. Recall that a family of measures  $\{\mu_\alpha, \alpha \in \mathcal{A}\} \subset \mathcal{P}(X)$  is said to be *tight* if for any  $\varepsilon > 0$  there is a compact subset  $K \subset X$  such that  $\mu_\alpha(K) \geq 1 - \varepsilon$  for any  $\alpha \in \mathcal{A}$ .

**Theorem 2.5.1.** *Under the above condition, assume that  $\lambda \in \mathcal{P}(X)$  is a measure for which the family  $\{\bar{\lambda}_t, t \geq 0\}$  is tight. Then the Markov family  $(u_t, \mathbb{P}_v)$  has at least one stationary measure.*

*Proof.* In view of the Prokhorov compactness criterion (see Theorem 1.2.14), there is a subsequence  $\{t_n\} \subset \mathcal{T}_+$  going to  $+\infty$  such that  $\{\bar{\lambda}_{t_n}\}$  converges weakly to a measure  $\mu \in \mathcal{P}(X)$ . That is,

$$(f, \bar{\lambda}_{t_n}) \rightarrow (f, \mu) \quad \text{as } n \rightarrow \infty \text{ for any } f \in C_b(X).$$

We claim that  $\mu$  is a stationary measure for  $(u_t, \mathbb{P}_v)$ . Indeed, in view of this convergence, Exercise 1.3.23, and the definition of  $\bar{\lambda}_t$ , for any  $f \in C_b(X)$  and  $r \in \mathcal{T}_+$  we have

$$\begin{aligned} (f, \mathfrak{P}_r^* \mu) &= (\mathfrak{P}_r f, \mu) = \lim_{n \rightarrow \infty} (\mathfrak{P}_r f, \bar{\lambda}_{t_n}) \\ &= \lim_{n \rightarrow \infty} \frac{1}{t_n} \int_0^{t_n} (\mathfrak{P}_r f, \mathfrak{P}_s^* \lambda) ds = \lim_{n \rightarrow \infty} \frac{1}{t_n} \int_0^{t_n} (f, \mathfrak{P}_{s+r}^* \lambda) ds \\ &= \lim_{n \rightarrow \infty} \frac{1}{t_n} \left( \int_0^{t_n} (f, \mathfrak{P}_s^* \lambda) ds + \int_{t_n}^{t_n+r} (f, \mathfrak{P}_s^* \lambda) ds - \int_0^r (f, \mathfrak{P}_s^* \lambda) ds \right) \\ &= (f, \mu). \end{aligned}$$

Since this is true for any  $f \in C_b(X)$ , we see that  $\mathfrak{P}_r^* \mu = \mu$  for each  $r \in \mathcal{T}_+$ . So,  $\mu$  is a stationary measure.  $\square$

*Exercise 2.5.2.* Let  $(u_t, \mathbb{P}_v)$ ,  $v \in H$ , be a family of Markov processes in  $X$  satisfying the above conditions. Suppose that there is an initial point  $v \in X$ , an increasing sequence of compact subsets  $K_m \subset X$  and finite times  $t_m \in \mathcal{T}_+$  such that

$$\sup_{t \geq t_m} \mathbb{P}_v \{u_t \notin K_m\} \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

Show that the hypotheses of Theorem 2.5.1 are satisfied.

## 2.5.2 Application to Navier–Stokes equations

Let us consider the problem of existence of a stationary measure for the kick-forced Navier–Stokes system (2.76).

**Theorem 2.5.3.** *Let  $h \in H$  be a deterministic function and let  $\{\eta_k\}$  be a sequence of i.i.d. random variables in  $H$  such that  $\mathbb{E}|\eta_1|_2 < \infty$ . Then Eq. (2.76)*



possesses at least one stationary measure. Moreover, every stationary measure  $\mu \in \mathcal{P}(H)$  satisfies

$$\int_H |u|_2 \mu(du) < \infty. \quad (2.153)$$

*Proof. Step 1.* To prove the existence of a stationary measure, we shall show that the hypotheses of Exercise 2.5.2 hold for the initial point  $v = 0$ . In view of relation (2.79), it suffices to construct, for any given  $\varepsilon > 0$ , two compact subsets  $K^{(1)}, K^{(2)} \subset H$  such that

$$\mathbb{P}\{S_T(u_{k-1}) \notin K^{(1)}\} \leq \varepsilon, \quad \mathbb{P}\{\eta_k \notin K^{(2)}\} \leq \varepsilon,$$

where  $\{u_k\}$  is the trajectory of (2.79) with  $u_0 = 0$ . Taking then  $K = K^{(1)} + K^{(2)}$ , we see that  $\mathbb{P}\{u_k \notin K\} \leq 2\varepsilon$  for any  $k \geq 1$ .

The existence of a compact set  $K^{(2)}$  with the required property follows immediately from Ulam's theorem on regularity of probability measures on a Polish space; see Section 1.2.1. To construct  $K^{(1)}$ , we note that (see (2.81))

$$\mathbb{E}|u_k|_2 \leq C \quad \text{for all } k \geq 0.$$

Hence, by Chebyshev's inequality, for any  $\varepsilon > 0$  there is  $R_\varepsilon > 0$  such that

$$\mathbb{P}\{|u_{k-1}| > R_\varepsilon\} \leq R_\varepsilon^{-1} \mathbb{E}|u_{k-1}|_2 \leq \varepsilon \quad \text{for } k \geq 1. \quad (2.154)$$

By part (ii) of Proposition 2.1.25, the operator  $S_T$  is continuous from  $H$  to  $V$ . Denoting by  $B_H(r)$  the closed ball in  $H$  of radius  $r$  centred at zero, we conclude that the set  $K^{(1)} := S_T(B_H(R_\varepsilon))$  is compact in  $H$ . Inequality (2.154) implies that the probability of the event  $\{S_T(u_{k-1}) \notin K^{(1)}\}$  does not exceed  $\varepsilon$ .

*Step 2.* We now prove (2.153) for any stationary measure  $\mu \in \mathcal{P}(H)$ . Let us fix a constant  $R > 0$  and consider a function  $f_R : H \rightarrow \mathbb{R}$  defined by

$$f_R(u) = \begin{cases} |u|_2, & |u|_2 \leq R, \\ R, & |u|_2 > R. \end{cases}$$

Then, for any  $k \geq 0$ , we have

$$\int_H f_R(u) \mu(du) = \int_H \int_H P_k(u, dv) f_R(v) \mu(du). \quad (2.155)$$

Let us estimate the right-hand side of (2.155). If  $|u|_2 \leq \rho$ , then by inequality (2.81) with  $m = 1$  we have

$$\int_H P_k(u, dv) f_R(v) \leq \mathbb{E}|u_k|_2 \leq q^k \rho + C, \quad (2.156)$$

where  $\{u_k\}$  is the trajectory of (2.79) with  $u_0 = u$ , and  $C > 0$  and  $q < 1$  are constants not depending on  $u$ ,  $k$ , and  $R$ . Substituting (2.156) into (2.155) and using the inequality  $f_R \leq R$ , for any  $R > 0$  and  $\rho > 0$  we derive

$$\int_H f_R(u) \mu(du) \leq R \mu(H \setminus B_H(\rho)) + q^k \rho + C.$$

Passing to the limit as  $k \rightarrow \infty$  and  $\rho \rightarrow \infty$ , we see that

$$\int_H f_R(u) \mu(du) \leq C.$$

Fatou's lemma implies now the required inequality (2.153).  $\square$

*Exercise 2.5.4.* Show that if the law of the random variables  $\eta_k$  satisfies the hypotheses of part (i) of Proposition 2.3.5 with some integer  $s \geq 0$ , then for any stationary measure  $\mu \in \mathcal{P}(H)$  of (2.76) we have

$$\int_H \|u\|_s^m \mu(du) < \infty \quad \text{for any } m \geq 1.$$

In particular, if (2.88) holds for any  $s > 0$ , then any stationary measure is concentrated on the space  $C^\infty(\mathbb{T}^2; \mathbb{R}^2)$ . Similarly, show that, under the hypotheses of part (ii) of Proposition 2.3.5, any stationary measure  $\mu \in \mathcal{P}(H)$  satisfies the inequality

$$\int_H \exp(\varkappa_s \|u\|_s^{p_s}) \mu(du) < \infty,$$

where  $\varkappa_s$  and  $p_s$  are some positive constants, and  $p_0 = p_1 = 2$ . *Hint:* Use the argument of the proof of inequality (2.153).

Theorem 2.5.3 is valid for the Navier–Stokes equations in a bounded domain, and its proof remains literally the same. Moreover, the results announced in Exercise 2.5.4 are also true in this case, except that  $p_1 < 2$ .

We now turn to the existence of a stationary measure and some estimates for its moments in the case of the Navier–Stokes system with a spatially regular white noise. Since the proofs are based on similar ideas, we shall mostly confine ourselves to formulating corresponding results and outlining their proofs.

**Theorem 2.5.5.** *Under the hypotheses of Theorem 2.4.6, the stochastic Navier–Stokes system (2.98) with an arbitrary  $\nu > 0$  has a stationary measure. Moreover, any stationary measure  $\mu_\nu \in \mathcal{P}(H)$  satisfies the relations*

$$\int_H \left( \nu \|u\|_1^2 + \exp(\varkappa \nu |u|_2^2) \right) \mu_\nu(du) \leq C(\mathfrak{B} + \nu^{-1} |h|_2^2), \quad (2.157)$$

$$\nu \int_H \|u\|_1^2 \mu_\nu(du) = \frac{\mathfrak{B}}{2} + \int_H \langle u, h \rangle \mu_\nu(du), \quad (2.158)$$

where positive constants  $\varkappa$  and  $C$  do not depend on  $\nu$ .

The left-hand side of (2.158) is called the *rate of dissipation of energy*; cf. (2.119).

*Proof.* We first prove the existence of a stationary measure. Let us denote by  $u(t, x)$  the solution of (2.98) issued from  $u_0 = 0$  and by  $\lambda_t$  the law of  $u(t)$  regarded as a random variable in  $H$ . Thus, we have  $\lambda_t = \mathfrak{P}_t^* \delta_0$  for  $t \geq 0$ , where  $\mathfrak{P}_t^* : \mathcal{P}(H) \rightarrow \mathcal{P}(H)$  stands for the Markov semigroup associated with

the equation and  $\delta_0$  denotes the Dirac measure concentrated at zero. By Theorem 2.5.1, the existence of a stationary measure will be established if we show that the family  $\{\bar{\lambda}_t, t \geq 0\}$  is tight, where  $\bar{\lambda}_t$  is defined by relation (2.152) with  $X = H$ . Since the embedding  $V \subset H$  is compact (see Property 1.1.2), it suffices to prove that

$$\sup_{t \geq 0} \bar{\lambda}_t(H \setminus B_V(R)) \rightarrow 0 \quad \text{as } R \rightarrow \infty, \quad (2.159)$$

where  $B_V(R)$  stands for the ball in  $V$  of radius  $R$  centred at zero. To this end, let us note that relation (2.119) with  $u_0 = 0$  implies that

$$\mathbb{E} \int_0^t |\nabla u(s)|_2^2 ds \leq t(\mathfrak{B} + \nu^{-1}|h|_2^2), \quad t \geq 0.$$

Combining this with Chebyshev's inequality, we derive

$$\begin{aligned} \bar{\lambda}_t(H \setminus B_V(R)) &= \frac{1}{t} \int_0^t \lambda_s(H \setminus B_V(R)) ds = \frac{1}{t} \int_0^t \mathbb{P}\{\|u(s)\|_1 > R\} ds \\ &\leq \frac{1}{t} R^{-2} \mathbb{E} \int_0^t |\nabla u(s)|_2^2 ds \leq R^{-2}(\mathfrak{B} + \nu^{-1}|h|_2^2), \end{aligned}$$

whence follows convergence (2.159).

To prove that any stationary measure satisfies (2.157) and (2.158), we first note that inequality (2.125) and the argument used in derivation of (2.153) enable one to prove that

$$\int_H \exp(\nu|u|_2^2) \mu_\nu(du) \leq C_1(\mathfrak{B} + \nu^{-1}|h|_2^2).$$

Let  $u_0$  be an  $H$ -valued random variable independent of  $\zeta$  whose law coincides with  $\mu$  and let  $u(t)$  be a solution of (2.98) issued from  $u_0$ . Then  $u(t)$  is a stationary process, and it follows from (2.119) that

$$2\nu t \mathbb{E}\|u\|_1^2 = \mathfrak{B}t + 2t \mathbb{E}\langle u, h \rangle.$$

Dividing the above relation by  $2t$ , we obtain (2.158). Combining (2.158) with Schwarz's and Friedrichs' inequalities, we derive the required upper bound for the first term on the left-hand side of (2.157). We leave the details to the reader as an exercise.  $\square$

*Exercise 2.5.6.* Under the hypotheses of Theorem 2.5.5, show that any stationary solution  $u(t, x)$  of (2.98) satisfies the inequality

$$\begin{aligned} \mathbb{E} \sup_{t_0 \leq t \leq t_0+T} \left( |u(t)|_2^2 + \nu \int_{t_0}^t \|u(s)\|_1^2 ds \right)^m \\ \leq C_m (T^m \mathfrak{B}^m + \nu^{-m} (\gamma^{-m} + \alpha_1^{-m} T^m |h|_2^{2m})), \end{aligned}$$

for any constants  $t_0 \geq 0$ ,  $T \geq 1$ ,  $\nu > 0$  and any integer  $m \geq 1$ . Here  $C_m$  is a constant depending only on  $m$ , and  $\gamma > 0$  is defined in Proposition 2.4.10. *Hint:* Use inequalities (2.125), (2.135), and (2.157).

The above theorem and exercise are true both for periodic and Dirichlet boundary conditions. The following exercise gives some extra a priori estimates for stationary measures for the Navier–Stokes system on the torus.

*Exercise 2.5.7.* In addition to the hypotheses of Theorem 2.5.5, assume that  $\mathfrak{B}_1 < \infty$  and  $h \in V$ . Show that there are positive constants  $\varkappa$  and  $C$  not depending on  $\nu$  such that any stationary measure  $\mu_\nu \in \mathcal{P}(H)$  for (2.98) satisfies the relations

$$\int_H \exp(\varkappa \nu \|u\|_1^2) \mu_\nu(du) \leq C (\mathfrak{B}_1 + \nu^{-1} \|h\|_1^2), \quad (2.160)$$

$$\nu \int_H |Lu|_2^2 \mu_\nu(du) = \frac{\mathfrak{B}_1}{2} + \int_H \langle \nabla u, \nabla h \rangle \mu_\nu(du). \quad (2.161)$$

Combine (2.160) with an analogue of (2.125) and (2.135) for  $u_0 \in V$  to obtain some bounds for  $\mathbb{E} \sup_t \mathcal{E}_u(1, t)^m$ , where  $u(t, x)$  is a stationary solution for (2.98). *Hint:* To prove (2.160), combine (2.141) with the idea of the proof of (2.153).

The left-hand side of (2.161) is called the *rate of dissipation of enstrophy*; cf. (2.158).

In case of higher regularity of the deterministic and random forces, some further bounds for various moments of stationary solutions in terms of the viscosity can be found in [BKL00, Shi02, KS03]. In Chapter 5, we shall need the result announced in the following exercise.

*Exercise 2.5.8.* Let the hypotheses of Proposition 2.4.12 be fulfilled for some integer  $k \geq 1$ . Show that, for any  $m \geq 1$  and  $T > 0$ , there is a constant  $C(k, m, T) > 0$  such that any stationary solution of (2.98) satisfies the inequality

$$\mathbb{E} \sup_{0 \leq t \leq T} \mathcal{E}_u(k, t)^m \leq C(k, m, T) (1 + \nu^{-m(7k+2)}). \quad (2.162)$$

In particular, if  $h \in C^\infty(\mathbb{T}^2; \mathbb{R}^2)$  and (2.137) holds for all  $k \geq 1$ , then any stationary measure is concentrated on the space  $C^\infty(\mathbb{T}^2; \mathbb{R}^2)$ . *Hint:* Use inequality (2.138) and the scheme of the proof of Theorem 2.5.3.

We conclude this chapter by the following result on the uniqueness of a stationary measure for the Navier–Stokes system with a small noise.

*Exercise 2.5.9.* Show that for any  $\nu > 0$  there is  $\varepsilon > 0$  such that if  $|h|_2 + \mathfrak{B} \leq \varepsilon$ , then the RDS  $\Phi$  associated with the stochastic Navier–Stokes system (2.98) satisfies inequality (1.75). In particular,  $\Phi$  has a unique stationary measure  $\mu_\nu$ . Show also that the transition function converges to  $\mu_\nu$  exponentially fast in the dual-Lipschitz distance. *Hint:* Use Propositions 1.3.31 and 2.4.10, as well as the proof of (2.54).

## 2.6 Appendix: some technical proofs

*Proof of Proposition 2.1.25.* We shall confine ourselves to the formal derivation of (2.54) and (2.55). The calculations can be justified with the help of Galerkin approximations.

(i) Let us set  $u_i(t) = S_t(u_{0i}, f_i)$ ,  $i = 1, 2$ . Then the difference  $u = u_1 - u_2$  satisfies the equation

$$\dot{u} + Lu + B(u, u_1) + B(u_2, u) = f_1 - f_2. \quad (2.163)$$

Let us take the scalar product of this equation with  $2u(t)$  in  $H$ . Using (2.11) and the inequalities

$$\begin{aligned} |(B(u, u_1), u)| &\leq c_1 \|u\|_{1/2}^2 \|u_1\|_1 \leq c_2 |u|_2 \|u\|_1 \|u_1\|_1 \leq \frac{1}{4} \|u\|_1^2 + c_2^2 \|u_1\|_1^2 |u|_2^2, \\ |(f_1 - f_2, u)| &\leq \|f_1 - f_2\|_{-1} \|u\|_1 \leq \frac{1}{4} \|u\|_1^2 + \|f_1 - f_2\|_{-1}^2, \end{aligned}$$

we derive from (2.163) the differential inequality

$$\partial_t \left( |u|_2^2 + \int_0^t \|u\|_1^2 ds \right) \leq 2c_2^2 \|u_1\|_1^2 \left( |u|_2^2 + \int_0^t \|u\|_1^2 ds \right) + 2\|f_1 - f_2\|_{-1}^2. \quad (2.164)$$

Applying the Gronwall inequality, we arrive at

$$\begin{aligned} |u(t)|_2^2 + \int_0^t \|u\|_1^2 ds &\leq \exp\left(2c_2^2 \int_0^t \|u_1\|_1^2 dr\right) |u_{01} - u_{02}|_2^2 \\ &\quad + 2 \int_0^t \exp\left(2c_2^2 \int_s^t \|u_1\|_1^2 dr\right) \|f_1(s) - f_2(s)\|_{-1}^2 ds. \end{aligned} \quad (2.165)$$

This implies, in particular, inequality (2.54).

(ii) We now take the scalar product of (2.163) with  $2tLu$  in  $H$ :

$$\partial_t (t|u|_1^2) + 2t|Lu|_2^2 = \|u\|_1^2 - 2t(B(u, u_1), Lu) - 2t(B(u_2, u), Lu) + 2t(f_1 - f_2, Lu). \quad (2.166)$$

Let us use the inequalities

$$\|v\|_\infty^2 \leq c_3 |v|_2 |Lv|_2, \quad \|v\|_1^2 \leq |v|_2 |Lv|_2$$

to estimate the second and third terms on the right-hand side of (2.166):

$$\begin{aligned} |(B(u, u_1), Lu)| &\leq c_4 \|u\|_\infty \|u_1\|_1 |Lu|_2 \leq c_5 |u|_2^{1/2} |Lu|_2^{3/2} |u_1|_2^{1/2} |Lu|_2^{1/2} \\ &\leq \frac{1}{8} |Lu|_2^2 + c_6 |u|_2^2 |u_1|_2^2 |Lu|_2^2, \end{aligned} \quad (2.167)$$

$$\begin{aligned} |(B(u_2, u), Lu)| &\leq c_4 \|u_2\|_\infty \|u\|_1 |Lu|_2 \leq c_5 |u_2|_2^{1/2} |Lu|_2^{1/2} |u|_2^{1/2} |Lu|_2^{3/2} \\ &\leq \frac{1}{8} |Lu|_2^2 + c_6 |u|_2^2 |u_2|_2^2 |Lu|_2^2. \end{aligned} \quad (2.168)$$

Furthermore, by the Cauchy inequality,

$$|(f_1 - f_2, Lu)| \leq \frac{1}{4} |Lu|_2^2 + |f_1 - f_2|_2^2. \quad (2.169)$$

Substituting (2.167) – (2.169) into (2.166) and integrating in time, we derive

$$\begin{aligned} t\|u\|_1^2 + \int_0^t s|Lu|_2^2 ds &\leq \int_0^t \|u\|_1^2 ds \\ &+ c_7 \int_0^t s|u|_2^2(|u_1|_2^2|Lu_1|_2^2 + |u_2|_2^2|Lu_2|_2^2) ds + 2 \int_0^t s|f_1 - f_2|_2^2 ds. \end{aligned}$$

Combining this with (2.165), we see that the required inequality will be established if we show that

$$\int_0^t s|u_i|_2^2|Lu_i|_2^2 ds \leq c_8 t^{-2} \exp\left(c_8 \int_0^t (\|u_i\|_1^2 + |f_i|_2^2) ds\right), \quad (2.170)$$

where  $i = 1, 2$  and  $0 < t \leq 1$ .

To prove (2.170), we first note that, by (2.43),

$$\int_0^t s|Lu_i|_2^2 ds \leq |u_{0i}|_2^2 + 2 \int_0^t |f|_2^2 ds \quad \text{for } 0 \leq t \leq 1. \quad (2.171)$$

Furthermore, taking the scalar product of (2.19) with  $2u$  and integrating in time, we derive

$$|u_i(t)|_2^2 + 2 \int_0^t \|u_i\|_1^2 ds = |u_{0i}|_2^2 + 2 \int_0^t \langle u_i, f_i \rangle ds.$$

Combining this with (2.51), we conclude that

$$(1 - e^{-\alpha_1 t})|u_{0i}|_2^2 \leq c_9 \int_0^t (\|u_i\|_1^2 + \|f_i\|_{-1}^2) ds. \quad (2.172)$$

It follows from (2.51), (2.171), and (2.172) that

$$\begin{aligned} \int_0^t s|u_i|_2^2|Lu_i|_2^2 ds &\leq c_{10} \left( |u_{0i}|_2^2 + \int_0^t |f_i|_2^2 ds \right)^2 \\ &\leq c_{11} \left( (1 - e^{-\alpha_1 t})^{-1} \int_0^t (\|u_i\|_1^2 + \|f_i\|_{-1}^2) ds + \int_0^t |f_i|_2^2 ds \right)^2. \end{aligned}$$

This implies the required inequality (2.170), and the proof of the proposition is complete.  $\square$

*Proof of Proposition 2.4.12.* The proof is by induction in  $k$ . For  $k = 0$ , a stronger result is established in Corollary 2.4.11. We now assume that  $k = n \geq 1$  and that for  $k \leq n - 1$  inequality (2.138) is already proved. We wish to apply Itô's formula to the functional

$$F_n(t, u) = t^n \|u\|_n^2 = \langle L^n u, u \rangle.$$

To this end, we use Theorem 7.7.5 with  $H = V^n$ ,  $V = V^{n+1}$ , and  $V^* = V^{n-1}$ . The fact that the hypotheses of Theorem 7.7.5 are fulfilled can be checked with

the help of an argument similar to that used in the proof of Propositions 2.4.8 and 2.4.10 (which correspond to the case  $n = 0$ ). Therefore we shall confine ourselves to a formal derivation of estimates.

Let us set  $f_n(t) = F_n(t, u(t))$ . By Itô's formula (7.37), we have (cf. (2.131))

$$\begin{aligned} f_n(t) = & \int_0^t (ns^{n-1}\|u\|_n^2 + 2s^n \langle L^n u, h - \nu Lu - B(u) \rangle + \mathfrak{B}_n s^n) ds \\ & + \sum_{j=1}^{\infty} 2b_j \int_0^t s^n \langle L^n u, e_j \rangle d\beta_j(s). \end{aligned} \quad (2.173)$$

We now distinguish between two cases:  $n = 1$  and  $n \geq 2$ . In the first case, by Lemma 2.1.16, we have  $\langle Lu, B(u) \rangle = 0$ . Combining this with (2.173) and the relations

$$2s |\langle Lu, h \rangle| \leq \|u\|_1^2 + s^2 \|h\|_1^2, \quad \langle Lu, e_j \rangle = \alpha_j \langle u, e_j \rangle,$$

we derive

$$\mathcal{E}_u(1, t) \leq 2 \int_0^t \|u\|_1^2 ds + \frac{1}{2} \mathfrak{B}_1 t^2 + \frac{1}{3} \|h\|_1^2 t^3 + K_1(t),$$

where we set

$$K_1(t) = \sum_{j=1}^{\infty} 2b_j \alpha_j \int_0^t s \langle u, e_j \rangle d\beta_j(s) - \nu \int_0^t s \|u\|_2^2 ds. \quad (2.174)$$

It follows that

$$\sup_{0 \leq t \leq T} \mathcal{E}_u(1, t) \leq 2 \int_0^T \|u\|_1^2 ds + \mathfrak{B}_1 T^2 + \|h\|_1^2 T^3 + \sup_{0 \leq t \leq T} K_1(t). \quad (2.175)$$

The first term on the right-hand side of (2.175) can be estimated with the help of (2.135). To bound the last term, we repeat the argument used in the proof of (2.130). Namely, we denote by  $M(t)$  the martingale defined by the sum in (2.174) and remark that its quadratic variation is equal to

$$\langle M \rangle_t = \sum_{j=1}^{\infty} 4b_j^2 \alpha_j^2 \int_0^t s^2 \langle u, e_j \rangle^2 ds \leq 4T \left( \sup_{j \geq 1} b_j^2 \right) \int_0^t s \|u(s)\|_2^2 ds.$$

Therefore, setting  $\gamma = (2T \sup_j b_j^2)^{-1}$ , we obtain

$$K_1(t) \leq M(t) - \frac{\gamma \nu}{2} \langle M \rangle_t.$$

By the supermartingale inequality (7.57), we have

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq T} K_1(t) \geq \rho \right\} = \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \exp(\gamma \nu M(t) - \frac{(\gamma \nu)^2}{2} \langle M \rangle_t) \geq e^{\gamma \nu \rho} \right\} \leq e^{-\gamma \nu \rho},$$

whence it follows that (cf. proof of Corollary 2.4.11)

$$\begin{aligned} \mathbb{E} \sup_{0 \leq t \leq T} \mathcal{E}_u(1, t)^m &\leq C_1 (\mathfrak{B}_1^m T^{2m} + \|h\|_1^{2m} T^{3m}) \\ &\quad + C_1 \left( \nu^{-m} \mathbb{E} \mathcal{E}_u(T)^m + \int_0^\infty e^{-\nu\gamma\lambda^{1/m}} d\lambda \right). \end{aligned} \quad (2.176)$$

Combining this with (2.135), we obtain the inequality

$$\begin{aligned} \mathbb{E} \sup_{0 \leq t \leq T} \mathcal{E}_u(1, t)^m &\leq C_2 (\mathfrak{B}_1^m T^{2m} + \|h\|_1^{2m} T^{3m}) \\ &\quad + C_2 \nu^{-m} (1 + \mathbb{E} |u_0|_2^{2m} + T^m \mathfrak{B}^m) + C_2 \nu^{-2m} (1 + T^m \alpha_1^{-m} |h|_2^{2m}), \end{aligned} \quad (2.177)$$

which gives the required estimate (2.138) with  $k = 1$ .

We now assume that  $n \geq 2$ . In this case, the scalar product  $\langle L^n u, B(u) \rangle$  does not vanish, and we need to estimate it. In view of Lemma 2.1.20, we have

$$\begin{aligned} |\langle L^n u, B(u) \rangle| &\leq C_2 \|u\|_{n+1}^{(4n-1)/2n} \|u\|_1^{(n+1)/2n} |u|_2^{1/2} \\ &\leq \frac{\nu}{4} \|u\|_{n+1}^2 + C_2^2 \nu^{-4n} \|u\|_1^{2(n+1)} |u|_2^{2n}. \end{aligned}$$

Combining this with (2.173) and the relations

$$2s |\langle L^n u, h \rangle| \leq s^{n-1} \|u\|_n^2 + s^{n+1} \|h\|_n^2, \quad \langle L^n u, e_j \rangle = \alpha_j^n \langle u, e_j \rangle,$$

we derive

$$\begin{aligned} \mathcal{E}_u(n, t) &\leq (n+1) \int_0^t s^{n-1} \|u\|_n^2 ds + \frac{1}{n+1} \mathfrak{B}_n t^{n+1} + \frac{1}{n+2} \|h\|_n^2 t^{n+2} \\ &\quad + C_3 \nu^{-4n} \int_0^t s^n \|u\|_1^{2(n+1)} |u|_2^{2n} ds + K_n(t), \end{aligned}$$

where we set

$$K_n(t) = \sum_{j=1}^{\infty} 2b_j \alpha_j^n \int_0^t s^n \langle u, e_j \rangle d\beta_j(s) - \frac{\nu}{2} \int_0^t s^n \|u\|_{n+1}^2 ds.$$

Since

$$\sup_{0 \leq t \leq T} \int_0^t s^n \|u\|_1^{2(n+1)} |u|_2^{2n} ds \leq \nu^{-1} \sup_{0 \leq t \leq T} \mathcal{E}_u(t)^{n+1} \mathcal{E}_u(1, t)^n,$$

we conclude that

$$\begin{aligned} \sup_{0 \leq t \leq T} \mathcal{E}_u(n, t) &\leq C_4 (\nu^{-1} \mathcal{E}_u(n-1, T) + \mathfrak{B}_n T^{n+1} + \|h\|_n^2 T^{n+2}) \\ &\quad + C_4 \nu^{-(4n+1)} \sup_{0 \leq t \leq T} (\mathcal{E}_u(t)^{n+1} \mathcal{E}_u(1, t)^n) + \sup_{0 \leq t \leq T} K_n(t). \end{aligned}$$



Setting  $P(m, n) = \mathbb{E} \sup \mathcal{E}_u(n, t)^m$  and repeating the argument used in the derivation of (2.176), we prove that

$$P(m, n) \leq C_5 \nu^{-m} (P(m, n-1) + \nu^{-4mn} (P(2m(n+1), 0) P(2mn, 1))^{1/2}) + C_5 (1 + \nu^{-m}). \quad (2.178)$$

Recalling that (see (2.135) and (2.177))

$$P(2m(n+1), 0) \leq C_6 (\mathbb{E} |u_0|_2^{4m(n+1)} + \nu^{-2m(n+1)} + 1), \\ P(2mn, 1) \leq C_7 (1 + \nu^{-2mn} \mathbb{E} |u_0|_2^{4mn} + \nu^{-4mn}),$$

and using the induction hypothesis, we see that (2.178) implies the required estimate (2.138) with  $k = n$ . This completes the proof of the proposition.  $\square$

*Proof of Theorem 2.4.19. Step 1.* Suppose we have shown that

$$\theta_s(\Omega_*) = \Omega_* \quad \text{for any } s \in \mathbb{R}. \quad (2.179)$$

To prove that  $\Phi$  is an RDS, we need to check the continuity and cocycle properties of Definition 1.3.14. The continuity of  $\varphi_t^\omega(\cdot)$  is trivial for  $\omega \notin \Omega_*$ . To prove it for  $\omega \in \Omega_*$ , let us note that, in view of (2.151), we have

$$\varphi_t^\omega u_0 = \tilde{z}^\omega(t) + \mathcal{R}_t(u_0, \tilde{z}^\omega(\cdot)) \quad \text{for } t \geq 0, \omega \in \Omega_*, \quad (2.180)$$

where  $\mathcal{R}_t : H \times \mathcal{X} \rightarrow H$  stands for the operator that takes  $(u_0, z)$  to the solution at time  $t$  of problem (2.111), (2.112). Proposition 2.4.5 now implies the required continuity  $\omega \in \Omega_*$ .

Let us prove the cocycle property (1.62). The first relation is obvious, as is the second one with  $\omega \notin \Omega_*$  (in view of (2.179)). We need to show that

$$\varphi_{t+s}^\omega u_0 = \varphi_t^{\theta_s \omega} (\varphi_s^\omega u_0) \quad \text{for } t, s \geq 0, \omega \in \Omega_*, u_0 \in H. \quad (2.181)$$

To this end, note that, in view of (2.150) and (2.102), we have

$$\tilde{z}^{\theta_s \omega}(t) = \tilde{z}_s^\omega(s+t) \quad \text{for } t \geq 0, s \in \mathbb{R}, \omega \in \Omega_*, \quad (2.182)$$

where  $\tilde{z}_s^\omega(t)$  is the solution of the problem

$$\dot{z} + \nu Lz = \partial_t \tilde{\zeta}, \quad z(s) = 0$$

Combining (2.180) and (2.182), we write

$$\varphi_t^{\theta_s \omega} (\varphi_s^\omega u_0) = \tilde{z}_s^\omega(t+s) + \mathcal{R}_t(\varphi_s^\omega u_0, \tilde{z}_s^\omega(s, \cdot)). \quad (2.183)$$

For any  $s \in \mathbb{R}$  and  $\omega \in \Omega_*$ , the right-hand side of (2.183) regarded as a function of  $t$  is a solution of the problem

$$\dot{u} + \nu Lu + B(u) = h + \partial_t \tilde{\zeta}^\omega, \quad u(s) = \varphi_s^\omega u_0. \quad (2.184)$$

On the other hand, the function  $\varphi_{+s}^\omega u_0$  also satisfies (2.184). By uniqueness, the two functions must coincide, and we obtain (2.181).

*Step 2.* Let us prove (2.179). To this end, it suffices to show that  $\theta_s \omega \in \tilde{\Omega}$  for any  $s \in \mathbb{R}$ . Let  $\omega \in \Omega_*$  and let  $n \in \mathbb{Z}$  be such that  $n + s \geq 0$ . In view of (2.182), the difference

$$w(t) = \tilde{z}^{\theta-n\omega}(s+n+t) - \tilde{z}^{\theta_s\omega}(t) = \tilde{z}^{\theta-n\omega}(s+n+t) - \tilde{z}_s^\omega(s+t)$$

regarded as a function of  $t$  satisfies the equations

$$\dot{w} + \nu Lw = 0, \quad w(0) = \tilde{z}^\omega(s+n) \in H.$$

It follows that  $w \in \mathcal{X}$ . Since  $\tilde{z}^{\theta-n\omega}(s+n+\cdot) \in \mathcal{X}$ , we conclude that  $\tilde{z}^{\theta_s\omega} = \tilde{z}^\omega(s+n+\cdot) - w \in \mathcal{X}$  and, hence,  $\theta_s \omega \in \tilde{\Omega}$  for any  $s \in \mathbb{R}$ .

*Step 3.* We now prove that  $\Phi$  is Markov. In view of Definition 1.3.18, we need to show for any finite sets  $u_i^\pm \in H$ ,  $t_i^\pm, s_i^\pm \in \mathbb{R}$ ,  $i = 1, \dots, N^\pm$ , such that

$$s_i^- < 0, \quad 0 < t_i^- \leq -s_i^-, \quad s_i^+ \geq 0, \quad t_i^+ > 0,$$

and any Borel subsets  $\Gamma_i^\pm \subset H$ , the events<sup>10</sup>

$$A^\pm = \{\omega \in \Omega : \varphi(t_i^\pm, \theta_{s_i^\pm} \omega) u_i^\pm \in \Gamma_i^\pm, i = 1, \dots, N^\pm\}$$

are independent. To this end, note that (cf. (2.180) and (2.102))

$$\begin{aligned} \varphi_t^{\theta_s\omega} u_0 &= \tilde{z}_s^\omega(s+t) + \mathcal{R}_t(u_0, \tilde{z}_s^\omega(s+\cdot)), \\ \tilde{z}_s^\omega(s+t) &= \tilde{\zeta}^\omega(s+t) - e^{-\nu t L} \tilde{\zeta}^\omega(s) - \nu \int_s^{s+t} L e^{-\nu(t-s)L} \tilde{\zeta}^\omega(s) ds. \end{aligned}$$

where  $s \in \mathbb{R}$ ,  $t \geq 0$ , and  $\omega \in \Omega_*$ . It follows that, up to an event of zero measure,  $A^-$  depends only on the path  $\{\omega_s, s \leq 0\}$ , whereas  $A^+$  depends only on  $\{\omega_s, s \geq 0\}$ . Since the process  $\omega_t$  has independent increments and vanishes at  $t = 0$ , we conclude that  $A^-$  and  $A^+$  are independent.

Finally, the fact that the law of  $\{\varphi_t u_0, t \geq 0\}$  is the same as that of the solution  $u(t)$  for (2.98), (2.114) follows immediately from (2.151) and the construction of  $u$  (see Theorem 2.4.6). The proof of the theorem is now complete.  $\square$

## Notes and comments

The results presented in Section 2.1 are very well known. The well-posedness of the 2D Navier–Stokes system in a bounded domain and the regularity of solutions was proved due to contributions of many researchers, starting from the pioneering paper of Leray [Ler34], followed by Ladyzhenskaya [Lad59, Lad63], Lions and Prodi [LP59, Lio69], Temam [Tem68] and others. We refer the reader

<sup>10</sup>We write here  $\varphi(t, \theta_s \omega)$  instead of  $\varphi_t^{\theta_s \omega}$  to avoid triple subscripts.

to the books [Tem79, CF88, Soh01] for a comprehensive study of the initial-boundary value problem for Navier–Stokes equations.

The Foias–Prodi estimates presented in Section 2.1.8 are of fundamental importance for this book. They show that the dynamics of the 2D Navier–Stokes system in a bounded domain is determined by finitely many modes. This type of result was first established by Foias and Prodi [FP67]. It is used in various problems related to the large-time behaviour of solutions, including the theory of attractors and inertial manifolds; see the books [Tem88, CFNT89, BV92] and references therein.

The theory of well-posedness for Navier–Stokes equations with various types of additive noises is a rather straightforward consequence of deterministic results, due to the simple reductions described in Sections 2.3 and 2.4. The study of the Navier–Stokes system perturbed by a spatially regular white noise was initiated by Bensoussan, Temam [BT73], Viot [Vio75, Vio76], and Vishik, Komech, Fursikov [VKF79, VF88]. They proved that the 2D Navier–Stokes system perturbed by a white noise force defines a Markov process in a suitable function space and that Itô’s formula applies to its solutions and results in the *a priori* estimates (2.119) and (2.135) with  $m = 1$ . These inequalities with  $m \geq 2$  and other estimates for more regular random forces were derived by E, Mattingly, Sinai [EMS01, Mat02b] and Kuksin, Shirikyan [KS02a, KS03]. The case of an analytic random force was investigated by Bricmont, Kupiainen, Lefevre [BKL00], Mattingly [Mat02a], and Shirikyan [Shi02]. A more delicate task is to prove the existence and uniqueness of a solution for white forces of low spatial smoothness. This situation was studied, for instance, by Flandoli [Fla94], Ferrario [Fer03], Da Prato, Debussche [DD02], and Brzeźniak, Ferrario [BF09]. Navier–Stokes equations perturbed by random kicks are less traditional, but their investigation is technically simpler. The estimates described in Section 2.3 are taken from [KS01b, Shi04]. Theorem 2.4.15 and relation (2.148) belong to a large group of results in stochastic PDE’s motivated by numerical methods, since a popular way to calculate solutions of a PDE perturbed by a white force is to replace it by high-frequency normalised random kicks. This approach is called *splitting up method* in numerical analysis. We refer the reader to the paper of Gyöngy and Krylov [GK03] for some results in this direction that concern, in particular, the rate of convergence.

Finally, the idea of studying limit points of the averaged laws of solutions to construct a stationary distribution of Markov processes goes back to Kryloff and Bogoliouboff [KB37]. It is a simple, but effective tool for producing stationary distributions in various problems; see the books [VF88, KH95] and references therein. The question of uniqueness of a stationary distribution is, in general, much more complicated. However, in the laminar case, the RDS generated by the 2D Navier–Stokes system is globally asymptotically stable, which implies the uniqueness and exponential mixing of a stationary measure (see Exercise 2.5.9). This fact was proved by Mattingly [Mat99].



## Chapter 3

# Uniqueness of stationary measure and mixing

This chapter contains some results on uniqueness of a stationary distribution and the property of exponential mixing. Their proof is based on a development of the classical coupling argument introduced by Doeblin in late thirties; see [Doe38, Doe40]. Without going into details, let us describe two essentially equivalent versions of Doeblin's approach to study ergodic properties of Markov chains.

Let  $X$  be a complete metric space and let  $(u_k, \mathbb{P}_u)$ ,  $k \in \mathbb{Z}_+$ , be a Feller family of Markov chains in  $X$  parametrised by the initial point  $u \in X$ . We shall denote by  $P_k(u, \Gamma)$ ,  $u \in X$ ,  $\Gamma \in \mathcal{B}(X)$ , the transition function associated with the Markov family and by  $\mathfrak{P}_k : C_b(X) \rightarrow C_b(X)$  and  $\mathfrak{P}_k^* : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$  the corresponding Markov operators (see Section 1.3.3). Let us assume that

$$\|P_1(u, \cdot) - P_1(u', \cdot)\|_{\text{var}} \leq \gamma \quad (3.1)$$

for any  $u, u' \in X$ , where  $\|\cdot\|_{\text{var}}$  denotes the total variation distance, and  $\gamma < 1$  is a constant not depending on  $u$  and  $u'$ . In this case, the uniqueness of a stationary distribution and exponential mixing for the family  $(u_k, \mathbb{P}_u)$  can be proved using the coupling lemma in one of the two equivalent forms given in Section 1.2.4.

**Contraction of the space of measures.** Let us endow  $\mathcal{P}(X)$  with the total variation distance and consider the operator  $\mathfrak{P}_1^* : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$ . We claim that this is a contraction:

$$\|\mathfrak{P}_1^* \lambda - \mathfrak{P}_1^* \lambda'\|_{\text{var}} \leq \gamma \|\lambda - \lambda'\|_{\text{var}} \quad \text{for any } \lambda, \lambda' \in \mathcal{P}(X). \quad (3.2)$$

Indeed, let us use Corollary 1.2.25 to write

$$\lambda = (1 - \delta)\nu + \delta\tilde{\lambda}, \quad \lambda' = (1 - \delta)\nu + \delta\tilde{\lambda}',$$

where  $\delta = \|\lambda - \lambda'\|_{\text{var}}$ ,  $\nu \in \mathcal{P}(X)$ , and  $\tilde{\lambda}$  and  $\tilde{\lambda}'$  are mutually singular measures.

Using (3.1) and (1.22), for any  $\Gamma \in \mathcal{B}(X)$  we obtain

$$\begin{aligned} \mathfrak{P}_1^* \lambda(\Gamma) - \mathfrak{P}_1^* \lambda'(\Gamma) &= \delta (\mathfrak{P}_1^* \tilde{\lambda}(\Gamma) - \mathfrak{P}_1^* \tilde{\lambda}'(\Gamma)) \\ &= \delta \iint_{\tilde{X} \times X} (P_1(u, \Gamma) - P_1(u', \Gamma)) \tilde{\lambda}(du) \tilde{\lambda}'(du') \\ &\leq \delta \iint_{\tilde{X} \times X} \gamma \tilde{\lambda}(du) \tilde{\lambda}'(du') \\ &= \|\lambda - \lambda'\|_{\text{var}} \gamma. \end{aligned}$$

Using the symmetry and taking the supremum over  $\Gamma \in \mathcal{B}(X)$ , we arrive at (3.2).

Since  $\mathfrak{P}_1^*$  is a contraction of the complete metric space  $(\mathcal{P}(X), \|\cdot\|_{\text{var}})$ , there is a unique measure  $\mu \in \mathcal{P}(X)$  such that  $\mathfrak{P}_1^* \mu = \mu$ , and for any  $\lambda \in \mathcal{P}(X)$  we have

$$\|\mathfrak{P}_k^* \lambda - \mu\|_{\text{var}} \leq \gamma^k, \quad k \geq 0. \quad (3.3)$$

When (3.3) holds, we say that the Markov process (which defines the semi-group) is *exponentially mixing* in the total variation norm. Similarly, if inequality (3.3) holds with the total variation distance replaced by the dual-Lipschitz distance, then we say that the Markov process is *exponentially mixing* in the dual-Lipschitz norm.

**Coupling argument.** Let  $(\mathcal{R}(u, u', \cdot), \mathcal{R}'(u, u', \cdot))$  be a pair of random variables in  $X$  that depend on  $u, u' \in X$  and form a maximal coupling for the measures  $P_1(u, \cdot)$  and  $P_1(u', \cdot)$ . That is,  $\mathcal{D}(\mathcal{R}) = P_1(u, \cdot)$ ,  $\mathcal{D}(\mathcal{R}') = P_1(u', \cdot)$ , and

$$\mathbb{P}\{\mathcal{R}(u, u') \neq \mathcal{R}'(u, u')\} = \|P_1(u, \cdot) - P_1(u', \cdot)\|_{\text{var}} \quad \text{for all } u, u' \in X; \quad (3.4)$$

cf. Section 1.2.4. In particular,  $\mathcal{R}(u, u) = \mathcal{R}'(u, u)$  for any  $u \in X$ . Such random variables exist in view of Theorem 1.2.28. Let us denote by  $\Omega$  the direct product of countably many copies of the probability space on which  $\mathcal{R}$  and  $\mathcal{R}'$  are defined and consider a family of Markov chains  $\{\mathbf{u}_k\}$  in  $\mathbf{X} = X \times X$  given by the rule

$$\mathbf{u}_0(\omega) = \mathbf{u}, \quad \mathbf{u}_k(\omega) = (\mathcal{R}(\mathbf{u}_{k-1}, \omega_k), \mathcal{R}'(\mathbf{u}_{k-1}, \omega_k)) \quad \text{for } k \geq 1. \quad (3.5)$$

Here  $\omega = (\omega_j, j \geq 1) \in \Omega$  denotes the random parameter and  $\mathbf{u} \in \mathbf{X}$  is an initial point. Writing  $\mathbf{u} = (u, u')$  and  $\mathbf{u}_k = (u_k, u'_k)$ , we derive from inequalities (3.1), (3.4) and the Markov property that

$$\mathbb{P}_{\mathbf{u}}\{u_{k+1} \neq u'_{k+1} \mid \mathcal{F}_k\} \leq \gamma \quad \text{for any } \mathbf{u} \in \mathbf{X}, k \geq 0, \quad (3.6)$$

where  $\mathcal{F}_k$  denotes the  $\sigma$ -algebra generated by  $\mathbf{u}_1, \dots, \mathbf{u}_k$ . Iterating inequality (3.6) and using the fact that  $u_{k-1} = u'_{k-1}$  implies  $u_k = u'_k$ , we obtain by induction

$$\begin{aligned} \mathbb{P}_{\mathbf{u}}\{u_k \neq u'_k\} &= \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\{u_{k-1} \neq u'_{k-1}\}} \mathbb{P}_{\mathbf{u}}\{u_k \neq u'_k \mid \mathcal{F}_{k-1}\}) \\ &\leq \gamma \mathbb{P}_{\mathbf{u}}\{u_{k-1} \neq u'_{k-1}\} \leq \gamma^k \end{aligned} \quad (3.7)$$

for any  $\mathbf{u} \in \mathbf{X}$ ,  $k \geq 0$ . So,

$$\|P_k(u, \cdot) - P_k(u', \cdot)\|_{\text{var}} \leq \gamma^k. \quad (3.8)$$

Combining this with the Kolmogorov–Chapman relation, we see that for any two measures  $\lambda$  and  $\lambda'$  on  $X$  the total variation distance between  $\mathfrak{P}_k^* \lambda$  and  $\mathfrak{P}_k^* \lambda'$  goes to zero exponentially fast. In particular, there is at most one stationary distribution. Moreover, the sequence  $\{P_k(u, \cdot)\}$  converges to a limiting measure  $\mu$ , which is stationary for  $(u_k, \mathbb{P}_u)$ . Finally, inequality (3.3) also follows from (3.8).

Inequality (3.1) is satisfied for a number of problems with compact phase space, e.g., for stochastic differential equations (SDE) with non-degenerate diffusion on a compact manifold. On the other hand, condition (3.1) is rather restrictive if the phase space is not compact. For instance, in the case of SDE's in  $\mathbb{R}^n$ , it is fulfilled only if there is a strong nonlinear drift towards a bounded ball. However, sometimes one can overcome this difficulty with the help of the following modification of the coupling argument.

Let  $X$  be a Polish space and let  $(u_k, \mathbb{P}_u)$  be a Feller family of Markov chains in  $X$ . Retaining the notation used above, suppose we can find a closed subset  $B \subset X$  for which the two properties below are satisfied:

*Recurrence:* The first hitting time  $\tau_B$  of the set  $B$  is almost surely finite for any initial point  $u \in X$ , and there is  $\delta > 0$  such that

$$\mathbb{E}_u \exp(\delta \tau_B) < \infty \quad \text{for all } u \in X. \quad (3.9)$$

*Squeezing:* Inequality (3.1) with a constant  $\gamma < 1$  holds for any  $u, u' \in B$ .

Let  $(\mathcal{R}, \mathcal{R}')$  be the family of random variables in  $\mathbf{X}$  defined above and let  $\{\mathbf{u}_k\}$  be the family of Markov chains given by (3.5). Denote by  $\rho_n$  the  $n^{\text{th}}$  instant when the trajectory  $\mathbf{u}_k$  enters the set  $\mathbf{B} := B \times B$ . Then, using (3.1), (3.4), and the strong Markov property, we get (cf. (3.6))

$$\mathbb{P}\{u_{\rho_{n+1}} \neq u'_{\rho_{n+1}} | \mathcal{F}_{\rho_n}\} \leq \gamma \quad \text{for any } \mathbf{u} \in \mathbf{X}, n \geq 1, \quad (3.10)$$

where  $\mathcal{F}_{\rho_n}$  denotes the  $\sigma$ -algebra associated with the Markov time  $\rho_n$ . Iteration of (3.10) results in (cf. (3.7))

$$\mathbb{P}_{\mathbf{u}}\{u_{\rho_{n+1}} \neq u'_{\rho_{n+1}}\} \leq \gamma^n \quad \text{for any } \mathbf{u} \in \mathbf{X}, n \geq 1.$$

Combining this with (3.9), one can prove inequality (3.8) with a larger constant  $\gamma < 1$ , and this implies all the properties established for the case in which (3.1) holds uniformly. Thus, Doeblin's method applies also in the case of an unbounded phase space, provided that inequality (3.1) is satisfied on a subset that can be reached from any initial point at a random time with finite exponential moment.

Application of the above technique to stochastic PDE's encounters one essential difficulty: inequality (3.1) cannot be true (even for close  $u$  and  $u'$ ), unless

some very restrictive conditions are imposed on the diffusion; see Section 3.5.1 below for the case of Navier–Stokes equations. In this chapter, we show how to develop Doebelin’s approach to be able to treat the 2D stochastic Navier–Stokes equations and other dissipative SPDE’s with degenerate diffusion. In Section 3.1, some general criteria are established for the uniqueness of a stationary distribution and mixing in the dual-Lipshcitz norm. In Sections 3.2 – 3.4 we apply these criteria to the Navier–Stokes system with various random perturbations. Section 3.5 is devoted to some further results on uniqueness and mixing, a discussion of the Navier–Stokes system perturbed by a compound Poisson process, and a description of an alternative proof for mixing of the model with random kicks. Finally, in Section 3.7, we clarify the importance of the results of this chapter for physics.

### 3.1 Three results on uniqueness and mixing

In this section, we establish some sufficient conditions for uniqueness of a stationary measure for a Markov process. We begin with the case in which there is a Kantorovich functional (see Section 1.2.5) decaying along any pair of trajectories. This property immediately implies that the Markov process in question defines a contraction in the space of measures, and therefore has a unique stationary measure, which is exponentially mixing. We next discuss a simple criterion for uniqueness and mixing. Roughly speaking, we prove that if a Markov process is recurrent, and the dual-Lipschitz distance between the laws of trajectories issued from close points remains small (without any contraction!), then there is at most one stationary measure, and if it exists, then solution converges to it in distribution. Our third result gives a sufficient condition for uniqueness and exponential mixing. Its proof is based on a stopping times technique well known in the theory of renewal processes; see Chapters XI in [Fel71] or Chapter 5 in [PSS89].

#### 3.1.1 Decay of a Kantorovich functional

Let  $X$  be a Polish space, let  $\mathcal{T} = \mathbb{R}$  or  $\mathbb{Z}$ , let  $\mathcal{T}_+ = \{t \in \mathcal{T} : t \geq 0\}$ , and let  $(u_t, \mathbb{P}_u)$  be a Feller family of Markov processes in  $X$ . Recall that, given a measurable symmetric function  $F : X \times X \rightarrow \mathbb{R}_+$  satisfying (1.40), we define the Kantorovich functional  $\mathcal{K}_F$  with the density  $F$  by relation (1.41). In what follows, we suppose that there is a constant  $C > 0$  and a point  $u_0 \in X$  such that

$$F(u_1, u_2) \leq C (1 + \text{dist}_X(u_1, u_0) + \text{dist}_X(u_2, u_0)) \quad \text{for any } u_1, u_2 \in X. \quad (3.11)$$

Recall that  $\mathcal{P}_1(X)$  is the class of measures  $\mu \in \mathcal{P}(X)$  such that

$$\mathfrak{m}_1(\mu) = \int_X \text{dist}_X(u, u_0) \mu(du) < \infty.$$

Inequality (3.11) implies that if  $\xi_1$  and  $\xi_2$  are two random variables whose laws belong to  $\mathcal{P}_1(X)$ , then  $\mathbb{E} F(\xi_1, \xi_2) < \infty$ .



**Theorem 3.1.1.** *Let  $(u_t, \mathbb{P}_u)$  be a Feller family of Markov processes in  $X$  satisfying the following two conditions.*

**A priori estimate:** *There is a constant  $C > 0$  and a point  $u_0 \in X$  such that*

$$\int_X \text{dist}_X(z, u_0) P_t(u, dz) \leq C (1 + \text{dist}_X(u, u_0)) \quad \text{for any } u \in X, t \in \mathcal{T}_+. \quad (3.12)$$

**Contraction:** *There is a time  $s \in \mathcal{T}_+$ , a constant  $\gamma < 1$ , and a Kantorovich density  $F : X \times X \rightarrow \mathbb{R}_+$  satisfying (3.11) such that*

$$\mathcal{K}_F(\mathfrak{P}_s^* \lambda, \mathfrak{P}_s^* \lambda') \leq \gamma \mathcal{K}_F(\lambda, \lambda') \quad \text{for any } \lambda, \lambda' \in \mathcal{P}_1(X). \quad (3.13)$$

*Then the Markov family  $(u_t, \mathbb{P}_u)$  has a unique stationary distribution  $\mu \in \mathcal{P}_1(X)$ . Moreover, there are positive constants  $C$  and  $\alpha$  such that*

$$\|\mathfrak{P}_t^* \lambda - \mu\|_L^* \leq C e^{-\alpha t} (1 + \mathfrak{m}_1(\lambda)) \quad \text{for } t \in \mathcal{T}_+, \quad (3.14)$$

where  $\lambda \in \mathcal{P}_1(X)$  is an arbitrary measure.

In what follows, we shall say that a Markov process is *exponentially mixing* in the dual-Lipschitz norm if it satisfies (3.14).

*Proof. Step 1.* Let us endow  $\mathcal{P}(X)$  with the dual-Lipschitz distance. By Theorem 1.2.15, this is a complete metric space. Integrating (3.12) with respect to  $\lambda(du)$ , we see that

$$\mathfrak{m}_1(\mathfrak{P}_t^* \lambda) \leq C (1 + \mathfrak{m}_1(\lambda)) \quad \text{for } t \in \mathcal{T}_+. \quad (3.15)$$

Inequalities (1.42) and (3.13) imply that

$$\|\mathfrak{P}_{ks}^* \lambda - \mathfrak{P}_{ks}^* \lambda'\|_L^* \leq \mathcal{K}_F(\mathfrak{P}_{ks}^* \lambda, \mathfrak{P}_{ks}^* \lambda') \leq \gamma^k \mathcal{K}_F(\lambda, \lambda') \quad \text{for } k \geq 0. \quad (3.16)$$

Let us show that  $\{P_{ks}(u, \cdot), k \geq 1\}$  is a Cauchy sequence in  $\mathcal{P}(X)$  for any  $u \in X$ . Indeed, applying the Kolmogorov–Chapman relation, for  $u, u' \in X$ ,  $l \geq k$ , and for an arbitrary function  $f \in L_b(X)$  with  $\|f\|_L \leq 1$ , we obtain

$$\begin{aligned} & (f, P_{ls}(u', \cdot)) - (f, P_{ks}(u, \cdot)) \\ &= \int_X P_{(l-k)s}(u', dz) \int_X (P_{ks}(z, dw) - P_{ks}(u, dw)) f(w) \\ &\leq \gamma^k \int_X P_{(l-k)s}(u', dz) \mathcal{K}_F(\delta_z, \delta_u). \end{aligned}$$

Using the fact that  $\mathcal{K}_F(\delta_{u_1}, \delta_{u_2}) = F(u_1, u_2)$  and recalling inequality (3.11), we see that

$$\begin{aligned} & (f, P_{ls}(u', \cdot)) - (f, P_{ks}(u, \cdot)) \\ &\leq C_1 \gamma^k \int_X P_{(l-k)s}(u', dz) (1 + \text{dist}_X(z, u_0) + \text{dist}_X(u, u_0)) \\ &\leq C_2 \gamma^k (1 + \text{dist}_X(u', u_0) + \text{dist}_X(u, u_0)). \end{aligned}$$

Integrating this inequality with respect to  $\lambda(du)\lambda'(du')$ , where  $\lambda, \lambda' \in \mathcal{P}_1(X)$  are arbitrary measures, using the symmetry with respect to  $\lambda$  and  $\lambda'$ , and taking the supremum over  $f$ , for  $l \geq k$  we derive

$$\begin{aligned} \|\mathfrak{P}_{ls}^* \lambda' - \mathfrak{P}_{ks}^* \lambda\|_L^* &\leq C_2 \gamma^k \int_X ((1 + \text{dist}_X(u', u_0) + \text{dist}_X(u, u_0)) \lambda(du) \lambda(du')) \\ &\leq C_2 \gamma^k (1 + \mathbf{m}_1(\lambda) + \mathbf{m}_1(\lambda')). \end{aligned} \quad (3.17)$$

In particular, taking  $\lambda = \lambda' = \delta_u$ , we see that  $\{P_{ks}(u, \cdot)\}$  is a Cauchy sequence in  $\mathcal{P}(X)$ , and therefore it converges to a limit  $\mu \in \mathcal{P}(X)$ . It follows from (3.17) that  $\mu$  does not depend on  $u \in X$ . Since  $\mu = \lim_{k \rightarrow \infty} (\mathfrak{P}_s^*)^k \delta_u$ , applying  $\mathfrak{P}_s^*$  to this relation, we see that  $\mathfrak{P}_s^* \mu = \mu$ .

*Step 2.* Let us show that  $\mu \in \mathcal{P}_1(X)$ . We choose any increasing sequence  $\{f_n\} \subset C_b(X)$  converging pointwise to the function  $\text{dist}_X(\cdot, u_0)$ . Then, using (3.12), we can write

$$(f_n, \mu) = \lim_{k \rightarrow \infty} (f_n, P_{ks}(u_0, \cdot)) \leq \limsup_{k \rightarrow \infty} \int_X \text{dist}_X(z, u_0) P_{ks}(u_0, dz) \leq C.$$

Applying the monotone convergence theorem and passing to the limit as  $n \rightarrow \infty$ , we see that  $\mathbf{m}_1(\mu) \leq C$ .

*Step 3.* Applying inequality (3.17) with  $l = k$ ,  $\lambda' = \mathfrak{P}_t^* \mu$ ,  $\lambda = \mu$  and using (3.15), we obtain

$$\|\mathfrak{P}_t^* \mu - \mu\|_L^* = \|\mathfrak{P}_{ks}^* \mathfrak{P}_t^* \mu - \mathfrak{P}_{ks}^* \mu\|_L^* \leq C_3 \gamma^k (1 + \mathbf{m}_1(\mu)).$$

Letting  $k \rightarrow \infty$ , we obtain  $\mathfrak{P}_t^* \mu = \mu$  for each  $t \in \mathcal{T}_+$ , and we see that  $\mu$  is a stationary distribution. If  $\lambda \in \mathcal{P}_1(X)$  is another stationary measure, then by (3.17) we have  $\lambda = \mu$ . So  $\mu$  is a unique stationary distribution for the Markov family  $(u_t, \mathbb{P}_u)$  in the space  $\mathcal{P}_1(X)$ .

*Step 4.* To complete the proof of the theorem, it remains to establish (3.14). To this end, take any  $t \in \mathcal{T}_+$  and write  $t = ks + r$ , where  $k \geq 0$  is an integer and  $0 \leq r < s$ . Using (3.15) with  $t = r$  and inequality (3.17) in which  $\lambda' = \mu$  and  $\lambda$  is replaced by  $\mathfrak{P}_r^* \lambda$ , we derive

$$\|\mathfrak{P}_t^* \lambda - \mu\|_L^* = \|\mathfrak{P}_{ks}^* (\mathfrak{P}_r^* \lambda) - \mu\|_L^* \leq C_3 \gamma^k (1 + \mathbf{m}_1(\mathfrak{P}_r^* \lambda)) \leq C_4 \gamma^k (1 + \mathbf{m}_1(\lambda)).$$

This inequality readily implies (3.14).  $\square$

### 3.1.2 Coupling method: uniqueness and mixing

As before, we denote by  $(u_t, \mathbb{P}_u)$ ,  $t \in \mathcal{T}_+$ , a Feller family of Markov processes in  $X$ . Let  $\mathbf{X} = X \times X$  be the direct product of two copies of  $X$ . We write  $\mathbf{u} = (u, u') \in \mathbf{X}$  and denote by  $\Pi : \mathbf{u} \mapsto u$  and  $\Pi' : \mathbf{u} \mapsto u'$  the natural projections. Let us consider a family of Markov processes  $(\mathbf{u}_t, \mathbb{P}_u)$ ,  $t \in \mathcal{T}_+$ , in  $\mathbf{X}$ . We denote by  $\mathbf{P}_t(\mathbf{u}, \Gamma)$  its the transition function and by  $\mathbf{P}_t$  and  $\mathbf{P}_t^*$  the corresponding Markov operators. We shall write  $\mathbf{u}_t = (u_t, u'_t)$ .

**Definition 3.1.2.** The Markov family  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$  is called a *coupling of two copies* of  $(u_t, \mathbb{P}_u)$  if for any  $\mathbf{u} = (u, u') \in \mathbf{X}$  the laws under  $\mathbb{P}_{\mathbf{u}}$  of the processes  $\Pi \mathbf{u}_t$  and  $\Pi' \mathbf{u}_t$  (regarded as measures on the space of functions from  $\mathcal{T}_+$  to  $X$ ) coincide with those of  $u_t$  under  $\mathbb{P}_u$  and  $\mathbb{P}_{u'}$ , respectively. In this case, we shall also say that  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$  is an *extension* of  $(u_t, \mathbb{P}_u)$ .

That is, if  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$  is an extension of  $(u_t, \mathbb{P}_u)$ , then

$$\Pi_* \mathbf{P}_t(\mathbf{u}, \cdot) = P_t(u, \cdot), \quad \Pi'_* \mathbf{P}_t(\mathbf{u}, \cdot) = P_t(u', \cdot). \quad (3.18)$$

For a closed subset  $B \subset X$ , we denote by  $\mathbf{B}$  the direct product  $B \times B$  and by  $\tau(\mathbf{B})$  the first hitting time of  $\mathbf{B}$  for  $\mathbf{u}_t$ :

$$\tau(\mathbf{B}) = \min\{t \geq 0 : u_t \in B, u'_t \in B\} = \min\{t \geq 0 : \mathbf{u}_t \in \mathbf{B}\}.$$

The following result gives a sufficient condition for the uniqueness of a stationary distribution and mixing in the dual-Lipschitz distance.

**Theorem 3.1.3.** *Let  $(u_t, \mathbb{P}_u)$  be a Markov process and let  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$  be its extension. Let us assume that for any integer  $m \geq 1$  there is a closed subset  $B_m \subset X$  and a constant  $\delta_m > 0$  such that  $\delta_m \rightarrow 0$  as  $m \rightarrow \infty$ , and the following two properties hold.*

**Recurrence:** *For any  $\mathbf{u} = (u, u') \in \mathbf{X}$  and  $m \geq 1$ , we have*

$$\mathbb{P}_{\mathbf{u}}\{\tau(\mathbf{B}_m) < \infty\} = 1. \quad (3.19)$$

**Stability:** *There is a constant  $T_m \in \mathcal{T}_+$  such that*

$$\sup_{t \geq T_m} \|P_t(u, \cdot) - P_t(u', \cdot)\|_L^* \leq \delta_m \quad \text{for any } \mathbf{u} \in \mathbf{B}_m = B_m \times B_m. \quad (3.20)$$

*Then, for any  $u, u' \in X$ , we have*

$$\|P_t(u, \cdot) - P_t(u', \cdot)\|_L^* \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (3.21)$$

*Moreover, if  $\mu$  is a stationary distribution for the family  $(u_t, \mathbb{P}_u)$ , then it is unique, and for any  $\lambda \in \mathcal{P}(X)$  we have*

$$\|\mathfrak{P}_t^* \lambda - \mu\|_L^* \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (3.22)$$

That is, the Markov process is *mixing* in the dual-Lipschitz norm.

*Proof. Step 1.* We first assume that (3.21) is proved and establish the claims concerning the stationary distribution. The uniqueness is a straightforward consequence of (3.22). To prove (3.22), we take any function  $f \in L_b(X)$  and write

$$\begin{aligned} |(f, \mathfrak{P}_t^* \lambda - \mu)| &= |(f, \mathfrak{P}_t^* \lambda - \mathfrak{P}_t^* \mu)| \\ &= \left| \iint_{X \times X} (\mathfrak{P}_t f(u) - \mathfrak{P}_t f(u')) \lambda(du) \mu(du') \right| \\ &\leq \iint_{X \times X} |\mathfrak{P}_t f(u) - \mathfrak{P}_t f(u')| \lambda(du) \mu(du'). \end{aligned} \quad (3.23)$$

Taking into account (3.21) and using the Lebesgue theorem on dominated convergence, we see that the right-hand side converges to zero as  $t \rightarrow +\infty$  for any  $f \in L_b(X)$ . In view of assertion (ii) of Theorem 1.2.15, this is equivalent to (3.22).

*Step 2.* We now prove (3.21). For any  $m \geq 1$  and  $t \geq T_m$ , write

$$\tau(m, t) = \tau(B_m) \wedge t, \quad p(\mathbf{u}, m, t) = \mathbb{P}_{\mathbf{u}}\{\tau(B_m) + T_m > t\}.$$

Applying the strong Markov property (1.57) to the family  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$  and taking into account the first relation in (3.18), for any integer  $m \geq 1$  we obtain

$$\begin{aligned} P_t(u, \Gamma) &= \mathbf{P}_t(\mathbf{u}, \Gamma \times X) = \mathbb{E}_{\mathbf{u}} \mathbf{P}_{t-\tau(m, t)}(\mathbf{u}_{\tau(m, t)}, \Gamma \times X) \\ &= \mathbb{E}_{\mathbf{u}} P_{t-\tau(m, t)}(u_{\tau(m, t)}, \Gamma). \end{aligned}$$

A similar relation holds for  $P_t(u', \Gamma)$ . It follows that

$$\|P_t(u, \cdot) - P_t(u', \cdot)\|_L^* \leq \mathbb{E}_{\mathbf{u}} g(t - \tau(m, t), \mathbf{u}_{\tau(m, t)}), \quad (3.24)$$

where for  $s \geq 0$  and  $\mathbf{z} = (z, z') \in \mathbf{X}$  we set

$$g(s, \mathbf{z}) = \|P_s(z, \cdot) - P_s(z', \cdot)\|_L^*.$$

Now note that, by the stability condition (3.20), we have

$$\sup_{s \geq T_m} g(s, \mathbf{z}) \leq \delta_m \quad \text{for } \mathbf{z} \in \mathbf{B}_m. \quad (3.25)$$

Combining inequalities (3.24) and (3.25) and taking into account that the relation  $\tau(B_m) + T_m \leq t$  implies  $g(t - \tau(m, t), \mathbf{u}_{\tau(m, t)}) \leq \delta_m$ , we get

$$\|P_t(u, \cdot) - P_t(u', \cdot)\|_L^* \leq \delta_m + p(\mathbf{u}, m, t) \quad \text{for } t \geq T_m.$$

In view of (3.19), the right-hand side of this inequality can be made arbitrarily small by choosing  $m$  first and then taking  $t$  to be sufficiently large. This completes the proof of Theorem 3.1.3.  $\square$

Analysing the above proof, it is easy to see that if  $\mathbf{Y} \subset \mathbf{X}$  is a Borel subset such that

$$\sup_{\mathbf{u} \in \mathbf{Y}} \mathbb{P}_{\mathbf{u}}\{\tau(B_m) > t\} \leq p(m, t), \quad (3.26)$$

where  $p(m, t) \rightarrow 0$  as  $t \rightarrow \infty$  for any integer  $m \geq 1$ , then

$$\sup_{\mathbf{u} \in \mathbf{Y}} \|P_t(u, \cdot) - P_t(u', \cdot)\|_L^* \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (3.27)$$

This simple observation enables one to prove the following result on uniform convergence to the stationary distribution.

**Theorem 3.1.4.** *Under the hypotheses of Theorem 3.1.3, assume that (3.26) holds for any compact set  $Y \subset X$  and any integer  $m \geq 1$ , where  $p(m, t) \rightarrow 0$  as  $t \rightarrow \infty$ . Then for any compact subset  $\Lambda$  of the space  $\mathcal{P}(H)$  endowed with the dual-Lipschitz metric convergence (3.22) holds uniformly with respect to  $\lambda \in \Lambda$ .*

*Proof.* Let us fix a constant  $\varepsilon > 0$  and use Prokhorov's theorem to find a compact subset  $Y \subset X$  such that  $\lambda(Y) \geq 1 - \varepsilon$  for any  $\lambda \in \Lambda \cup \{\mu\}$ . Let  $f \in L_b(X)$  be an arbitrary function such that  $\|f\|_L \leq 1$ . Then it follows from (3.23) that

$$|(f, \mathfrak{P}_t^* \lambda - \mu)| \leq \sup_{u, u' \in Y} |\mathfrak{P}_t f(u) - \mathfrak{P}_t f(u')| + 4\varepsilon.$$

Using inequality (3.27) with  $Y = Y \times Y$  and recalling that  $\varepsilon$  and  $f$  are arbitrary, we arrive at the required uniform convergence.  $\square$

Let us also mention that the rate of convergence in (3.22) can be estimated in terms of the mean value of the hitting time  $\tau(B)$  and the sequence  $\{\delta_m\}$  (see the recurrence and stability conditions of Theorem 3.1.3). In particular, if the Markov process in question depends on a parameter  $\theta$  in such a way that the above-mentioned quantities are estimated uniformly in  $\theta$ , then the rate of convergence to the stationary measure is also uniform in  $\theta$ .

An application of Theorem 3.1.3 to the Navier–Stokes system is presented in Section 3.3. Here we discuss the case of the viscous *Burgers equation*, for which the stability holds uniformly with respect to the viscosity.

*Example 3.1.5.* Let us consider the Burgers equation with periodic boundary conditions:

$$\dot{u} - \nu \partial_x^2 u + \partial_x u^2 = \frac{\partial}{\partial t} \zeta^\omega(t, x), \quad x \in \mathbb{T}, \quad \int_{\mathbb{T}} u \, dx \equiv \int_{\mathbb{T}} \zeta \, dx \equiv 0, \quad (3.28)$$

where  $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ . We abbreviate  $\dot{H}^m(\mathbb{T}; \mathbb{R}) = \dot{H}^m$  for  $m \in \mathbb{R}$  (see Section 1.1.1) and denote by  $\{e_j, j \in \mathbb{Z}_0\}$  the usual  $L^2$ -normalised trigonometric basis of these spaces. As for the white-forced Navier–Stokes system, we choose  $\zeta^\omega(t, x) = \sum_j b_j \beta_j^\omega(t) e_j(x)$ . For simplicity, assume that all quantities  $\mathfrak{B}_m = \sum_j j^{2m} b_j^2$  are finite. Then Eq. (3.28) is well posed in each space  $\dot{H}^m$ ,  $m \geq 1$ , and defines there a Markov process; see [DZ96]. Consider the RDS  $\{\varphi_t^\omega, t \geq 0\}$  defined by (3.28). It is known that for a.e.  $\omega$  the maps  $\varphi_t^\omega$  extend by continuity to non-expanding transformations  $\dot{L}^1 \rightarrow \dot{L}^1$ , where  $\dot{L}^1$  denotes the space of integrable functions on  $\mathbb{T}$  with zero mean value; e.g., see Lemma 3.2.2 in [Hör97]. So (3.28) defines a Markov process in  $\dot{L}^1$ . Theorem 3.1.3 applies to it. Indeed, choose  $B_m = B_{\dot{L}^1}(1/m)$ . Then the recurrence follows from the same simple argument as for the Navier–Stokes system (see below Section 3.3.2 and the discussion after Theorem 3.3.1). The stability with  $T_m = 0$  and  $\delta_m = 1/m$  immediately follows from the fact that the maps  $\varphi_t^\omega : \dot{L}^1 \rightarrow \dot{L}^1$  are non-expanding.

If all numbers  $b_j$  are non-zero, then the rate of recurrence may be chosen to be independent of  $\nu$  (cf. Section 3.3.2 below). Since characteristics of stability are

also independent of  $\nu$ , in this case the rate of mixing for (3.28) does not depend on the viscosity. See [Bor12] for further properties of the Burgers equation (3.28) and [BK07] for its relevance as a physical model.

### 3.1.3 Coupling method: exponential mixing

This subsection can be omitted at first reading, since its results are used only in the proof of exponential mixing for equations with unbounded noise. To simplify the presentation, we assume here that  $X$  is a separable Banach space with a norm  $\|\cdot\|$ . The main result of this section—Theorem 3.1.7—gives a sufficient condition for exponential mixing for Markov processes in  $X$ . In addition to the coupling, the proof uses some techniques of the theory of renewal processes; cf. [Fel71, Chapter XI] or [PSS89, Chapter 5].

Let  $(u_t, \mathbb{P}_u)$  be a Feller family of Markov processes in  $X$ , let  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$  be its extension, and let  $B$  be a closed subset in  $X$ . Recall that  $\tau(B)$  stands for the first hitting time of the set  $\mathbf{B} = B \times B$  for  $\mathbf{u}_t$ . We also introduce the stopping time

$$\sigma = \inf\{t \in \mathcal{T}_+ : \|u_t - u'_t\| \geq C e^{-\beta t}\}, \quad (3.29)$$

where  $C$  and  $\beta$  are some fixed positive constants. In other words,  $\sigma$  is the first instant when the curves  $u_t$  and  $u'_t$  “stop converging” to each other exponentially fast. In particular, if  $\sigma = \infty$ , then

$$\|u_t - u'_t\| \leq C e^{-\beta t} \quad \text{for } t \geq 0. \quad (3.30)$$

**Definition 3.1.6.** We shall say that the family  $(u_t, \mathbb{P}_u)$  satisfies the *coupling hypothesis* if there is an extension  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$ , a closed set  $B \subset X$ , and an increasing function  $g(r) \geq 1$  of the variable  $r \geq 0$  such that the following two properties hold.

**Recurrence:** There is  $\delta > 0$  such that

$$\mathbb{E}_{\mathbf{u}} \exp(\delta \tau(B)) \leq G(\mathbf{u}) \quad \text{for all } \mathbf{u} = (u, u') \in \mathbf{X}, \quad (3.31)$$

where we set  $G(\mathbf{u}) = g(\|u\|) + g(\|u'\|)$ .

**Exponential squeezing:** There are positive constants  $\delta_1, \delta_2, c, K$ , and  $q > 1$  such that, for any  $\mathbf{u} \in \mathbf{B}$ , we have

$$\mathbb{P}_{\mathbf{u}}\{\sigma = \infty\} \geq \delta_1, \quad (3.32)$$

$$\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\sigma < \infty\}} \exp(\delta_2 \sigma)\} \leq c, \quad (3.33)$$

$$\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\sigma < \infty\}} G(\mathbf{u}_\sigma)^q\} \leq K. \quad (3.34)$$

Any extension of  $(u_t, \mathbb{P}_u)$  satisfying the above properties will be called a *mixing extension*.

Note that  $\sigma \equiv 0$  for any  $\mathbf{u} \in \mathbf{X}$  with  $\|u - u'\| > C$ . It follows from (3.32) that the set  $\mathbf{B}$  must belong to the  $C$ -neighbourhood of the diagonal, that is,  $\mathbf{B} \subset \{\mathbf{u} : \|u - u'\| \leq C\}$ .

Before formulating the main result of this subsection, we make some comments on the above definition. Let us take an arbitrary initial point  $\mathbf{u} \in \mathbf{B}$ . Then, in view of (3.32), with  $\mathbb{P}_{\mathbf{u}}$ -probability  $\geq \delta_1$ , we have  $\sigma = \infty$ , and therefore, with the same probability, the trajectories  $u_t$  and  $u'_t$  converge exponentially fast (see (3.30)). On the other hand, if they do not, inequality (3.33) says that the first instant  $\sigma$  when the trajectories “stop converging” is not very large. Moreover, by (3.34), we have some control over  $\mathbf{u}_t$  at the instant  $t = \sigma$ . If the initial point  $\mathbf{u} \in \mathbf{X}$  does not belong to  $\mathbf{B}$ , we cannot claim that the above properties hold. However, we know that, with probability 1, any trajectory hits the set  $\mathbf{B}$ , and by (3.31), the first hitting time  $\tau(\mathbf{B})$  has a finite exponential moment.

These observations make it plausible that, for any initial point  $\mathbf{u} \in \mathbf{X}$ , the trajectories  $u_t$  and  $u'_t$  converge exponentially fast. In fact, we have the following result.

**Theorem 3.1.7.** *Let  $(u_t, \mathbb{P}_{\mathbf{u}})$  be a Feller family of Markov processes that possesses a mixing extension  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$ . Then there is a random time  $\ell \in \mathcal{T}_+$  such that, for any  $\mathbf{u} \in \mathbf{X}$ , with  $\mathbb{P}_{\mathbf{u}}$ -probability 1, we have*

$$\|u_t - u'_t\| \leq C_1 e^{-\beta(t-\ell)} \quad \text{for } t \geq \ell, \quad (3.35)$$

$$\mathbb{E}_{\mathbf{u}} e^{\alpha \ell} \leq C_1 G(\mathbf{u}), \quad (3.36)$$

where  $\mathbf{u} \in \mathbf{X}$  is an arbitrary initial point,  $g(r)$  is the function in Definition 3.1.6, and  $C_1$ ,  $\alpha$ , and  $\beta$  are positive constants not depending on  $\mathbf{u}$  and  $t$ . If, in addition, there is an increasing function  $\tilde{g}(r) \geq 1$  such that

$$\mathbb{E}_{\mathbf{u}} g(\|u_t\|) \leq \tilde{g}(\|u\|) \quad \text{for } u \in X, t \geq 0, \quad (3.37)$$

then the family  $(u_t, \mathbb{P}_{\mathbf{u}})$  has a unique stationary measure  $\mu \in \mathcal{P}(X)$ , and there is a constant  $\gamma > 0$  such that

$$\|P_t(u, \cdot) - \mu\|_L^* \leq V(\|u\|) e^{-\gamma t} \quad \text{for } t \geq 0, u \in X, \quad (3.38)$$

where  $V$  is given by the relation

$$V(r) = 3C_1(g(r) + \tilde{g}(0)). \quad (3.39)$$

*Proof.* We first show that inequalities (3.35) – (3.37) imply the existence and uniqueness of a stationary measure and the mixing property (3.38). Namely, we shall derive from these relations that, for any  $u, u' \in X$ ,

$$\|P_t(u, \cdot) - P_t(u', \cdot)\|_L^* \leq 3C_1 G(\mathbf{u}) e^{-\gamma t}, \quad t \geq 0. \quad (3.40)$$

To this end, we fix an arbitrary functional  $f \in L_b(X)$  with  $\|f\|_L \leq 1$  and note that

$$\begin{aligned} |(f, P_t(u, \cdot) - P_t(u', \cdot))| &\leq \mathbb{E}_{\mathbf{u}} |f(u_t) - f(u'_t)| \\ &\leq 2\mathbb{P}_{\mathbf{u}}\{\ell > \tfrac{t}{2}\} + \mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\ell \leq \frac{t}{2}\}} |f(u_t) - f(u'_t)|\}. \end{aligned} \quad (3.41)$$

In view of (3.36) and Chebyshev's inequality, we have

$$\mathbb{P}_{\mathbf{u}}\{\ell > \frac{t}{2}\} \leq C_1 G(\mathbf{u}) e^{-\frac{\alpha t}{2}}.$$

Furthermore, it follows from the assumption  $\|f\|_L \leq 1$  and inequality (3.35) that the second term on the right-hand side of (3.41) does not exceed

$$\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\ell \leq \frac{t}{2}\}} \|u_t - u'_t\|\} \leq C_1 e^{-\frac{\beta t}{2}}.$$

Substituting the last two estimates into (3.41) and using that  $G \geq 2$ , we obtain the required inequality (3.40) with  $\gamma = \frac{1}{2}(\alpha \wedge \beta)$ .

We now use (3.40) to prove the existence and uniqueness of a stationary measure and inequality (3.38). Let us fix arbitrary points  $u, u' \in X$  and a functional  $f \in L_b(X)$  such that  $\|f\|_L \leq 1$ . By the Kolmogorov–Chapman relation and inequality (3.40), for  $t \leq s$  we have

$$\begin{aligned} |(f, P_t(u, \cdot) - P_s(u', \cdot))| &= \left| \int_X P_{s-t}(u', dz) \int_X (P_t(u, dv) - P_t(z, dv)) f(v) \right| \\ &\leq 3C_1 e^{-\gamma t} \int_X P_{s-t}(u', dz) [g(\|u\|) + g(\|z\|)] \\ &= 3C_1 e^{-\gamma t} [g(\|u\|) + \mathbb{E}_{u'} g(\|u_{s-t}\|)]. \end{aligned}$$

Taking into account (3.37), we conclude that

$$\|P_t(u, \cdot) - P_s(u', \cdot)\|_L^* \leq 3C_1 e^{-\gamma t} (g(\|u\|) + \tilde{g}(\|u'\|)). \quad (3.42)$$

Since  $\mathcal{P}(X)$  is a complete metric space with respect to the dual-Lipschitz distance, we conclude that  $P_t(u, \cdot)$  converges, as  $t \rightarrow +\infty$ , to a measure  $\mu \in \mathcal{P}(X)$ , which does not depend on  $u$  and is stationary. Setting  $u' = 0$  in (3.42) and passing to the limit as  $s \rightarrow +\infty$ , we obtain inequality (3.38) with  $V$  given by (3.39).

Thus, we need to establish inequalities (3.35) and (3.36). Their proof is divided into four steps.

*Step 1.* We introduce the stopping time

$$\rho = \inf\{t \in \mathcal{T}_+ : t \geq \sigma, \mathbf{u}_t \in \mathbf{B}\}. \quad (3.43)$$

In other words, we wait until the first instant  $\sigma$  when the trajectories  $u_t$  and  $u'_t$  “stop converging” and denote by  $\rho$  the first hitting time of  $\mathbf{B}$  after  $\sigma$ . Let  $\delta$ ,  $\delta_1$  and  $\delta_2$  be the constants in (3.31), (3.32), and (3.33). We claim that, for any  $\mathbf{u} \in \mathbf{B}$ ,

$$\mathbb{P}_{\mathbf{u}}\{\rho = \infty\} \geq \delta_1, \quad (3.44)$$

$$\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\rho < \infty\}} e^{\alpha \rho}\} \leq q, \quad (3.45)$$

where  $\alpha \leq \delta_2 \wedge \delta$  and  $q < 1$  are positive constants not depending on  $\mathbf{u}$ . Indeed, the definition of  $\rho$  and (3.31) imply that  $\{\rho = \infty\} = \{\sigma = \infty\}$ , so (3.44) is an immediate consequence of (3.32).



To prove (3.45), we first show that

$$\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\rho < \infty\}}e^{\delta_3\rho}\} \leq M \quad \text{for any } \mathbf{u} \in \mathbf{B}, \quad (3.46)$$

where  $\delta_3 = \frac{(q-1)(\delta_2 \wedge \delta)}{q}$  and  $M > 0$  is a constant not depending on  $\mathbf{u}$ . Indeed, using relation (3.43), the strong Markov property, and inequality (3.31), we derive

$$\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\rho < \infty\}}e^{\delta_3\rho}\} = \mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\sigma < \infty\}}e^{\delta_3\sigma}(\mathbb{E}_{\mathbf{u}_\sigma}e^{\delta_3\tau_B})\} \leq \mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\sigma < \infty\}}e^{\delta_3\sigma}G(\mathbf{u}_\sigma)\}.$$

Combining this with (3.33) and (3.34), we conclude that

$$\begin{aligned} \mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\rho < \infty\}}e^{\delta_3\rho}\} &\leq (\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\sigma < \infty\}}e^{\delta_2\sigma}\})^{\frac{q-1}{q}} (\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\sigma < \infty\}}G(\mathbf{u}_\sigma)^q\})^{\frac{1}{q}} \\ &\leq (c^{q-1}K)^{\frac{1}{q}} =: M. \end{aligned}$$

To derive (3.45), let us set  $\alpha = \varepsilon\delta_3$  and note that, in view of (3.44) and (3.46), we have

$$\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\rho < \infty\}}e^{\alpha\rho}\} \leq (\mathbb{P}_{\mathbf{u}}\{\rho < \infty\})^{1-\varepsilon} (\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\rho < \infty\}}e^{\delta_3\rho}\})^\varepsilon \leq (1 - \delta_1)^{1-\varepsilon} M^\varepsilon.$$

The right-hand side of this inequality is less than 1 if  $\varepsilon > 0$  is sufficiently small.

*Step 2.* We now consider the iterations of  $\rho$ . Namely, we define two sequences of stopping times  $\rho_k$  and  $\rho'_k$  by the formulas

$$\begin{aligned} \rho_0 &= \tau_B, \\ \rho'_k &= \inf\{t \in \mathcal{T}_+ : t \geq \rho_{k-1}, \|u_t - u'_t\| \geq C e^{-\beta(t-\rho_{k-1})}\}, \quad k \geq 1, \\ \rho_k &= \inf\{t \in \mathcal{T}_+ : t \geq \rho'_k, \mathbf{u}_t \in \mathbf{B}\}, \quad k \geq 1. \end{aligned}$$

That is,  $\rho_k$  is the first occurrence of  $\rho$  after  $\rho_{k-1}$ . We claim that

$$\mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\rho_k < \infty\}}e^{\alpha\rho_k}\} \leq q^k G(\mathbf{u}) \quad \text{for any } \mathbf{u} \in \mathbf{X}. \quad (3.47)$$

Indeed, inequality (3.47) is true for  $k = 0$  in view of (3.31). Let us assume that (3.47) holds for  $\rho_0, \dots, \rho_{k-1}$ , where  $k \geq 1$ . Since  $\mathbf{u}_{\rho_k} \in \mathbf{B}$ , inequality (3.45) and the strong Markov property imply that

$$\begin{aligned} \mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\rho_k < \infty\}}e^{\alpha\rho_k}\} &\leq \mathbb{E}_{\mathbf{u}}\left\{\mathbb{I}_{\{\rho_{k-1} < \infty\}}e^{\alpha\rho_{k-1}} \sup_{\mathbf{v} \in \mathbf{B}} \mathbb{E}_{\mathbf{v}}(\mathbb{I}_{\{\rho < \infty\}}e^{\alpha\rho})\right\} \\ &\leq q \mathbb{E}_{\mathbf{u}}\{\mathbb{I}_{\{\rho_{k-1} < \infty\}}e^{\alpha\rho_{k-1}}\}. \end{aligned}$$

Combining this with the induction hypothesis, we arrive at (3.47).

*Step 3.* We now note that, if  $\rho_k < \infty$  and  $\rho_{k+1} = \infty$  for an integer  $k \geq 0$ , then

$$\|u_t - u'_t\| \leq C e^{-\beta(t-\rho_k)} \quad \text{for } t \geq \rho_k. \quad (3.48)$$

For any  $\mathbf{u} \in \mathbf{X}$ , let us set  $\bar{k} = \sup\{k \geq 0 : \rho_k < \infty\}$ . We wish to show that, for any  $\mathbf{u} \in \mathbf{X}$ ,

$$\bar{k} < \infty \quad \mathbb{P}_{\mathbf{u}}\text{-almost surely.} \quad (3.49)$$

To this end, note that, in view of (3.44) and the strong Markov property,

$$\mathbb{P}_u\{\rho_k < \infty\} \leq (1 - \delta_1)\mathbb{P}_u\{\rho_{k-1} < \infty\} \leq (1 - \delta_1)^k \mathbb{P}_u\{\rho_0 < \infty\} \leq (1 - \delta_1)^k.$$

Hence, the Borel–Cantelli lemma implies (3.49).

*Step 4.* Let us set

$$\ell = \begin{cases} \rho_{\bar{k}} & \text{if } \bar{k} < \infty, \\ +\infty & \text{if } \bar{k} = \infty. \end{cases}$$

Inequality (3.35) follows immediately from (3.48), the definition of  $\rho_k$ , and the fact that  $\rho_{\bar{k}+1} = \infty$ . To prove (3.36), we write

$$\mathbb{E}_u e^{\alpha\ell} = \sum_{k=0}^{\infty} \mathbb{E}_u \{ \mathbb{I}_{\{\bar{k}=k\}} e^{\alpha\rho_k} \} \leq \sum_{k=0}^{\infty} \mathbb{E}_u \{ \mathbb{I}_{\{\rho_k < \infty\}} e^{\alpha\rho_k} \} \leq (1 - q)^{-1} G(u),$$

where we used inequality (3.47) and the fact that  $\ell < \infty$  with  $\mathbb{P}_u$ -probability 1 for any  $u \in \mathbf{X}$ . This completes the proof of Theorem 3.1.7.  $\square$

*Remark 3.1.8.* Analysing the proof given above, it is not difficult to see that Theorem 3.1.7 remains valid if  $\sigma$  is replaced with any other stopping time  $\tilde{\sigma}$  such that

$$\mathbb{P}_u\{\tilde{\sigma} \leq \sigma\} = 1 \quad \text{for any } u \in B.$$

In other word, if inequalities (3.32) – (3.35) hold with  $\sigma$  replaced by  $\tilde{\sigma}$ , then the conclusion of Theorem 3.1.7 is true. To see this, it suffices to repeat the arguments above, replacing everywhere  $\sigma$  by  $\tilde{\sigma}$ .

## 3.2 Dissipative RDS with bounded kicks

### 3.2.1 Main result

Let  $H$  be a separable Banach space with a norm  $\|\cdot\|$  and let  $S : H \rightarrow H$  be a continuous (nonlinear) operator. We consider an RDS defined by the relation

$$u_k = S(u_{k-1}) + \eta_k, \quad k \geq 1, \quad (3.50)$$

where  $\{\eta_k\}$  is a sequence of independent identically distributed (i.i.d.) random variables in  $H$  whose law  $\mathcal{D}(\eta_1)$  has a bounded support. As it was explained in Example 1.3.15, one can regard (3.50) as a Markov RDS in  $H$ . Our goal in this section is to study the Markov chain defined by (3.50) in  $H$  (See Example 1.3.6). In what follows, we assume that the following four conditions are satisfied.

**Condition 3.2.1.** For any  $R > r > 0$  there are positive constants  $C = C(R)$ ,  $a = a(R, r) < 1$  and an integer  $n_0 = n_0(R, r) \geq 1$  such that

$$\|S(u_1) - S(u_2)\| \leq C(R)\|u_1 - u_2\| \quad \text{for all } u_1, u_2 \in B_H(R), \quad (3.51)$$

$$\|S^n(u)\| \leq \max\{a\|u\|, r\} \quad \text{for } u \in B_H(R), \quad n \geq n_0, \quad (3.52)$$

where  $S^n$  stands for the  $n^{\text{th}}$  iteration of  $S$ .

Let us denote by  $\mathcal{K}$  the support of the measure  $\mathcal{D}(\eta_1)$  and define the sets of attainability from  $B \subset H$  by the relations

$$\mathcal{A}_0(B) = B, \quad \mathcal{A}_k(B) = S(\mathcal{A}_{k-1}(B)) + \mathcal{K} \quad \text{for } k \geq 1.$$

**Condition 3.2.2.** There is a constant  $\rho > 0$  and a non-decreasing continuous function  $k_0 = k_0(R) > 0$  such that

$$\mathcal{A}_k(B_H(R)) \subset B_H(\rho) \quad \text{for any } R > 0 \text{ and } k \geq k_0(R), \quad (3.53)$$

where  $B_H(r)$  stands for the closed ball of radius  $r$  centred at origin.

Let us introduce the set

$$\mathcal{A} = \overline{\bigcup_{k \geq 0} \mathcal{A}_k(B_H(\rho))}, \quad (3.54)$$

where  $\rho$  is the constant defined in Condition 3.2.2. A straightforward consequence of (3.51) and (3.53) is that  $\mathcal{A}$  is a bounded invariant absorbing set for the RDS (3.50); see Section 1.3.4.

**Condition 3.2.3.** There is a finite-dimensional subspace  $E \subset H$ , a continuous linear projection operator  $P : H \rightarrow H$  onto  $E$ , and a constant  $\gamma < 1$  such that

$$\|Q(S(u_1) - S(u_2))\| \leq \gamma \|u_1 - u_2\| \quad \text{for all } u_1, u_2 \in \mathcal{A}, \quad (3.55)$$

where  $Q = I - P$ .

Note that, in view of the Hahn–Banach theorem, a continuous projection onto a finite-dimensional subspace of a Banach space always exists. The last condition concerns the law of random variables  $\eta_k$ .

**Condition 3.2.4.** The support of  $\mathcal{D}(\eta_1)$  contains the origin and the random variables  $P_*\eta_1$  and  $Q_*\eta_1$  are independent. Furthermore, the measure  $P_*\mathcal{D}(\eta_1)$  has a density  $p(x)$  with respect to the Lebesgue measure on  $E$ , and there is a constant  $C$  such that

$$\int_E |p(x+y) - p(x)| dx \leq C \|y\| \quad \text{for } y \in E. \quad (3.56)$$

Before formulating the main result of this section, we make some comments on the above conditions. Inequality (3.51) is nothing else but the Lipschitz continuity of  $S$  on bounded subsets of  $H$ , while (3.52) expresses the property of dissipativity for  $S$ . In particular, it follows from (3.52) that  $u = 0$  is a stable fixed point for the deterministic dynamical system generated by  $S$ . It is easy to see that inequality (3.52) of Condition 3.2.1 is fulfilled if  $S$  satisfies the inequality

$$\|S(u)\| \leq q \|u\| \quad \text{for all } u \in H,$$

where  $q < 1$  and  $C$  are positive constants not depending on  $u$ . Condition 3.2.2 means that the RDS (3.50) has a bounded absorbing set, and the function  $k_0(R)$

provides an upper bound for the time of absorption. In applications to evolutionary PDE's, Condition 3.2.3 (sometimes referred to as the *squeezing property*) manifests itself as a smoothing property of flow maps. It is usually satisfied for quasilinear parabolic PDE's, such as Ginzburg–Landau and reaction–diffusion equations. It also holds for the 2D Navier–Stokes system, which can be written as a non-local parabolic PDE; see (2.19). Finally, Condition 3.2.4 expresses the non-degeneracy of the random perturbation and the continuous dependence on the shifts of the law of its projection to  $E$ . The latter property is certainly satisfied if the density  $p$  belongs to the Sobolev class  $W^{1,1}(E)$ .<sup>1</sup>

We now formulate the main result of this section. Its proof is based on Theorem 3.1.1 and is given in the next two subsections. An application to the randomly forced Navier–Stokes system is discussed in Section 3.2.4.

**Theorem 3.2.5.** *Suppose that Conditions 3.2.1 – 3.2.4 are satisfied. Then the RDS (3.50) has a unique stationary distribution  $\mu \in \mathcal{P}(H)$ . Moreover, there are positive constants  $C$  and  $\alpha_0$  such that, for any  $\alpha \in (0, \alpha_0]$  we have*

$$\|\mathfrak{P}_k^* \lambda - \mu\|_L^* \leq C e^{-\alpha k} \int_H \exp(\alpha k_0(\|u\|)) \lambda(du), \quad k \geq 0, \quad (3.57)$$

where  $\lambda \in \mathcal{P}(H)$  is an arbitrary measure for which the right-hand side of (3.57) is finite, and  $k_0$  is defined in Condition 3.2.2.

### 3.2.2 Coupling

Let us denote by  $\chi(du)$  the law of the random variable  $\eta_1$ , by  $\chi_E(dx)$  its image under the projection  $P$ , and by  $\chi_E(u, dx)$ ,  $u \in H$ , the law of  $P(S(u) + \eta_1)$ . We shall need the following auxiliary result.

**Lemma 3.2.6.** *Let  $S : H \rightarrow H$  be a continuous operator, and suppose that Condition 3.2.4 is satisfied. Then there is a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  such that for any pair  $v, v' \in H$  there are two  $H$ -valued random variables  $\zeta = \zeta(v, v', \omega)$  and  $\zeta' = \zeta'(v, v', \omega)$  possessing the following properties.*

- (i) *The laws of  $\zeta$  and  $\zeta'$  coincide with  $\chi$ .*
- (ii) *The random variables  $(P\zeta, P\zeta')$  and  $(Q\zeta, Q\zeta')$  are independent. Furthermore, the projections  $Q\zeta$  and  $Q\zeta'$  coincide for all  $\omega \in \Omega$  and do not depend on  $(v, v')$ .*
- (iii) *The pair*

$$V = P(S(v) + \zeta), \quad V' = P(S(v') + \zeta')$$

*is a maximal coupling for  $(\chi_E(v, \cdot), \chi_E(v', \cdot))$ . Moreover,*

$$\mathbb{P}\{V \neq V'\} \leq C_1 \|S(v) - S(v')\| \quad \text{for any } v, v' \in H, \quad (3.58)$$

*where  $C_1 > 0$  is a constant not depending on  $v$  and  $v'$ .*

<sup>1</sup>Inequality (3.56) means precisely that the density  $p$  belongs to the Nikolski–Besov space  $\Lambda_1^{1,\infty}(E)$ ; see Chapter V in [Ste70].

(iv) The functions  $\zeta$  and  $\zeta'$  are measurable with respect to  $(v, v', \omega)$ .

*Proof.* For any  $v, v' \in H$ , let  $(V, V')$  be a maximal coupling for the pair of measures  $(\chi_E(v, \cdot), \chi_E(v', \cdot))$ . By Theorem 1.2.28, we can assume that  $V$  and  $V'$  are defined on the same probability space  $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$  for all  $v, v' \in H$  and are measurable functions of  $(v, v', \omega_1)$ . Let  $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$  be the probability space on which the random variables  $\eta_k$  are defined. We denote by  $(\Omega, \mathcal{F}, \mathbb{P})$  the direct product of these two spaces and, for any  $\omega = (\omega_1, \omega_2) \in \Omega$ , set

$$\begin{aligned}\zeta(v, v', \omega) &= V(v, v', \omega_1) - \mathbf{P}S(v) + \mathbf{Q}\eta_1(\omega_2), \\ \zeta'(v, v', \omega) &= V'(v, v', \omega_1) - \mathbf{P}S(v') + \mathbf{Q}\eta_1(\omega_2).\end{aligned}\tag{3.59}$$

Assertions (i), (ii), (iv) and the first part of (iii) are straightforward consequences of the construction. To prove inequality (3.58), we note that  $\chi_E(v, dy)$  is absolutely continuous with respect to the Lebesgue measure on  $E$ , and the corresponding density is given by  $p(x - \mathbf{P}S(v))$ . Since  $(V, V')$  is a maximal coupling, inequality (3.56) and relation (1.23) imply that

$$\begin{aligned}\mathbb{P}\{V \neq V'\} &= \|\chi_E(v, \cdot) - \chi_E(v', \cdot)\|_{\text{var}} \\ &= \frac{1}{2} \int_E |p(x - \mathbf{P}S(v)) - p(x - \mathbf{P}S(v'))| dx \\ &\leq \frac{1}{2} C \|\mathbf{P}(S(v) - S(v'))\| \leq C_1 \|S(v) - S(v')\|,\end{aligned}$$

where we used the continuity of the projection  $\mathbf{P}$ . This completes the proof of Lemma 3.2.6.  $\square$

We now fix a constant  $d > 0$  and use the following rules to define *coupling operators*  $(\mathcal{R}, \mathcal{R}')$  on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  (constructed in the proof of Lemma 3.2.6): for  $v, v' \in \mathcal{A}$  with  $\|v - v'\| \leq d$ , we set

$$\mathcal{R}(v, v', \omega) = S(v) + \zeta(v, v', \omega), \quad \mathcal{R}'(v, v', \omega) = S(v') + \zeta'(v, v', \omega); \tag{3.60}$$

for  $v, v' \in \mathcal{A}$  with  $\|v - v'\| > d$ , we set

$$\mathcal{R}(v, v', \omega) = S(v) + \eta_1(\omega_2), \quad \mathcal{R}'(v, v', \omega) = S(v') + \eta_1(\omega_2). \tag{3.61}$$

Let us recall that the concept of extension for a family of Markov chains was introduced in Section 3.1.2. We now define an extension of the Markov chain associated with the RDS (3.50). Denote by  $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$  the direct product of countably many copies of the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  in Lemma 3.2.6. We shall write  $\widehat{\omega} = (\omega^1, \omega^2, \dots)$  for the points of  $\widehat{\Omega}$ .

Let us consider the following RDS  $\Phi_k = (\Phi_k, \Phi'_k)$  on the product space  $H \times H$ :

$$\begin{aligned}\Phi_0(\widehat{\omega})(v, v') &= v, \quad \Phi'_0(\widehat{\omega})(v, v') = v', \\ \Phi_k(\widehat{\omega})(v, v') &= \mathcal{R}(\Phi_{k-1}(v, v'), \Phi'_{k-1}(v, v'), \omega^k), \\ \Phi'_k(\widehat{\omega})(v, v') &= \mathcal{R}'(\Phi_{k-1}(v, v'), \Phi'_{k-1}(v, v'), \omega^k).\end{aligned}$$

Note that  $\Phi_k$  depends only on  $(\omega^1, \dots, \omega^k)$ . Arguing by induction and using the definition of the coupling operators  $\mathcal{R}$  and  $\mathcal{R}'$ , we see that the Markov chain  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  associated with  $\Phi_k$  is an extension<sup>2</sup> for the restriction to  $\mathcal{A}$  of the Markov chain  $(u_k, \mathbb{P}_u)$  corresponding to (3.50). In what follows, we shall drop the hat from the notations and write  $\mathbb{P}$ ,  $\omega$ , etc. The following two propositions are crucial points of the proof of Theorem 3.2.5.

**Proposition 3.2.7.** *Suppose that Conditions 3.2.1 – 3.2.4 are satisfied. Then, for sufficiently small  $d > 0$  there is a constant  $K > 0$  such that, for any pair of vectors  $v, v' \in \mathcal{A}$  with  $\|v - v'\| \leq d$ , we have*

$$\mathbb{P}\{\|\mathcal{R}(v, v') - \mathcal{R}'(v, v')\| \leq \gamma \|v - v'\|\} \geq 1 - K \|v - v'\|, \quad (3.62)$$

$$\mathbb{P}\{\|\Phi_k(v, v') - \Phi'_k(v, v')\| \leq \gamma^k \|v - v'\| \text{ for } k \geq 0\} \geq 1 - K_1 \|v - v'\|, \quad (3.63)$$

where  $K_1 = \frac{K}{1-\gamma}$ .

*Proof.* We first prove (3.62). The definition of the operators  $\mathcal{R}$  and  $\mathcal{R}'$  implies that if  $v, v' \in \mathcal{A}$  and  $\|v - v'\| \leq d$ , then

$$\|\mathcal{R}(v, v') - \mathcal{R}'(v, v')\| \leq \|V(v, v') - V'(v, v')\| + \|\mathbf{Q}(S(v) - S(v'))\|. \quad (3.64)$$

In view of (3.51) and (3.58), the first term on the right-hand side of (3.64) vanishes with probability no less than  $1 - K \|v - v'\|$ . Furthermore, it follows from (3.55) that

$$\|\mathbf{Q}(S(v) - S(v'))\| \leq \gamma \|v - v'\|.$$

Combining this with (3.64), we arrive at (3.62).

We now iterate (3.62) to prove (3.63). Let us assume that  $d > 0$  is so small that  $K_1 d \leq 1/2$ . For any integer  $k \geq 1$ , we set

$$G_k = \{\|\Phi_k(v, v') - \Phi'_k(v, v')\| \leq \gamma \|\Phi_{k-1}(v, v') - \Phi'_{k-1}(v, v')\|\},$$

$$\overline{G}_k = \bigcap_{l=1}^k G_l, \quad k \leq +\infty.$$

It is clear that the event on the left-hand side of (3.63) coincides with  $\overline{G}_\infty$ . Since  $\overline{G}_1 \supset \overline{G}_2 \supset \dots$ , inequality (3.63) will be established once we show that

$$\mathbb{P}(\overline{G}_k) \geq 1 - K \|v - v'\| \sum_{l=0}^{k-1} \gamma^l =: \varkappa_k \quad \text{for each } k \geq 1. \quad (3.65)$$

The proof of (3.65) is by induction in  $k$ . For  $k = 1$ , inequality (3.65) coincides with (3.62). Assume now that (3.65) is established for some  $k$ . Let us write  $\omega = (\omega_k, \omega'_k)$ , where  $\omega_k := (\omega^1, \dots, \omega^k)$  and  $\omega'_k := (\omega^j, j \geq k+1)$ , and denote

<sup>2</sup>In other words, the laws of the components of the random variables  $\mathbf{u}_k = \Phi_k(\omega)(v, v')$  coincide with  $P_k(v, \cdot)$  and  $P_k(v', \cdot)$  for any  $v, v' \in \mathcal{A}$ .

by  $\mathbb{P}_k$  the projections of  $\mathbb{P}$  to the first  $k$  components of  $\omega$ . Since  $\Phi_k$  depends only on  $\omega_k$ , the event  $\overline{G}_k$  can be written as

$$\overline{G}_k = \{\omega = (\omega_k, \omega'_k) \in \Omega : \omega_k \in g_k\},$$

where  $g_k$  is a measurable subset of the direct product of  $k$  copies of the probability space constructed in Lemma 3.2.6. By the induction hypothesis, we have  $\mathbb{P}^k(g_k) \geq \varkappa_k$ , and by (3.62), for each  $\omega_k \in g_k$  the probability of the event  $\{\omega^{k+1} : (\omega_k, \omega^{k+1}) \in g_{k+1}\}$  is minorised by  $1 - K\gamma^k\|v - v'\|$ . Therefore,

$$\mathbb{P}_{k+1}(g_{k+1}) \geq \varkappa_k(1 - K\gamma^k\|v - v'\|) > \varkappa_{k+1},$$

whence we conclude that (3.65) holds with  $k$  replaced by  $k + 1$ . This completes the proof of the proposition.  $\square$

**Proposition 3.2.8.** *Suppose that Conditions 3.2.1 – 3.2.4 are satisfied and that  $2K_1d \leq 1$ , where  $K_1$  is the constant constructed in Proposition 3.2.7. Then for any  $\delta > 0$  there is an integer  $\ell = \ell_\delta \geq 1$  and a constant  $\theta_\delta > 0$  such that*

$$\mathbb{P}\{\|\Phi_\ell(v, v') - \Phi'_\ell(v, v')\| \leq \delta\} \geq \theta_\delta \quad \text{for any } v, v' \in \mathcal{A}. \quad (3.66)$$

*Proof.* We fix  $\delta > 0$  and assume, without loss of generality, that  $\delta \leq d \wedge (2K_1)^{-1}$ . Let us introduce the stopping time

$$\tau = \tau(v, v') = \min\{k \geq 0 : \|\Phi_k(v, v') - \Phi'_k(v, v')\| \leq d\}.$$

Choose  $m \geq 0$  so large that  $d\gamma^m \leq \delta$ . In view of the strong Markov property and inequality (3.63), for any integer  $\ell \geq 1$  we have

$$\begin{aligned} \mathbb{P}\{\|\Phi_{\ell+m} - \Phi'_{\ell+m}\| \leq \delta\} &\geq \mathbb{P}(\{\tau \leq \ell\} \cap \{\|\Phi_{\ell+m} - \Phi'_{\ell+m}\| \leq \delta\}) \\ &= \mathbb{E}(\mathbb{I}_{\{\tau \leq \ell\}} \mathbb{P}(\{\|\Phi_{\ell+m} - \Phi'_{\ell+m}\| \leq \delta\} | \mathcal{F}_\tau)) \\ &\geq (1 - K_1d) \mathbb{P}\{\tau \leq \ell\} \\ &\geq \frac{1}{2} \mathbb{P}\{\tau \leq \ell\}. \end{aligned}$$

Thus, inequality (3.66) will be established if we show that

$$\mathbb{P}\{\tau(v, v') > \ell\} \leq p \quad \text{for all } v, v' \in \mathcal{A}, \quad (3.67)$$

where  $\ell \geq 1$  is a suitable integer and  $p < 1$ .

To prove (3.67), we use inequality (3.52) with  $r = d/4$  to find an integer  $\ell \geq 1$  such that

$$\|S^\ell(u)\| \leq \frac{d}{4} \quad \text{for all } u \in \mathcal{A}.$$

The Lipschitz continuity of  $S$  on bounded subsets implies that if  $\varepsilon > 0$  is sufficiently small and  $\|\eta_k\| \leq \varepsilon$  for  $k = 1, \dots, \ell$ , then the trajectory  $\{u_k\}$  of (3.50) issued from any point  $u \in \mathcal{A}$  satisfies the inequality  $\|u_\ell\| \leq d/2$ . By Condition 3.2.4, we have  $\nu = \mathbb{P}\{\|\eta_1\| \leq \varepsilon\} > 0$ . Let us define random variables  $\zeta_k(v, v')$  and  $\zeta'_k(v, v')$  by the relations

$$\zeta_k = \Phi_k - S(\Phi_{k-1}), \quad \zeta'_k = \Phi'_k - S(\Phi'_{k-1}), \quad k \geq 1.$$

The construction implies that each of the sequences  $\{\zeta_k\}$  and  $\{\zeta'_k\}$  consists of i.i.d. random variables whose distributions coincide with that of  $\eta_1$ . Let us consider the events

$$G = \{\|\zeta_k\| \leq \varepsilon, 1 \leq k \leq \ell\}, \quad G' = \{\|\zeta'_k\| \leq \varepsilon, 1 \leq k \leq \ell\}.$$

What has been said above implies that  $\mathbb{P}(G) = \mathbb{P}(G') = \nu^\ell$  and

$$\|\Phi_\ell(v, v')\| \vee \|\Phi'_\ell(v, v')\| \leq d/2 \quad \text{for } \omega \in G \cap G',$$

whence it follows that

$$\|\Phi_\ell(v, v') - \Phi'_\ell(v, v')\| \leq d \quad \text{for } \omega \in G \cap G'.$$

Thus, we see that

$$\tau(v, v') \leq \ell \quad \text{for } \omega \in G \cap G' \text{ and any } v, v' \in \mathcal{A}. \quad (3.68)$$

On the other hand, the definition of  $\Phi_k$  and  $\Phi'_k$  implies that if  $\omega \in \{\tau > \ell\}$ , then  $\zeta_k = \zeta'_k$  for  $1 \leq k \leq \ell$ . It follows that

$$G \cap G' \cap \{\tau > \ell\} = G \cap \{\tau > \ell\} = G' \cap \{\tau > \ell\}. \quad (3.69)$$

Assume that  $\mathbb{P}\{\tau > \ell\} > 1 - \nu^\ell$ . Then the event in (3.69) is not empty, which contradicts (3.68). So (3.67) is proved with  $p = 1 - \nu^\ell$ .  $\square$

### 3.2.3 Proof of Theorem 3.2.5

*Step 1.* Let us recall that the set  $\mathcal{A}$  defined by (3.54) is invariant and absorbing for the RDS (3.50). Therefore, by Lemma 1.3.30, the support of any stationary measure is contained in  $\mathcal{A}$ . Let us fix any stationary measure  $\mu$  and suppose we have proved that

$$\|P_k(u, \cdot) - \mu\|_L^* \leq C e^{-\alpha_0 k} \quad \text{for } k \geq 0, u \in \mathcal{A}, \quad (3.70)$$

where  $C$  and  $\alpha_0$  are some positive constants. In this case, if  $\mu'$  is another stationary measure, then by the Lebesgue theorem on dominated convergence, for any  $f \in L_b(H)$  we have

$$(f, \mu') = (f, \mathfrak{P}_k^* \mu') = (\mathfrak{P}_k f, \mu') = \int_H (f, P_k(u, \cdot)) \mu'(du) \rightarrow (f, \mu) \quad \text{as } k \rightarrow \infty.$$

Thus,  $(f, \mu) = (f, \mu')$  for any  $f \in L_b(H)$ , and therefore  $\mu' = \mu$ .

We now prove that (3.70) implies (3.57). By Condition 3.2.2, we have

$$u_k \in \mathcal{A} \quad \text{for any } u \in H \text{ and } k \geq k_0(\|u\|).$$

Therefore, the support of the measure  $P_k(u, \cdot)$  is contained in  $\mathcal{A}$  for  $k \geq k_0(\|u\|)$ . Combining this with inequality (3.70) and the Kolmogorov–Chapman relation,



we obtain

$$\begin{aligned}
|(f, P_k(u, \cdot)) - (f, \mu)| &= \left| \int_H P_{k_0}(u, dz) ((f, P_{k-k_0}(z, \cdot)) - (f, \mu)) \right| \\
&\leq \int_H P_{k_0}(u, dz) \sup_{z \in \mathcal{A}} |(f, P_{k-k_0}(z, \cdot)) - (f, \mu)| \\
&\leq C e^{-\alpha_0(k-k_0)}, \tag{3.71}
\end{aligned}$$

where  $f \in L_b(H)$  is any function with  $\|f\|_L \leq 1$  and  $k \geq k_0 = k_0(\|u\|)$ . Furthermore, inequality (3.71) remains valid for  $0 \leq k < k_0$ , provided that  $C \geq 2$ . It follows that, for any  $u \in H$ ,  $k \geq 0$ , and  $\alpha \in (0, \alpha_0]$ , we have

$$|(f, P_k(u, \cdot)) - (f, \mu)| \leq C e^{-\alpha(k-k_0)}.$$

Integrating this inequality with respect to  $\lambda(du)$  and taking the supremum over  $f \in L_b(H)$  with  $\|f\|_L \leq 1$ , we arrive at (3.57).

*Step 2.* We now prove (3.70). To this end, we apply Theorem 3.1.1, in which  $X = \mathcal{A}$ , and the Markov family in question is the one generated by the restriction of the RDS (3.50) to  $\mathcal{A}$ . Since the metric space  $\mathcal{A}$  is bounded, inequality (3.12) is trivial, and we only need to check (3.13).

Let us choose a small constant  $d > 0$  and set  $d_{-1} = +\infty$  and  $d_m = \gamma^m d$  for  $m \geq 0$ , where  $\gamma > 0$  is defined in Proposition 3.2.7. We now define a Kantorovich density  $F$  by the relation

$$F(u_1, u_2) = \begin{cases} R & \text{for } \|u_1 - u_2\| > d_0, \\ R^2 d_m & \text{for } d_{m+1} < \|u_1 - u_2\| \leq d_m, \end{cases}$$

where  $m \geq 0$ , and  $R \geq 1$  stands for a constant that will be chosen later in such a way that  $Rd \leq 1/2$ , so that  $F(u_1, u_2) \leq R$ . We claim that if  $d > 0$  is sufficiently small, then

$$\mathcal{K}_F(\mathfrak{P}_\ell^* \lambda, \mathfrak{P}_\ell^* \lambda') \leq \gamma' \mathcal{K}_F(\lambda, \lambda') \quad \text{for any } \lambda, \lambda' \in \mathcal{P}(\mathcal{A}), \tag{3.72}$$

where  $\ell \geq 1$  is the integer defined in Proposition 3.2.8 for  $\delta = d$ , and  $\gamma' < 1$  is a constant not depending on the measures  $\lambda$  and  $\lambda'$ . Once this inequality is established, the required result will follow from Theorem 3.1.1.

For any pair of vectors  $(v, v') \in H \times H$ , we set

$$f(v, v') = \mathbb{E} F(\Phi_\ell(v, v'), \Phi'_\ell(v, v')),$$

where the RDS  $\Phi_k = (\Phi_k, \Phi'_k)$  is defined in Section 3.2.2. Define the subsets

$$\mathcal{A}_m = \{(v, v') \in \mathcal{A} \times \mathcal{A} : d_{m+1} < \|v - v'\| \leq d_m\}, \quad m \geq -1.$$

Since  $F \leq R$ , for  $(v, v') \in \mathcal{A}_m$  we have

$$\begin{aligned}
f(v, v') &\leq R^2 d_{m+1} \mathbb{P}(Q_m(v, v')) + R(1 - \mathbb{P}(Q_m(v, v'))) \\
&= R^2 d_{m+1} P_m(v, v') + R(1 - P_m(v, v')),
\end{aligned}$$

where  $m \geq -1$ ,  $Q_m(v, v') = \{\omega \in \Omega : \|\Phi_\ell(v, v') - \Phi'_\ell(v, v')\| \leq d_{m+1}\}$ , and  $P_m(v, v') = \mathbb{P}(Q_m(v, v'))$ . It follows that if  $(u, u')$  is a coupling for the pair of measures  $\lambda, \lambda' \in \mathcal{P}(\mathcal{A})$  which is independent of  $\{\Phi_k, k \geq 1\}$ , then

$$\begin{aligned} \mathbb{E} F(\Phi_\ell(u, u'), \Phi'_\ell(u, u')) &= \sum_{m=-1}^{\infty} \mathbb{E} \left\{ \mathbb{I}_{\{(u, u') \in \mathcal{A}_m\}} f(u, u') \right\} \\ &\leq \sum_{m=-1}^{\infty} \mathbb{E} \left\{ \mathbb{I}_{\{(u, u') \in \mathcal{A}_m\}} \left( R^2 d_{m+1} P_m(u, u') + R(1 - P_m(u, u')) \right) \right\} \\ &= R \mathbb{E} \left\{ \mathbb{I}_{\{(u, u') \in \mathcal{A}_{-1}\}} (1 - P_{-1}(u, u')(1 - Rd)) \right\} \\ &\quad + \sum_{m=0}^{\infty} R^2 d_m \mathbb{E} \left\{ \mathbb{I}_{\{(u, u') \in \mathcal{A}_m\}} (\gamma P_m(u, u') + (Rd_m)^{-1}(1 - P_m(u, u'))) \right\}. \end{aligned}$$

By (3.66) and (3.63) for  $(u, u') \in \mathcal{A}_{-1}$  and  $(u, u') \in \mathcal{A}_m$ ,  $m \geq 0$ , we have  $P_{-1} \geq \theta_d$  and  $P_m \geq 1 - K_1 d$ . So choosing  $R$  and  $d$  so that  $\gamma + R^{-1}K_1 < 1$  and  $2Rd \leq 1$ , we derive

$$\begin{aligned} \mathbb{E} F(\Phi_\ell(u, u'), \Phi'_\ell(u, u')) &\leq R(1 - \theta_d(1 - Rd)) \mathbb{P}\{(u, u') \in \mathcal{A}_{-1}\} \\ &\quad + \sum_{m=0}^{\infty} R^2 d_m (\gamma + R^{-1}K_1) \mathbb{P}\{\{(u, u') \in \mathcal{A}_m\} \leq \gamma' \mathbb{E} F(u, u')\}, \end{aligned} \quad (3.73)$$

where  $\theta_d$  is constructed in Proposition 3.2.8, and  $\gamma' < 1$  is a constant not depending on  $u$  and  $u'$ . Since inequality (3.73) is true for some coupling  $(u, u')$  of the pair of measures  $(\lambda, \lambda')$ , taking the infimum over all couplings, we arrive at (3.72). The proof of the theorem is complete.

### 3.2.4 Application to Navier–Stokes equations

Let us consider the homogeneous Navier–Stokes system perturbed by a random kick force (cf. (2.76)):

$$\dot{u} + \nu Lu + B(u) = \sum_{k=1}^{\infty} \eta_k \delta(t - kT). \quad (3.74)$$

Here  $\{\eta_k\}$  is a sequence of i.i.d. random variables in  $H$  of the form

$$\eta_k(x) = \sum_{j=1}^{\infty} b_j \xi_{jk} e_j(x), \quad (3.75)$$

where  $\{e_j\}$  stands for an orthonormal basis in  $H$  consisting of the eigenfunctions of the Stokes operator  $L$ ,  $\xi_{jk}$  are independent scalar random variables whose ranges are included in  $[-1, 1]$  and laws are independent of  $k$ , and  $b_j \geq 0$  are some constants such that

$$\mathfrak{B} := \sum_{j=1}^{\infty} b_j^2 < \infty. \quad (3.76)$$

As was explained in Section 2.3, the restriction of trajectories for (3.74) to  $T\mathbb{Z}_+$  satisfies relation (3.50), where  $u_k = u(kT)$ , and  $S : H \rightarrow H$  stands for the operator that takes  $u_0$  to the solution at time  $T$  of the homogeneous Navier–Stokes system. We denote by  $\Phi = \{\varphi_k, k \geq 0\}$  the Markov RDS associated with (3.50) and by  $\mathfrak{P}_k$  and  $\mathfrak{P}_k^*$  the corresponding Markov semigroups; see Section 1.3.3. Recall that a function  $p : [-1, 1] \rightarrow \mathbb{R}$  is said to have a *bounded variation* if<sup>3</sup>

$$\text{Var}(p) = \sup_{\{r_j\}} \sum_{j=1}^m |p(r_j) - p(r_{j-1})|, \quad (3.77)$$

where the supremum is taken over all partitions  $r_0 = -1 < r_1 < \dots < r_m = 1$  of the interval  $[-1, 1]$ .

**Theorem 3.2.9.** *Under the above conditions, assume that the laws of  $\xi_{jk}$  are absolutely continuous with respect to the Lebesgue measure, and the corresponding densities  $p_j$  are functions of bounded variation such that*

$$\int_{-\varepsilon}^{\varepsilon} p_j(r) dr > 0 \quad \text{for any } \varepsilon > 0. \quad (3.78)$$

*In this case, there is an integer  $N = N(\mathfrak{B}, \nu) \geq 1$  such that if*

$$b_j \neq 0 \quad \text{for } j = 1, \dots, N, \quad (3.79)$$

*then the RDS  $\Phi$  has a unique stationary measure  $\mu \in \mathcal{P}(H)$ . Moreover, there are positive constants  $C$  and  $\alpha$  such that*

$$\|\mathfrak{P}_k^* \lambda - \mu\|_L^* \leq C e^{-\alpha k} \left(1 + \int_H |u|_2 \lambda(du)\right), \quad k \geq 0, \quad (3.80)$$

*where  $\lambda \in \mathcal{P}(H)$  is an arbitrary measure with finite first moment.*

For example, the assumptions of Theorem 3.2.9 are met for any  $\nu > 0$  if the random variables  $\xi_{jk}$  are independent, each  $\xi_{jk}$  is uniformly distributed on a non-degenerate segment  $[r'_j, r''_j] \subset [-1, 1]$  containing the point  $r = 0$ , all numbers  $b_j$  are non-zero, and  $\mathfrak{B} < \infty$ .

*Proof.* We shall show that the Markov family generated by  $\Phi$  satisfies the hypotheses of Theorem 3.2.5. It follows from assertion (i) of Proposition 2.1.25 and inequality (2.24) that the operator  $S : H \rightarrow H$  is Lipschitz-continuous on bounded subsets. Moreover, inequality (2.51) implies that

$$|S(u)|_2 \leq q|u|_2 \quad \text{for any } u \in H, \quad (3.81)$$

where  $q = \exp(-\alpha_1 T/2)$ . Thus,  $S$  satisfies Condition 3.2.1.

<sup>3</sup>Equivalently, the derivative  $p'$  of the function  $p(x)$  in the sense of distributions is a signed measure, and  $\text{Var}(p) = 2\|p'\|_{\text{var}}$  (see (1.13)).

We now check Condition 3.2.2. It follows from (3.76) that  $|\eta_k|_2 \leq \sqrt{\mathfrak{B}}$  almost surely, and therefore the support of the law for  $\eta_k$  is contained in a ball  $B_H(r)$ . Combining this with inequality (3.81), we see that

$$|\varphi_1 u|_2 \leq q|u|_2 + r \quad \text{for any } u \in H \text{ and almost every } \omega \in \Omega.$$

It follows that

$$\sup_{u \in \mathcal{A}_k(B_H(R))} |u|_2 \leq q^k R + (1 - q)^{-1} r.$$

We conclude that Condition 3.2.2 holds with

$$\rho = 2r(1 - q)^{-1}, \quad k_0(R) = \lceil (\ln q^{-1})^{-1} \ln(\frac{2R}{r} + 1) \rceil + 1, \quad (3.82)$$

where  $\lceil a \rceil$  stands for the integer part of  $a$ .

Let us check Condition 3.2.3. To this end, we use the regularising property of  $S$ . Inequality (2.55) implies that

$$\|S(u_1) - S(u_2)\|_1 \leq C_\nu |u_1 - u_2|_2 \quad \text{for any } u_1, u_2 \in \mathcal{A},$$

where  $\mathcal{A}$  is defined by (3.54). Combining this with Poincaré's inequality, we see that

$$\begin{aligned} |(I - P_N)(S(u_1) - S(u_2))|_2 &\leq \frac{1}{\sqrt{\alpha_{N+1}}} \|S(u_1) - S(u_2)\|_1 \\ &\leq \frac{C_\nu}{\sqrt{\alpha_{N+1}}} |u_1 - u_2|_2, \end{aligned}$$

where  $\alpha_j$  denotes the eigenvalue of  $L$  corresponding to an eigenvector  $e_j$ , and  $P_N$  stands for the orthogonal projection to the space  $E := \text{span}\{e_1, \dots, e_N\}$ . Since  $\alpha_j \rightarrow \infty$  as  $j \rightarrow \infty$ , we conclude that Condition 3.2.3 is satisfied for  $P = P_N$  with a sufficiently large  $N \geq 1$ .

It remains to verify Condition 3.2.4. We first show that the support of  $\mathcal{D}(\eta_1)$  contains the origin. To this end, it suffices to prove that

$$\mathbb{P}\{|\eta_1|_2 < \delta\} > 0 \quad \text{for any } \delta > 0.$$

But this relation readily follows from (3.76) and (3.78). Indeed, if  $m$  is sufficiently large, then (3.76) implies that  $|(I - P_m)\eta_1|_2 \leq \delta/2$  for almost all  $\omega$ , and it follows from (3.78) that  $\mathbb{P}\{|P_m \eta_1|_2 < \delta/2\} > 0$  for any  $m \geq 1$ .

To check the second part of Condition 3.2.4, we note that the law of  $P_N \eta_1$  possesses a density (with respect to the Lebesgue measure on  $E$ ) given by the relation

$$p(x) = \prod_{j=1}^N q_j(x_j), \quad q_j(x_j) = b_j^{-1} p_j(x_j/b_j), \quad x = (x_1, \dots, x_N) \in E. \quad (3.83)$$

To prove (3.56), we first assume that  $p_j$  are  $C^1$ -smooth functions. In this case, we have

$$\begin{aligned} \int_E |p(x+y) - p(x)| dx &\leq |y| \int_E \int_0^1 |(\nabla p)(x + \theta y)| d\theta dx \\ &= |y| \int_E |(\nabla p)(x)| dx \leq |y| \sum_{j=1}^N \int_{\mathbb{R}} |\partial_{x_j} q_j(x_j)| dx_j \\ &= |y| \sum_{j=1}^N \text{Var}(q_j) = |y| \sum_{j=1}^N b_j^{-1} \text{Var}(p_j). \end{aligned}$$

In the general case, inequality (3.56) can be established by a standard approximation argument.

We have thus shown that the conclusion of Theorem 3.2.5 holds for the RDS (3.50) corresponding to the kicked Navier–Stokes system (3.74). This proves the uniqueness of a stationary measure and inequality (3.57). To establish (3.80), we choose  $\alpha \in (0, \alpha_0]$  such that  $\alpha \leq \ln q^{-1}$ . In this case, taking into account the second relation in (3.82), we see that

$$\exp(\alpha k_0(R)) \leq C_1(R+1).$$

Substituting this inequality into (3.57), we arrive at (3.80). The proof of Theorem 3.2.9 is complete.  $\square$

### 3.3 Navier–Stokes system perturbed by white noise

In this section, we consider the stochastic Navier–Stokes system in a bounded domain  $Q \subset \mathbb{R}^2$  (or on a torus  $\mathbb{R}^2/(a\mathbb{Z} \oplus b\mathbb{Z})$ , assuming that the space-mean values of the force and solutions vanish). Namely, we study the problem

$$\dot{u} + \nu Lu + B(u) = h + \eta(t), \quad (3.84)$$

$$u(0) = u_0, \quad (3.85)$$

where  $\nu > 0$ ,  $h \in H$  is a deterministic function, and  $\eta$  is a random process of the form (2.66), i.e.,

$$\eta(t, x) = \frac{\partial}{\partial t} \sum_{j=1}^{\infty} b_j \beta_j(t) e_j(x),$$

where  $\mathfrak{B} = \sum_j b_j^2 < \infty$ . We shall denote by  $u(t; u_0)$  the unique solution of (3.84), (3.85) constructed in Theorem 2.4.6. The main result of this section claims that, under some non-degeneracy assumptions, the Markov process associated with (3.84) has a unique stationary measure  $\mu$ , and all the trajectories of (3.84) converge to  $\mu$  in law. Among a number of theorems of this kind, we have chosen the one with the simplest proof. Stronger results and those with different assumptions on the random force  $\eta$  are discussed in Section 3.5.

### 3.3.1 Main result and scheme of its proof

Let us denote by  $(u_t, \mathbb{P}_u)$  the Markov family in  $H$  associated with the Navier–Stokes system (3.84) and denote by  $\mathfrak{P}_t$  and  $\mathfrak{P}_t^*$  the corresponding Markov semi-groups. Recall that if  $u_0$  is an  $H$ -valued random variable independent of  $\{\beta_j\}$ , then the law of the solution for (3.84), (3.85) coincides with  $\mathfrak{P}_t^* \lambda$ , where  $\lambda$  is the distribution of  $u_0$ .

**Theorem 3.3.1.** *Let us assume that the basis  $\{e_j\}$  in the definition of the process  $\eta$  consists of the eigenfunctions for the Stokes operator  $L$  and the coefficients  $b_j$  satisfies (2.67). Then the Markov family  $(u_t, \mathbb{P}_u)$  has a unique stationary measure  $\mu \in \mathcal{P}(H)$ , provided that  $b_j \neq 0$  for all  $j \geq 1$ . Moreover, for any compact subset  $\Lambda$  of the space  $\mathcal{P}(H)$  endowed with the dual-Lipschitz metric we have*

$$\sup_{\lambda \in \Lambda} \|\mathfrak{P}_t^* \lambda - \mu\|_L^* \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (3.86)$$

This theorem remains true under the weaker hypothesis that  $b_j \neq 0$  for  $j = 1, \dots, N$ , with a sufficiently large  $N$ . Indeed, the only point where we use that all the coefficients  $b_j$  (rather than finitely many of them) are non-zero is Lemma 3.3.11. However, we need the conclusion of that lemma to be true only for the family of balls  $B$  centred at a given point  $\hat{u} \in H$ . In this situation, finitely many non-zero components of the noise  $\eta$  are sufficient to be able to prove that, for an appropriately chosen function  $\hat{u} \in H$ , the probability of transition from a given point to any neighbourhood of  $\hat{u}$  is positive. For instance, if  $h = 0$ , then the function  $\hat{u} = 0$  is a globally exponentially stable fixed point for the unperturbed dynamics (see Corollary 2.1.24), and it is not difficult to see that the probability of transition from any point to an arbitrary small neighbourhood of origin is positive. We refer the reader to Exercise 3.3.12 for some more hints in the case of a non-zero function  $h$  and finitely many non-zero coefficients.

Convergence (3.86) and the a priori estimates for solutions of the Navier–Stokes system imply that for a large class of functionals  $f : H \rightarrow \mathbb{R}$  the following assertion holds: the mean value of  $f$  calculated at a solution at time  $t$  converges, as  $t \rightarrow \infty$ , to the mean value of  $f$  with respect to the stationary distribution. Namely, denote by  $\mathcal{W}$  the space of continuous functions  $w(r) > 0$  that are defined and non-decreasing on the positive half-line. In what follows, the elements of  $\mathcal{W}$  will be called *weight functions*. For any  $w \in \mathcal{W}$ , let  $C(H, w)$  be the space of continuous functionals  $f : H \rightarrow \mathbb{R}$  such that

$$|f|_w := \sup_{u \in H} \frac{|f(u)|}{w(|u|_2)} < \infty. \quad (3.87)$$

**Corollary 3.3.2.** *Under the hypotheses of Theorem 3.3.1, there is  $\varkappa > 0$  such that if  $f \in C(H, w)$  with  $w(r) = \exp(\varkappa r^2)$ , then*

$$\mathbb{E} f(u(t; v)) \rightarrow \int_H f(z) \mu(dz) \quad \text{as } t \rightarrow +\infty \text{ for any } v \in H. \quad (3.88)$$

*In particular, this convergence holds for the energy functional  $E(u) = \frac{1}{2}|u|_2^2$ .*

*Proof.* Let us denote by  $\varkappa_1$  and  $\varkappa_2$  the constants  $\varkappa > 0$  defined in Proposition 2.4.9 and Theorem 2.5.5, respectively, and let  $\varkappa = (\varkappa_1/2) \wedge \varkappa_2$ . Then for  $f \in C(H, w)$  the quantities  $\mathbb{E}f(u(t; v))$  and  $(f, \mu)$  are well defined in view of (2.125) and (2.157).

Convergence (3.86) with  $\lambda = \delta_v$  and assertion (ii) of Theorem 1.2.15 imply that

$$\mathfrak{P}_t g(v) = \mathbb{E}g(u(t; v)) \rightarrow (g, \mu) \quad \text{as } t \rightarrow +\infty, \quad (3.89)$$

where  $g \in C_b(H)$  is an arbitrary function. For any  $R > 0$ , let  $\chi_R \in C_b(H)$  be a function such that  $0 \leq \chi_R \leq 1$ ,  $\chi_R(u) = 1$  for  $|u|_2 \leq R$ , and  $\chi_R(u) = 0$  for  $|u|_2 \geq R + 1$ . Then for any  $f \in C(H, w)$  we can write

$$\mathbb{E}f(u(t; v)) = (\mathfrak{P}_t f_{\leq R})(u) + \mathbb{E}f_{\geq R}(u(t; v)), \quad (3.90)$$

where we set  $f_{\leq R} = \chi_R f$  and  $f_{\geq R} = (1 - \chi_R)f$ . Note that  $f_{\leq R} \in C_b(H)$ . By (3.89), the first term in the right-hand side of (3.90) satisfies the inequality

$$|(\mathfrak{P}_t f_{\leq R})(u) - (f_{\leq R}, \mu)| \leq \varepsilon_1(t, R, u),$$

where  $\varepsilon_1(t, R, u) \rightarrow 0$  as  $t \rightarrow \infty$  for any  $R > 0$  and  $u \in H$ . Furthermore, it follows from Proposition 2.4.9 and Chebyshev's inequality that

$$\begin{aligned} |\mathbb{E}f_{\geq R}(u(t; v))| &\leq |\mathbb{E}\{(1 - \chi_R(u(t; v))) \exp(\varkappa\nu|u(t; v)|_2^2)\}| \\ &\leq (\mathbb{P}\{|u(t; v)|_2 > R\})^{1/2} (\mathbb{E} \exp(2\varkappa\nu|u(t; v)|_2^2))^{1/2} \\ &\leq C(|u|_2)\varepsilon_2(R), \end{aligned}$$

where  $t \geq 0$  is arbitrary, and  $\varepsilon_2(R) \rightarrow 0$  as  $R \rightarrow \infty$ . We have thus shown that

$$\left| \mathbb{E}f(u(t; v)) - \int_H f(z)\mu(dz) \right| \leq |(f_{\geq R}, \mu)| + \varepsilon_1(t, R, u) + C(|u|_2)\varepsilon_2(R).$$

This inequality implies the required convergence, since  $(f_{\geq R}, \mu) \rightarrow 0$  as  $R \rightarrow \infty$  in view of Theorem 2.5.5 and Lebesgue's theorem on dominated convergence.  $\square$

Theorem 3.3.1 applies to the Navier–Stokes system (3.84) in a bounded domain or on the torus. In the latter case, we can consider the unique stationary distribution  $\mu$  as a measure on the space of locally square-integrable functions on  $\mathbb{R}^2$  that are  $2\pi$ -periodic in both variables. Hence, one may ask a question about the invariance of  $\mu$  with respect to the translations in space. It turns out that this is indeed true, provided that the random force is invariant. Namely, let us consider the white-forced Navier–Stokes system (3.84), in which<sup>4</sup>  $h \equiv 0$  and the coefficients of the random process

$$\eta(t, x) = \frac{\partial}{\partial t} \zeta(t, x), \quad \zeta(t, x) = \sum_{s \in \mathbb{Z}_0^2} b_s \beta_s(t) e_s(x),$$

<sup>4</sup>Note that the Dirac measure concentrated at a function  $h(x)$  is invariant with respect to translations if and only if  $h$  is constant. Since the mean value of  $h$  over the torus must be zero, the assumption that  $h \equiv 0$  is necessary for the law of the right-hand side of (3.84) to be invariant.

satisfy the relations

$$b_s = b_{-s} \quad \text{for all } s \in \mathbb{Z}_0^2. \quad (3.91)$$

We claim that in this case the law of the process  $\zeta(t, x)$  is invariant under translations with respect to  $x$ . Indeed, if  $a \in \mathbb{R}^2$  and  $T_a : \mathbb{T}^2 \rightarrow \mathbb{T}^2$  denotes the translation operator taking  $x$  to  $x + a \pmod{2\pi}$ , then

$$\begin{aligned} \zeta(t, T_a x) &= \sum_{s \in \mathbb{Z}_+^2} c_s b_s s^\perp (\beta_s(t) \cos\langle s, T_a x \rangle + \beta_{-s}(t) \sin\langle s, T_a x \rangle) \\ &= \sum_{s \in \mathbb{Z}_+^2} c_s b_s s^\perp (\tilde{\beta}_s(t) \cos\langle s, x \rangle + \tilde{\beta}_{-s}(t) \sin\langle s, x \rangle), \end{aligned}$$

where  $c_s$  and  $\mathbb{Z}_+^2$  are defined in Section 2.1.5,

$$\begin{aligned} \tilde{\beta}_s(t) &= \beta_s(t) \cos\langle s, a \rangle + \beta_{-s}(t) \sin\langle s, a \rangle, \\ \tilde{\beta}_{-s}(t) &= -\beta_s(t) \sin\langle s, a \rangle + \beta_{-s}(t) \cos\langle s, a \rangle. \end{aligned}$$

The following property of the multidimensional Brownian motion is a simple consequence of the definition.

*Exercise 3.3.3.* Let  $B(t)$  be a  $d$ -dimensional Brownian motion, that is, an  $\mathbb{R}^d$ -valued stochastic process whose components are independent standard Brownian motions. Then for any orthogonal  $d \times d$  matrix  $U$  the process  $UB(t)$  is also a Brownian motion. *Hint:* Show that  $UB(t)$  is a Gaussian process with the same correlations as  $B(t)$ .

Applying this property with  $d = 2$ , we see that  $\tilde{\beta}_s$  and  $\tilde{\beta}_{-s}$  are independent Brownian motions. Thus, the processes  $\zeta(t, T_a x)$  and  $\zeta(t, x)$  have exactly the same form and, hence, the same laws. A similar argument proves that the laws of  $\zeta(t, -x)$  and  $\zeta(t, x)$  also coincide. The following assertion shows that the same properties hold for the stationary measure.

**Corollary 3.3.4.** *Let us assume that  $h \equiv 0$  and the coefficients  $b_s \neq 0$  are such that  $\sum_s b_s^2 < \infty$  and (3.91) holds. Then for any  $\nu > 0$  the unique stationary measure  $\mu_\nu$  of (3.84) is invariant under translations of  $\mathbb{R}^2$  and the reflection  $x \mapsto -x$ .*

*Proof.* We first prove the invariance of  $\mu_\nu$  under translations. Let us consider Eq. (3.84) in which  $\eta$  is replaced by  $\eta_a(t, x) = \eta(t, T_a x)$ . Then, by Theorem 3.3.1 applied to the torus, the new equation has a unique stationary distribution, which must coincide with  $\mu_\nu$ , because the laws of  $\eta$  and  $\eta_a$  are the same. On the other hand, the process  $u_\nu^a(t, x) = u_\nu(t, T_a x)$  is a stationary solution of Eq. (3.84) with  $\eta$  replaced by  $\eta_a$ . By the uniqueness of a stationary measure, the law of  $u_\nu^a(t)$  coincides with  $\mu_\nu$ . A similar argument shows that  $\mu_\nu$  is invariant under the reflection  $x \mapsto -x$ .  $\square$

*Example 3.3.5.* Let us assume that the hypotheses of Corollary 3.3.4 are satisfied and the unique stationary measure  $\mu \in \mathcal{P}(H)$  for (3.84) is concentrated on  $V^{k+1}$



with some integer  $k \geq 1$ . (This is indeed the case if  $h \in V^k$  and (2.137) holds; see Exercise 2.5.8.) Consider a functional of the form  $f_x(u) = g(u(x))$ , where  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a continuous function and  $x \in \mathbb{T}^2$  is a fixed point. The functional  $f_x$  is well defined on  $V^2$ , and therefore its mean value  $(f_x, \mu)$  has a sense. Moreover, the invariance of  $\mu$  with respect to space translations implies that the function  $x \mapsto (f_x, \mu)$  is constant. On the other hand, by Fubini's theorem, we have

$$\langle (f_x, \mu) \rangle = (2\pi)^{-2} \int_{\mathbb{T}^2} \int_{V^{k+1}} g(u(x)) \mu(du) dx = \int_{V^{k+1}} \langle g(u(\cdot)) \rangle \mu(du).$$

In particular, taking  $g(v^1, v^2) = v^j$  with  $j = 1$  or  $2$ , we see that  $\langle u^j(\cdot) \rangle = 0$  for any  $u \in H$  and, hence,

$$\int_{V^{k+1}} u^j(x) \mu(du) = 0 \quad \text{for } x \in \mathbb{T}^2, j = 1, 2.$$

Thus, the mean velocity in the unique stationary regime is equal to zero. Let us also note that we did not use the uniqueness, and the same conclusion is true for any space-invariant stationary distribution concentrated<sup>5</sup> on  $V^{k+1}$  with  $k \geq 1$ .

Let us outline now the proof of Theorem 3.3.1. We wish to apply Theorem 3.1.4. To this end, we shall construct an extension  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$  for the family  $(u_t, \mathbb{P}_u)$  and a decreasing sequence of closed subsets  $B_m \subset H$  such that the recurrence and stability properties of Theorem 3.1.3 hold, together with inequality (3.26) in which  $\mathbf{Y} \subset \mathbf{H}$  is an arbitrary compact subset.

Let us first describe a trivial extension for  $(u_t, \mathbb{P}_u)$  that will be called a *pair of independent copies*. We denote by  $(\Omega, \mathcal{F})$  the measurable space on which the family  $(u_t, \mathbb{P}_u)$  is defined and by  $(\mathbf{\Omega}, \mathbf{\mathcal{F}})$  the direct product of two copies  $(\Omega, \mathcal{F})$ :

$$\mathbf{\Omega} = \Omega \times \Omega, \quad \mathbf{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}.$$

The points of  $\mathbf{\Omega}$  will be denoted by  $\boldsymbol{\omega} = (\omega, \omega')$ . We now define a process  $\mathbf{u}_t(\boldsymbol{\omega})$  in the space  $\mathbf{H} = H \times H$  by the formula

$$\mathbf{u}_t(\boldsymbol{\omega}) = (u_t(\omega), u_t(\omega'))$$

and, for any  $\mathbf{u} = (u, u') \in \mathbf{H}$ , we denote  $\mathbb{P}_{\mathbf{u}} = \mathbb{P}_u \times \mathbb{P}_{u'}$ . Clearly, the family  $(\mathbf{u}_t, \mathbb{P}_{\mathbf{u}})$  is an extension for  $(u_t, \mathbb{P}_u)$ , and its transition function  $\mathbf{P}_t$  satisfies the relation

$$\mathbf{P}_t(\mathbf{u}, \Gamma \times \Gamma') = P_t(u, \Gamma) P_t(u', \Gamma') \quad \text{for } \mathbf{u} = (u, u') \in \mathbf{H}, \Gamma, \Gamma' \in \mathcal{B}(H), \quad (3.92)$$

where  $P_t(u, \Gamma)$  stands for the transition function of  $(u_t, \mathbb{P}_u)$ .

Let us denote by  $B_m \subset H$  the closed ball of radius  $1/m$  centred at zero and define the stopping time

$$\boldsymbol{\tau} = \boldsymbol{\tau}(m) = \min\{t \geq 0 : \mathbf{u}_t \in B_m \times B_m\}.$$

---

<sup>5</sup>The fact that  $\mu$  is concentrated on  $V^{k+1}$  is not really important: one can give a meaning to the mean value  $(f_x, \mu)$  for any space-invariant measure  $\mu \in \mathcal{P}(H)$ , and the same conclusion will be valid; see [VF88].

We shall show in the next two subsections that the following propositions are true.

**Proposition 3.3.6.** *Under the hypotheses of Theorem 3.3.1, for any integer  $m \geq 1$  there are two positive constants  $C$  and  $\alpha$  such that*

$$\mathbb{E}_u \exp(\alpha\tau) \leq C(1 + |u|_2^2 + |u'|_2^2) \quad \text{for all } u, u' \in H. \quad (3.93)$$

*In particular, the stopping time  $\tau$  is almost surely finite for any initial point  $u \in H$ .*

**Proposition 3.3.7.** *Under the hypotheses of Theorem 3.3.1, there is a sequence  $\delta_m > 0$  going to zero such that*

$$\sup_{t \geq 0} \|P_t(u, \cdot) - P_t(u', \cdot)\|_L^* \leq \delta_m \quad \text{for any } u, u' \in B_m. \quad (3.94)$$

Once these two propositions are established, the required results will follow from Theorem 3.1.3.

Analysing the proof of the theorem, it is not difficult to see that convergence is uniform with respect to the random perturbations  $\eta$  that are bounded and “uniformly non-degenerate”. Namely, we have the following result that can be established by repeating step by step the proof of Theorem 3.3.1.

*Exercise 3.3.8.* Let  $\nu > 0$  be a constant, let  $h \in H$  be a deterministic function, and let  $\{\hat{b}_j\}$  be a sequence of positive numbers such that  $\sum_j \hat{b}_j^2 < \infty$ . Prove that for any  $\mathfrak{B} > 0$  there is  $\delta > 0$  such that if

$$\sum_{j=1}^{\infty} b_j^2 \leq \mathfrak{B}, \quad \sup_{j \geq 1} |b_j - \hat{b}_j| \leq \delta,$$

then the Markov family  $(u_t, \mathbb{P}_u)$  associated with Eq. (3.84) has a unique stationary measure  $\mu \in \mathcal{P}(H)$ , and for any compact subset  $\Lambda \subset \mathcal{P}(H)$  we have

$$\sup_{\lambda \in \Lambda} \|\mathfrak{P}_t^* \lambda - \mu\|_L^* \leq \alpha_\Lambda(t) \rightarrow 0 \quad \text{as } t \rightarrow \infty,$$

where  $\alpha_\Lambda$  is a function depending only on  $\mathfrak{B}$  and  $\{\hat{b}_j\}$ .

Finally, it is sometimes useful to know that laws of solutions in the space of trajectories also converge to a limiting measure. Namely, let  $\tilde{u}(t)$ ,  $t \geq 0$ , be a stationary solution of (3.84) and let  $\tilde{\mu} \in \mathcal{P}(C(\mathbb{R}_+; H))$  be its law.

*Exercise 3.3.9.* Under the hypotheses of Theorem 3.3.1, show that for any solution  $u(t)$  of (3.84) we have

$$\mathcal{D}(u(t + \cdot)) \rightarrow \tilde{\mu} \quad \text{as } t \rightarrow \infty,$$

where the convergence holds in the weak topology of the space  $\mathcal{P}(C(\mathbb{R}_+; H))$ .

### 3.3.2 Recurrence: proof of Proposition 3.3.6

For any  $R > 0$ , define the stopping time

$$T_R = \min\{t \geq 0 : \|\mathbf{u}_t\| \leq R\},$$

where we set

$$\|\mathbf{u}\| = |\nabla \mathbf{u}|, \quad |\mathbf{u}|^2 = |u|_2^2 + |u'|_2^2.$$

The proof of Proposition 3.3.6 is based on the following two lemmas.

**Lemma 3.3.10.** *For any  $\nu > 0$ ,  $h \in H$ , and  $\mathfrak{B} \geq 0$  there is  $R > 0$  such that if the coefficients  $b_j \geq 0$  satisfy the inequality  $\sum_j b_j^2 \leq \mathfrak{B}$ , then*

$$\mathbb{E}_{\mathbf{u}} \exp(\gamma T_R) \leq 1 + K |\mathbf{u}|^2 \quad \text{for any } \mathbf{u} \in \mathbf{H}, \quad (3.95)$$

where  $K$  and  $\gamma$  are positive constants. In particular,  $\mathbb{P}_{\mathbf{u}}\{T_R < \infty\} = 1$  for any  $\mathbf{u} \in \mathbf{H}$ .

**Lemma 3.3.11.** *Under the hypotheses of Theorem 3.3.1, for any  $R > 0$  and any non-degenerate ball  $B \subset H$  there is  $p > 0$  such that*

$$P_1(\mathbf{u}, B \times B) \geq p \quad \text{for any } \mathbf{u}, \mathbf{u}' \in B_V(R). \quad (3.96)$$

Let us emphasise that, in the first lemma, some of the constants  $b_j$  may be zero, whereas in the second lemma, the condition that all of them are positive is crucial for the result to be true with an *arbitrary* non-degenerate ball. Note, however, that if we wish (3.96) to be true for any ball *centred at a fixed point*, then it suffices to assume that a large, but finite number of coefficients  $b_j$  are non-zero; see below Exercise 3.3.12. Taking the above lemmas for granted, let us complete the proof of the proposition. It is divided into three steps.

*Step 1.* Let us introduce an increasing sequence of stopping times  $\sigma'_n$  by the rule

$$\sigma'_0 = T_R, \quad \sigma'_n = \min\{t \geq \sigma'_{n-1} + 1 : \|\mathbf{u}_t\| \leq R\}, \quad n \geq 1.$$

We also set  $\sigma_n = \sigma'_n + 1$ . In order to estimate the stopping time  $\tau = \tau(m)$  (which is the first instant when  $|u_t| \vee |u'_t| \leq 1/m$ ), we start with estimating the first integer  $n = n(m)$  for which  $|\mathbf{u}_{\sigma_n}| \leq 1/m$ . The probability of the event  $\{n(m) > k\}$  is equal to

$$P_{\mathbf{u}}(k) = \mathbb{P}_{\mathbf{u}} \left( \bigcap_{n=1}^k \{|\mathbf{u}_{\sigma_n}| > 1/m\} \right).$$

By the construction of  $\sigma'_n$  and the strong Markov property, we have

$$P_{\mathbf{u}}(k) \leq (1 - p)^k \quad \text{for any } \mathbf{u} \in \mathbf{H}, k \geq 1, \quad (3.97)$$

where  $p = p(m) > 0$  is the constant constructed in Lemma 3.3.11 for the ball  $B = B_m$ .

*Step 2.* We now show that

$$\mathbb{E}_{\mathbf{u}} \exp(\gamma \sigma_k) \leq C_1^k (1 + |\mathbf{u}|^2), \quad k \geq 1, \quad (3.98)$$

where  $C_1 > 0$  depends only on  $R$ . Indeed, define  $T'_R = \min\{t \geq 1 : \|\mathbf{u}_t\| \leq R\}$ . Then, by the Markov property and inequalities (3.95) and (2.120), for any  $\mathbf{u}$  such that  $|\mathbf{u}| \leq R$  we have

$$\mathbb{E}_{\mathbf{u}} e^{\gamma T'_R} = \mathbb{E}_{\mathbf{u}} \mathbb{E}_{\mathbf{u}} (e^{\gamma T'_R} | \mathcal{F}_1) = e^{\gamma} \mathbb{E}_{\mathbf{u}} (\mathbb{E}_{\mathbf{u}_1} e^{\gamma T'_R}) \leq e^{\gamma} \mathbb{E}_{\mathbf{u}} (1 + K |\mathbf{u}_1|^2) \leq C_2,$$

where  $\mathcal{F}_t$  is the filtration corresponding to the process  $\mathbf{u}_t$  and  $C_2 > 0$  is a constant depending only on  $R$ . Applying now the strong Markov property, for  $k \geq 1$  we obtain<sup>6</sup>

$$\mathbb{E}_{\mathbf{u}} e^{\gamma \sigma'_k} = \mathbb{E}_{\mathbf{u}} \mathbb{E}_{\mathbf{u}} (e^{\gamma \sigma'_k} | \mathcal{F}_1) = \mathbb{E}_{\mathbf{u}} (e^{\gamma \sigma'_{k-1}} \mathbb{E}_{\mathbf{u}(\sigma'_{k-1})} e^{\gamma T'_R}) \leq C_2 \mathbb{E}_{\mathbf{u}} e^{\gamma \sigma'_{k-1}},$$

where we used again the fact that  $\|\mathbf{u}_{\sigma'_n}\| \leq R$  for any  $n \geq 0$ . Iteration of this inequality results in

$$\mathbb{E}_{\mathbf{u}} e^{\gamma \sigma'_k} \leq C_2^k \mathbb{E}_{\mathbf{u}} e^{\gamma \sigma'_0}.$$

Combining this with (3.95), we arrive at (3.98).

*Step 3.* We can now prove (3.93). For any initial point  $\mathbf{u} \in \mathbf{H}$ , any constant  $M > 0$ , and any integer  $k \geq 1$ , we have

$$\begin{aligned} \mathbb{P}_{\mathbf{u}} \{\tau \geq M\} &= \mathbb{P}_{\mathbf{u}} \{\tau \geq M, \sigma_k < M\} + \mathbb{P}_{\mathbf{u}} \{\tau \geq M, \sigma_k \geq M\} \\ &\leq \mathbb{P}_{\mathbf{u}} \{\tau > \sigma_k\} + \mathbb{P}_{\mathbf{u}} \{\sigma_k \geq M\}. \end{aligned} \quad (3.99)$$

It follows from (3.97), (3.98), and Chebyshev's inequality that

$$\begin{aligned} \mathbb{P}_{\mathbf{u}} \{\tau > \sigma_k\} &\leq P_{\mathbf{u}}(k) \leq (1-p)^k, \\ \mathbb{P}_{\mathbf{u}} \{\sigma_k \geq M\} &\leq e^{-\gamma M} \mathbb{E}_{\mathbf{u}} \exp(\gamma \sigma_k) \leq C_1^k e^{-\gamma M} (1 + |\mathbf{u}|^2). \end{aligned}$$

Substitution of these estimates into (3.99) results in

$$\mathbb{P}_{\mathbf{u}} \{\tau \geq M\} \leq (1-p)^k + C_1^k e^{-\gamma M} (1 + |\mathbf{u}|^2).$$

Choosing  $k \sim \varepsilon M$ , where  $\varepsilon > 0$  is sufficiently small, we obtain

$$\mathbb{P}_{\mathbf{u}} \{\tau \geq M\} \leq C_3 e^{-\gamma' M} (1 + |\mathbf{u}|^2),$$

where  $C_3$  and  $\gamma' < \gamma$  are some positive constants. This inequality immediately implies (3.93), where  $\alpha < \gamma'$ . To complete the proof of Proposition 3.3.6, it remains to establish the two lemmas above.

*Proof of Lemma 3.3.10.* If all the coefficients  $b_j$  are zero, then we get a deterministic Navier–Stokes system, and the lemma is a straightforward consequence

<sup>6</sup>We shall sometimes write  $u(\sigma_n)$  instead of  $u_{\sigma_n}$  to avoid double subscript.

of (2.44) (with  $t = 1$  and  $m = 1$ ) and (2.52). Therefore, we shall assume that there are non-zero coefficients.

Recall that relation (2.131) was established in the proof of Proposition 2.4.10 as a result of application of Itô's formula to the functional  $F(u) = |u|_2^2$ . A similar argument applied to  $F(t, \mathbf{u}) = e^{\alpha_1 \nu t} |\mathbf{u}|^2$  implies that

$$\begin{aligned} e^{\alpha_1 \nu t} |\mathbf{u}_t|^2 + 2\nu \int_0^t e^{\alpha_1 \nu s} |\nabla \mathbf{u}_s|^2 ds &= |\mathbf{u}_0|^2 \\ &+ \int_0^t e^{\alpha_1 \nu s} (\alpha_1 \nu |\mathbf{u}_s|^2 + 2\mathfrak{B} + 2\langle u_s + u'_s, h \rangle) ds + M_t, \end{aligned}$$

where  $M_t$  is the corresponding stochastic integral. Let us fix a parameter  $N > 0$ , take  $t = T_R \wedge N =: T$ , and apply  $\mathbb{E}_{\mathbf{u}}$ . Using Doob's optional sampling theorem and the inequalities

$$|\mathbf{u}_s|^2 \leq \alpha_1^{-1} |\nabla \mathbf{u}_s|^2, \quad |\langle u_s + u'_s, h \rangle| \leq \frac{\alpha_1 \nu}{4} |\mathbf{u}_t|^2 + \frac{2}{\alpha_1 \nu} |h|_2^2,$$

we derive

$$\mathbb{E}_{\mathbf{u}}(e^{\alpha_1 \nu T} |\mathbf{u}_T|^2) + \mathbb{E}_{\mathbf{u}} \int_0^T e^{\alpha_1 \nu s} \left( \frac{\nu}{2} |\nabla \mathbf{u}_s|^2 - (2\mathfrak{B} + \frac{4}{\alpha_1 \nu} |h|_2^2) \right) ds \leq |\mathbf{u}|^2. \quad (3.100)$$

Now note that  $\|\mathbf{u}_s\|^2 \geq R^2$  for  $0 \leq t \leq T_R$ . Hence, if

$$R^2 \geq 8\nu^{-1} (\mathfrak{B} + \frac{2}{\alpha_1 \nu} |h|_2^2),$$

then (3.100) implies that

$$\mathbb{E}_{\mathbf{u}} \left( \frac{\nu R^2}{4} \int_0^T e^{\alpha_1 \nu s} ds \right) \leq |\mathbf{u}|^2.$$

Therefore,

$$\mathbb{E}_{\mathbf{u}} (\exp\{\alpha_1 \nu (T_R \wedge N)\} - 1) \leq \frac{4\alpha_1}{R^2} |\mathbf{u}|^2.$$

Passing to the limit as  $N \rightarrow \infty$  and using Fatou's lemma, we obtain the required inequality (3.95) with  $\gamma = \alpha_1 \nu$  and  $K = 4\alpha_1 R^{-2}$ .  $\square$

*Proof of Lemma 3.3.11.* In view of (3.92), it suffices to show that

$$\inf_{v \in B_V(R)} P_1(v, B) > 0.$$

In other words, we must prove that

$$\inf_{v \in B_V(R)} \mathbb{P}\{|u(1; v) - \hat{u}|_2 < \varepsilon\} > 0 \quad \text{for any } \hat{u} \in H, \varepsilon > 0. \quad (3.101)$$

Before giving a rigorous proof of (3.101), we describe the main idea. We wish to show that the probability of transition from any state  $v \in B_V(R)$  to the

$\varepsilon$ -neighbourhood of a given point  $\hat{u} \in H$  is separated from zero. Since our noise is non-degenerate in all Fourier modes, it is straightforward to find a path  $\hat{\zeta}$  that belongs to the support of the noise and stirs the solution from  $v$  to an arbitrarily small neighbourhood of  $\hat{u}$ . By continuity of the resolving operator, we can find  $\delta > 0$  such that for any path  $\zeta$  from the  $\delta$ -tube  $\mathcal{O}_\delta$  around  $\hat{\zeta}$  the corresponding trajectory ends up in the  $\varepsilon$ -neighbourhood of  $\hat{u}$ . Since  $\hat{\zeta}$  is in the support of the noise, the probability of the event  $\{\zeta \in \mathcal{O}_\delta\}$  is positive, and therefore so is the probability of  $\{|u(1; v) - \hat{u}|_2 < \varepsilon\}$ . The fact that these probabilities are separated from zero uniformly in  $v \in B_V(R)$  follows from some compactness and continuity arguments.

The accurate realisation of this scheme is rather simple when the external force is  $H^1$ -regular. To simplify the presentation, we give here the proof in this particular case; the general situation is considered in Section 3.6.1.

Let us assume that  $h \in V$  and  $\mathfrak{B}_1 < \infty$ . In view of Proposition 2.4.7, solutions of the stochastic Navier–Stokes systems are locally Lipschitz continuous with respect to the trajectories of the noise. In particular, a solution at time 1 can be written as  $u(1) = U_1(v, \zeta)$ , where  $U_1 : H \times C(0, T; V) \rightarrow H$  is a locally Lipschitz continuous mapping. We claim that, for any  $\varepsilon > 0$  and  $v \in B_V(R)$ , we can find  $\hat{\zeta}_v \in C(0, T; V)$  such that

$$\sup_{v \in B_V(R)} |U_1(v, \hat{\zeta}_v) - \hat{u}|_2 \leq \varepsilon/2, \quad (3.102)$$

and the mapping  $v \mapsto \hat{\zeta}_v$  is continuous from  $H$  to  $C(0, T; V)$ . Indeed, let  $\chi(t)$  be a smooth function equal to 1 for  $t \leq 0$  and to 0 for  $t \geq 1$ . For  $\delta > 0$  and  $0 \leq t \leq 1$ , we set

$$\begin{aligned} u_\delta(t) &= \chi(t)e^{(\nu t + \delta)L}v + (1 - \chi(t))e^{\delta L}\hat{u}, \\ \hat{\zeta}_v(t) &= u_\delta(t) - u_\delta(0) + \int_0^t (\nu Lu_\delta + B(u_\delta))ds - th. \end{aligned}$$

Then  $u_\delta(0) = e^{\delta L}v$  and  $U_1(e^{\delta L}v, \zeta_\delta) = e^{\delta L}\hat{u}$ . Choosing  $\delta > 0$  sufficiently small and using the uniform continuity of  $U_1$ , as well as the compactness of  $B_V(R)$  in  $H$ , we conclude that (3.102) holds. The continuous dependence of  $\hat{\zeta}_v$  on  $v$  is a straightforward consequence of the above relations.

We now choose  $\delta > 0$  so small that for any  $\xi \in C(0, T; V)$  satisfying the inequality  $\|\xi - \hat{\zeta}_v\|_{C(0, T; V)} < \delta$  we have (cf. (3.102))

$$\sup_{v \in B_V(R)} |U_1(v, \xi) - \hat{u}|_2 < \varepsilon,$$

It follows that

$$\{|u(1; v) - \hat{u}|_2 < \varepsilon\} \supset \{\|\zeta - \hat{\zeta}_v\|_{C(0, T; V)} < \delta\} =: \Gamma(v) \quad \text{for } v \in B_V(R),$$

whence we conclude that

$$\inf_{v \in B_V(R)} \mathbb{P}\{|u(1; v) - \hat{u}|_2 < \varepsilon\} \geq \inf_{v \in B_V(R)} \mathbb{P}(\Gamma(v)). \quad (3.103)$$

Since the support of the law for the restriction of  $\zeta$  to  $[0, T]$  coincides with  $C(0, T; V)$ , we have  $\mathbb{P}(\Gamma(v)) > 0$  for any  $v \in B_V(R)$ . Furthermore, in view of the portmanteau theorem, the function  $v \mapsto \mathbb{P}(\Gamma(v))$  is lower semicontinuous. It remains to note that the image of  $B_V(R)$  by the mapping  $v \mapsto \hat{\zeta}_v$  is a compact subset of  $C(0, T; V)$ , and therefore the right-hand side of (3.103) must be positive. This completes the proof of the lemma in the case of an  $H^1$ -regular external force.  $\square$

*Exercise 3.3.12.* Prove that, for any  $R, \nu > 0$  and  $h \in H$ , there is a function  $\hat{u} \in H$  and an integer  $N \geq 1$  such that, if  $b_j \neq 0$  for  $j = 1, \dots, N$ , then

$$P_T(\mathbf{u}, B \times B) \geq p \quad \text{for any } u, u' \in B_V(R), \quad (3.104)$$

where  $B \subset H$  is any ball centred at  $\hat{u}$ , and  $T > 0$  and  $p \in (0, 1]$  are some constants depending on  $B$ . *Hint:* If  $h = 0$ , then the claim is true with  $\hat{u} = 0$  for any coefficients  $b_j \geq 0$ . Similarly, if  $|h|_2 \ll 1$ , then the unperturbed problem has a unique stationary point  $\hat{u} \in V$ , which is globally exponentially stable as  $t \rightarrow \infty$ , and the claim is valid again for any  $b_j \geq 0$ . In the case of an arbitrary  $h \in H$ , one can use finitely many non-zero components of the noise to annihilate the low modes of  $h$ , while the contribution of high modes will be small.

### 3.3.3 Stability: proof of Proposition 3.3.7

We first outline the scheme of the proof, which is based on the Foias–Prodi estimate and Girsanov’s theorem. Without loss of generality, we can assume that  $u = 0$ , because the general case can be easily derived from this one with the help of the triangle inequality. Thus, we need to show that

$$\|\mathcal{D}(u_t) - \mathcal{D}(u'_t)\|_L^* \leq \delta_m \quad \text{for all } t \geq 0, \quad (3.105)$$

where  $u_t$  and  $u'_t$  stand for trajectories of the Navier–Stokes system (3.84) that are issued from  $u = 0$  and  $u' \in B_m = \{u \in H : |u|_2 \leq m^{-1}\}$ , respectively, and  $\delta_m \rightarrow 0$  as  $m \rightarrow \infty$ . Let us define an auxiliary process  $v_t$  as a solution of the problem

$$\dot{v} + \nu Lv + B(v) + \lambda P_N(v - u'_t) = h + \eta(t), \quad (3.106)$$

$$v_0 = 0, \quad (3.107)$$

where  $P_N : H \rightarrow H$  denotes the orthogonal projection on the subspace  $H_{(N)}$  spanned by  $e_j$ ,  $j = 1, \dots, N$ , and  $\lambda > 0$  and  $N \geq 1$  are large parameters that will be chosen later. Using the Foias–Prodi estimate (see Theorem 2.1.28), we shall show that, with high probability,

$$|v_t - u'_t|_2 \leq C e^{-t} |u'_t|_2 \quad \text{for } t \geq 0, \quad (3.108)$$

where  $C > 0$  is a deterministic constant. This inequality will imply that

$$\|\mathcal{D}(v_t) - \mathcal{D}(u'_t)\|_L^* \leq \delta_m^{(1)} \quad \text{for all } t \geq 0, \quad (3.109)$$

where  $\delta_m^{(1)} \rightarrow 0$  as  $m \rightarrow \infty$ . Thus, it suffices to compare the laws of  $u_t$  and  $v_t$ .

We shall prove that if  $|u'|_2 \leq \frac{1}{m}$ , then

$$\|\mathcal{D}(u_t) - \mathcal{D}(v_t)\|_L^* \leq \delta_m^{(2)} \quad \text{for all } t \geq 0, \quad (3.110)$$

where  $\delta_m^{(2)} \rightarrow 0$  as  $m \rightarrow \infty$ . Note that since inequality (3.110) concerns the law of solutions (and not the solutions themselves), we can choose the underlying probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  at our convenience. We assume that it coincides with the canonical space of the Wiener process  $\{\zeta(t), t \geq 0\}$ ; see (2.71). Namely,  $\Omega$  is the space of continuous functions  $\omega : \mathbb{R}_+ \rightarrow H$  endowed with the metric of uniform convergence on bounded intervals,  $\mathbb{P}$  is the law of  $\zeta$ , and  $\mathcal{F}$  is the completion of the Borel  $\sigma$ -algebra with respect to  $\mathbb{P}$  (cf. Example 1.3.13 and Section 3.4 in [Str93]). In this case,  $\zeta$  is the *canonical process* given by  $\zeta(t) = \omega_t$  for  $t \geq 0$ . We now define a transformation  $\Phi : \Omega \rightarrow \Omega$  by the relation

$$\Phi(\omega)_t = \omega_t - \lambda \int_0^t \mathbf{P}_N(v_s - u'_s) ds, \quad (3.111)$$

where  $v_t$  and  $u'_t$  are solutions of the corresponding equations with the right-hand side  $\eta(t) = \eta^\omega(t) = \partial_t \omega_t$ . We shall show that

$$\mathbb{P}\{u_t(\Phi(\omega)) = v_t(\omega) \text{ for all } t \geq 0\} = 1. \quad (3.112)$$

In view of Exercise 1.2.13 (ii) and inequality (1.16), to prove (3.110) it suffices to estimate the total variation distance between  $\mathbb{P}$  and  $\Phi_*(\mathbb{P})$ . This will be done with the help of Girsanov's theorem. We now turn to the accurate proof, which is divided into several steps.

*Step 1.* Let us prove (3.109). To this end, we fix any integer  $\varepsilon > 0$  and use Proposition 2.4.10 to find positive constants  $K$  and  $\rho_\varepsilon$  such that

$$\mathbb{P}\left\{\int_0^t \|u'_s\|_1^2 ds \leq \rho_\varepsilon + Kt \text{ for all } t \geq 0\right\} \geq 1 - \varepsilon \quad \text{for any } u' \in B_1.$$

Let  $\Omega_{K, \rho_\varepsilon}$  be the event in the left-hand side of this inequality. By Theorem 2.1.28 with  $M = 1$  and Remark 2.1.29, there is a constant  $\lambda > 0$  and an integer  $N \geq 1$  such that if  $v$  is a solution of (3.106), (3.107), then inequality (3.108) with  $C = C_\varepsilon = e^{c\rho_\varepsilon}$  holds for  $\omega \in \Omega_{K, \rho_\varepsilon}$ . In this case, for any  $t \geq 0$ ,  $u' \in B_1$ , and  $f \in L_b(H)$  with  $\|f\|_L \leq 1$  we have

$$|\mathbb{E}(f(v_t) - f(u'_t))| \leq 2\mathbb{P}(\Omega_{K, \rho_\varepsilon}^c) + |\mathbb{E} \mathbb{I}_{\Omega_{K, \rho_\varepsilon}}(f(v_t) - f(u'_t))| \leq 2\varepsilon + C_\varepsilon |u'|_2$$

Since  $f$  was arbitrary, we conclude that (3.109) holds.

*Step 2.* We now prove (3.112). Let us set  $\tilde{v}_t(\omega) = u_t(\Phi(\omega))$ . Since

$$\frac{\partial}{\partial t} \Phi(\omega) = \frac{\partial \zeta}{\partial t} - \lambda \mathbf{P}_N(v_t - u'_t),$$



the process  $\tilde{v}$  is a solution of problem (3.106), (3.107). By the uniqueness of a solution (see Theorem 2.4.6), we conclude that

$$\mathbb{P}\{\tilde{v}_t(\omega) = v_t(\omega) \text{ for all } t \geq 0\} = 1.$$

This relation coincides with (3.112).

*Step 3.* We have thus shown that, for any  $t \geq 0$ , the random variables  $u_t$  and  $v_t$  satisfy the hypotheses of Exercise 1.2.13 (ii). Therefore, to establish (3.110), it suffices to prove that the transformation  $\Phi : \Omega \rightarrow \Omega$  defined by (3.111) satisfies the inequality

$$\|\mathbb{P} - \Phi_*(\mathbb{P})\|_{\text{var}} \leq \delta_m^{(2)} \rightarrow 0 \quad \text{as } m \rightarrow \infty. \quad (3.113)$$

The proof of this fact is based on Girsanov's theorem. The space  $\Omega = C(\mathbb{R}_+, H)$  can be written as the direct sum of the closed subspaces  $\Omega_N = C(\mathbb{R}_+, H_{(N)})$  and  $\Omega_N^\perp = C(\mathbb{R}_+, H_{(N)}^\perp)$ . We shall accordingly write  $\Omega \ni \omega = (\omega^{(1)}, \omega^{(2)})$ . In this notation, the transformation  $\Phi$  takes the form

$$\Phi(\omega^{(1)}, \omega^{(2)}) = (\Psi(\omega^{(1)}, \omega^{(2)}), \omega^{(2)}),$$

where  $\Psi : \Omega \rightarrow \Omega_N$  is defined by the relation

$$\Psi(\omega^{(1)}, \omega^{(2)})_t = \omega_t^{(1)} + \int_0^t a(s; \omega^{(1)}, \omega^{(2)}) ds, \quad a(t) = -\lambda P_N(v_t - u_t'). \quad (3.114)$$

We now need the following lemma established at the end of this subsection.

**Lemma 3.3.13.** *Let  $\mathbb{P}_N$  and  $\mathbb{P}_N^\perp$  be the images of  $\mathbb{P}$  under the natural projections*

$$P_N : \Omega \rightarrow \Omega_N, \quad Q_N : \Omega \rightarrow \Omega_N^\perp.$$

*Then*

$$\|\Phi_*(\mathbb{P}) - \mathbb{P}\|_{\text{var}} \leq \int_{\Omega_N^\perp} \|\Psi_*(\mathbb{P}_N, \omega^{(2)}) - \mathbb{P}_N\|_{\text{var}} \mathbb{P}_N^\perp(d\omega^{(2)}), \quad (3.115)$$

*where  $\Psi_*(\mathbb{P}_N, \omega^{(2)})$  stands for the image of  $\mathbb{P}_N$  under the mapping  $\Psi(\cdot, \omega^{(2)})$ .*

Thus, to prove the required result, it suffices to derive an appropriate estimate for the right-hand side of (3.115). To do this, we may try to apply Theorem 7.10.1 to the processes  $y = \omega^{(1)}$  and  $\tilde{y} = \Psi(\omega^{(1)}, \omega^{(2)})$ . However, we cannot do it directly because the process  $a(t)$  does not necessarily satisfy Novikov's condition (7.51). To overcome this difficulty, we fix any  $\varepsilon > 0$  and find a constant  $C = C_\varepsilon > 0$  such that (3.108) holds with probability no less than  $1 - \varepsilon$  (cf. Step 1). We now introduce an auxiliary process defined as

$$\tilde{a}(t) = \chi\left((C|u'|_2)^{-1} \sup_{0 \leq s \leq t} (e^s |v_s - u_s'|_2)\right) a(t),$$

where  $\chi$  is the indicator function of the interval  $[0, 1]$ . In other words, the process  $\tilde{a}$  coincides with  $a$  if  $v - u'$  satisfies (3.108) and is zero starting from the first instant  $t \geq 0$  when (3.108) fails. This construction implies that

$$\mathbb{P}\{\tilde{a}(t) = a(t) \text{ for all } t \geq 0\} \geq 1 - \varepsilon. \quad (3.116)$$

We now define  $\tilde{\Phi}(\omega^{(1)}, \omega^{(2)}) = (\tilde{\Psi}(\omega^{(1)}, \omega^{(2)}), \omega^{(2)})$ , where  $\tilde{\Psi} : \Omega \rightarrow \Omega_N$  is given by relation (3.114) with  $a$  replaced by  $\tilde{a}$ . Inequality (3.116) implies that  $\Phi$  and  $\tilde{\Phi}$  coincide with probability  $\geq 1 - \varepsilon$ , whence it follows that

$$\|\Phi_*(\mathbb{P}) - \tilde{\Phi}_*(\mathbb{P})\|_{\text{var}} \leq \varepsilon. \quad (3.117)$$

We now estimate the total variation distance between  $\tilde{\Phi}_*(\mathbb{P})$  and  $\mathbb{P}$ . In view of (3.115), it suffices to compare  $\tilde{\Psi}_*(\mathbb{P}, \omega^{(2)})$  and  $\mathbb{P}_N$ . To this end, we apply Theorem 7.10.1. Define the processes

$$y(t) = \omega_t^{(1)}, \quad \tilde{y}(t) = \omega_t^{(1)} + \int_0^t \tilde{a}(s; \omega^{(1)}, \omega^{(2)}) ds$$

and note that  $\mathbb{P}_N = \mathcal{D}(y)$  and  $\tilde{\Psi}_*(\mathbb{P}, \omega^{(2)}) = \mathcal{D}(\tilde{y})$ . Furthermore, the definition of  $\tilde{a}$  implies that

$$|\tilde{a}(t)| \leq C e^{-t} |u'|_2 \quad \text{for all } t \geq 0 \text{ and a.e. } \omega \in \Omega, \quad (3.118)$$

and therefore Novikov's condition (7.51) holds for  $\tilde{a}$ . Hence, recalling that  $b_j > 0$  for all  $j \geq 1$  and using (7.52), we derive

$$\|\tilde{\Psi}_*(\mathbb{P}, \omega^{(2)}) - \mathbb{P}_N\|_{\text{var}} \leq \frac{1}{2} (\sqrt{Q_N} - 1)^{1/2}, \quad (3.119)$$

where we set

$$Q_N = \mathbb{E} \exp\left(6 \bar{b}_N^{-2} \int_0^\infty |\tilde{a}(t)|^2 dt\right), \quad \bar{b}_N = \min_{1 \leq j \leq N} b_j.$$

It follows from (3.118) that  $Q_N \leq \exp(C_{N,\varepsilon} |u'|_2^2)$ , where  $C_{N,\varepsilon} > 0$  does not depend on  $u'$ . Combining this with (3.117) and (3.119), we conclude that

$$\|\Phi_*(\mathbb{P}) - \mathbb{P}\|_{\text{var}} \leq \varepsilon + C'_{N,\varepsilon} |u'|_2 \quad \text{for any } u' \in B_1.$$

Since  $\varepsilon > 0$  was arbitrary, we arrive at (3.113). To complete the proof of Proposition 3.3.7, it remains to establish Lemma 3.3.13.

*Step 4: Proof of Lemma 3.3.13.* Let  $f : \Omega \rightarrow \mathbb{R}$  be a bounded continuous function with  $\|f\|_\infty \leq 1$ . Then

$$\begin{aligned} |\mathbb{E}(f(\Phi(\omega)) - f(\omega))| &= \left| \int_\Omega (f(\Psi(\omega), \omega^{(2)}) - f(\omega^{(1)}, \omega^{(2)})) \mathbb{P}(d\omega) \right| \\ &\leq \int_{\Omega_N^\perp} \left| \int_{\Omega_N} (f(\Psi(\omega^{(1)}, \omega^{(2)}), \omega^{(2)}) - f(\omega^{(1)}, \omega^{(2)})) \mathbb{P}_N(d\omega^{(1)}) \right| \mathbb{P}_N^\perp(d\omega^{(2)}) \\ &\leq \int_{\Omega_N^\perp} \|\Psi_*(\mathbb{P}_N, \omega^{(2)}) - \mathbb{P}_N\|_{\text{var}} \mathbb{P}_N^\perp(d\omega^{(2)}). \end{aligned}$$

Since  $f$  was arbitrary, we arrive at (3.115). The proof of Proposition 3.3.7 is complete.

### 3.4 Navier–Stokes system with unbounded kicks

In Section 3.2.4, we proved that the homogeneous Navier–Stokes system perturbed by a sufficiently non-degenerate bounded kick force has a unique stationary distribution, and the laws of all other solutions converge to it exponentially fast. The proof of this result was based on the exponential decay of a Kantorovich functional (see Theorem 3.1.1). In this section we establish these results for the case of unbounded kicks. To do this, we use Theorem 3.1.7, whose proof combines the coupling with a stopping times technique.

#### 3.4.1 Formulation of the result

Let us consider the Navier–Stokes system (3.84) in a bounded domain  $Q \subset \mathbb{R}^2$ . We assume that  $h \in H$  is a deterministic function and  $\eta$  is the random process given by (2.65), (3.75):

$$\eta(t, x) = \sum_{k=1}^{\infty} \eta_k(x) \delta(t - kT), \quad \eta_k(x) = \sum_{j=1}^{\infty} b_j \xi_{jk} e_j(x),$$

where the kicks  $\{\eta_k\}$  form a sequence of i.i.d. random variables in  $H$ . As was explained in Section 2.3, the problem in question is equivalent to the discrete-time RDS (3.50), in which  $S = S_T : H \rightarrow H$  stands for the operator that takes a function  $u_0 \in H$  to the value at time  $T$  of the solution of the deterministic Navier–Stokes system (2.77) supplemented with the initial condition  $u(0) = u_0$ .

We now formulate the main conditions imposed on the kicks. We assume that  $\{e_j\}$  is an orthonormal basis in  $H$  formed of the eigenfunctions of  $L$ ,  $b_j \geq 0$  are some constants satisfying (3.76), and  $\xi_{jk}$  are independent scalar random variables satisfying the inequality

$$\mathbb{E} \exp(\varkappa |\xi_{jk}|^2) \leq C \tag{3.120}$$

with some positive constants  $C$  and  $\varkappa$  not depending on  $j$  and  $k$ . It is straightforward to see that if this inequality holds, then  $\mathbb{E} \exp(\varkappa_0 |\eta_1|_2^2) < \infty$  for some  $\varkappa_0 \in (0, \varkappa]$ . In particular, the conclusions of Propositions 2.3.4 and 2.3.8 hold.

Recall that we denote by  $\{\varphi_k, k \geq 0\}$  the Markov RDS associated with (3.50); that is,  $\varphi_k u_0 = u_k$  for  $k \geq 0$ , where  $u_0, u_1, u_2, \dots$  are defined by (3.50). Let  $\mathfrak{P}_k$  and  $\mathfrak{P}_k^*$  be the corresponding Markov operators. The following theorem is the main result of this section (cf. Theorem 3.2.9).

**Theorem 3.4.1.** *Under the above conditions, assume that the laws of the random variables  $\xi_{jk}$  are absolutely continuous with respect to the Lebesgue measure, and the corresponding densities  $p_j$  are functions of bounded total variation that are positive almost everywhere. In this case, there is an integer  $N = N(\mathfrak{B}, \nu) \geq 1$  such that if*

$$b_j \neq 0 \quad \text{for } j = 1, \dots, N, \tag{3.121}$$

then the RDS  $\{\varphi_k, k \geq 0\}$  has a unique stationary measure  $\mu \in \mathcal{P}(H)$ . Moreover, there are positive constants  $C$  and  $\alpha$  such that

$$\|\mathfrak{P}_k^* \lambda - \mu\|_L^* \leq C e^{-\alpha k} \left(1 + \int_H |u|_2 \lambda(du)\right), \quad k \geq 0, \quad (3.122)$$

where  $\lambda \in \mathcal{P}(H)$  is an arbitrary measure with finite first moment.

The scheme of the proof of this result is described in the next subsection, and the details are given in Subsections 3.6.2 and 3.6.3. Here we make a comment on the conditions imposed on the external force of (3.84) and establish a corollary of Theorem 3.4.1.

In the case of bounded kicks, we assumed that the function  $h$  is zero, and the support of the law for the kicks contains zero. This hypothesis ensured that, with positive probability, any two solutions approach each other arbitrarily closely. In the present section, the function  $h \in H$  is arbitrary, however, we do assume that the noise is very efficient in the sense that the support of its law contains a *subspace* of high dimension (rather than a ball in a subspace of high dimension, as was the case in Theorem 3.2.9).

As in the case of white noise, the mixing and a priori estimates imply the convergence of the mean value of a functional  $f : H \rightarrow \mathbb{R}$ , calculated on solutions, to the mean value of  $f$  with respect to the stationary distribution. Namely, for any weight function  $w \in \mathcal{W}$  and a constant  $\gamma \in (0, 1]$ , we denote by  $C^\gamma(H, w)$  the space of Hölder-continuous functions  $f : H \rightarrow \mathbb{R}$  with finite norm

$$|f|_{w, \gamma} = \sup_{u \in H} \frac{|f(u)|}{w(|u|_2)} + \sup_{0 < |u-v|_2 \leq 1} \frac{|f(u) - f(v)|}{|u-v|_2^\gamma (w(|u|_2) + w(|v|_2))}.$$

Note that in the case  $\gamma = 1$  and  $w \equiv 1$  we obtain the space  $L_b(H)$ .

**Corollary 3.4.2.** *Under the conditions of Theorem 3.4.1, for any  $\gamma \in (0, 1]$ ,  $\nu > 0$ , and  $\mathfrak{B} > 0$  there are positive constants  $C$ ,  $\beta$ , and  $\varkappa$  such that, for any function  $f \in C^\gamma(H, w)$  with  $w(r) = \exp(\varkappa r^2)$ , we have*

$$\left| \mathbb{E} f(\varphi_k u) - \int_H f(v) \mu(dv) \right| \leq C |f|_{w, \gamma} w(|u|_2) e^{-\beta k} \quad \text{for } k \geq 0, u \in H. \quad (3.123)$$

*Proof.* We shall apply the scheme used in the proof of Corollary 3.3.2. The only difference is that now the estimates are more quantitative.

We first note that both terms on the left-hand side of (3.123) are well defined. Indeed, the fact that  $\mathbb{E} f(\varphi_k u)$  is finite was proved in Proposition 2.3.4. Furthermore, repeating the argument of the proof of Theorem 2.5.3 and using inequality (2.87) instead of (2.81), for  $\varkappa \ll 1$  one can show that

$$\int_H \exp(2\varkappa |v|_2^2) \mu(dv) < \infty. \quad (3.124)$$

This implies, in particular, that  $(f, \mu)$  is finite for any  $f \in C^\gamma(H, w)$ .

To prove (3.123), we fix any function  $f \in C^\gamma(H, w)$  and assume, without loss of generality, that  $|f|_{w, \gamma} = 1$ . Using representation (3.90) with  $t = k$ , we estimate the left-hand side  $\Delta_k$  of (3.123) as follows:

$$\Delta_k \leq |\mathfrak{P}_k f_{\leq R}(u) - (f_{\leq R}, \mu)| + |\mathbb{E} f_{\geq R}(\varphi_k u)| + |(f_{\geq R}, \mu)|. \quad (3.125)$$

Let us estimate each term of the right-hand side. Inequality (3.122) and Lemma 1.2.6 imply that

$$|\mathfrak{P}_k g(u) - (g, \mu)| \leq C_1 |g|_\gamma e^{-\gamma \alpha k} (1 + |u|_2), \quad k \geq 0, \quad (3.126)$$

where  $g \in C_b^\gamma(H)$  is an arbitrary function. Taking  $g = f_{\leq R}$ , we get

$$|\mathfrak{P}_k f_{\leq R}(u) - (f_{\leq R}, \mu)| \leq C_2 e^{-\gamma \alpha k} w(R+1) (1 + |u|_2).$$

Furthermore, it follows from inequality (2.87) with  $\varkappa$  replaced by  $2\varkappa$  that

$$\begin{aligned} |\mathbb{E} f_{\geq R}(\varphi_k u)| &= |\mathbb{E}(\mathbb{I}_{\{|\varphi_k u|_2 \geq R\}} f_{\geq R}(\varphi_k u))| \\ &\leq (\mathbb{P}\{|\varphi_k u|_2 \geq R\})^{1/2} (\mathbb{E} f_{\geq R}^2(\varphi_k u))^{1/2} \\ &\leq (\mathbb{P}\{w(\varphi_k u) \geq w(R)\})^{1/2} (\mathbb{E} \exp(2\varkappa |\varphi_k u|_2^2))^{1/2} \\ &\leq \frac{C_3 w(|u|_2)}{\sqrt{w(R)}}. \end{aligned}$$

Finally, inequality (3.124) implies that

$$|(f_{\geq R}, \mu)| \leq \frac{C_4}{w(R)}.$$

Substituting the above estimates into (3.125), we obtain

$$\Delta_k \leq C_5 \left( e^{-\gamma \alpha k} w(R+1) (1 + |u|_2) + \frac{1 + w(|u|_2)}{\sqrt{w(R)}} \right).$$

Choosing  $R$  such that  $R^2 = \frac{\gamma \alpha k}{2\varkappa}$ , we arrive at the required inequality (3.123) with  $\beta = \frac{\gamma \alpha}{4}$ .  $\square$

### 3.4.2 Proof of Theorem 3.4.1

Let  $(u_k, \mathbb{P}_u)$  be the Markov family that is associated with the RDS  $\{\varphi_k, k \geq 0\}$  corresponding to (3.50). It follows from Proposition 2.3.3 that

$$\mathbb{E} |u_k|_2 \leq C_1 (1 + |u|_2) \quad \text{for all } u \in H, k \geq 0$$

where  $C_1 > 0$  does not depend on  $u$  and  $k$ . By Theorem 3.1.7, the required result will be established if we construct a mixing extension  $(\mathbf{u}_k, \mathbb{P}_u)$  for the family  $(u_k, \mathbb{P}_u)$  and show that relations (3.31) – (3.34) hold with a function of the form

$$g(r) = C(1 + r). \quad (3.127)$$

It turns out that one can construct a suitable extension in the same way as for the case of bounded kicks (see Section 3.2.2). Namely, we fix an integer  $N \geq 1$  and denote by  $\mathbf{P}_N : H \rightarrow H$  and  $\mathbf{Q}_N : H \rightarrow H$  the orthogonal projections onto the subspaces  $H_{(N)}$  and  $H_{(N)}^\perp$ , respectively. The proof of Lemma 3.2.6 does not use the boundedness of the kick  $\eta_k$  and its assertions remain valid in the context of this section with  $E = H_{(N)}$ ,  $\mathbf{P} = \mathbf{P}_N$ , and  $\mathbf{Q} = \mathbf{Q}_N$ . Let us denote by  $\zeta = \zeta(v, v', \omega)$  and  $\zeta' = \zeta'(v, v', \omega)$  the random variables constructed in Lemma 3.2.6 and define *coupling operators*  $(\mathcal{R}, \mathcal{R}')$  be relations (3.60). Repeating the construction of Section 3.2.2, we obtain an RDS  $\Phi_k = (\Phi_k, \Phi'_k)$  in the product space  $H \times H$  such that the corresponding Markov family  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  is an extension for  $(u_k, \mathbb{P}_u)$ . Let us set

$$\sigma_0 = \min\{k \geq 0 : |u_k - u'_k|_2 \geq c e^{-k}\}, \quad (3.128)$$

where  $c > 0$  is a constant that will be chosen later. In Section 3.6, we shall show that the following two propositions are true.

**Proposition 3.4.3.** *Suppose that condition (3.121) is satisfied with a sufficiently large  $N \geq 1$ . Then for any  $d > 0$  the recurrence property of Definition 3.1.6 holds for the extension  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  with  $\mathbf{B} = \mathbf{B}_d = B_H(d) \times B_H(d)$  and a function  $g$  of the form (3.127).*

**Proposition 3.4.4.** *For sufficiently large integers  $N \geq 1$  there is a stopping time  $\sigma$  and a constant  $d > 0$  such that if condition (3.121) is satisfied, then*

$$\mathbb{P}_{\mathbf{u}}\{\sigma \leq \sigma_0\} = 1 \quad \text{for any } \mathbf{u} \in \mathbf{B}_d, \quad (3.129)$$

and relations (3.32) – (3.34) hold for  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  with  $\mathbf{B} = \mathbf{B}_d$ .

Combining these results with Theorem 3.1.7 and Remark 3.1.8, we easily derive the uniqueness of stationary distribution and inequality (3.122). Indeed, in view of the propositions, we can choose the parameters  $N \geq 1$  and  $d > 0$  so that, if condition (3.121) is fulfilled, then the Markov family  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  possesses the recurrence and exponential squeezing properties of Definition 3.1.6 with  $\mathbf{B} = \mathbf{B}_d$ , a stopping time  $\sigma$  satisfying (3.129), and a function  $g$  of the form (3.127). Therefore, by Remark 3.1.8, the conclusions of Theorem 3.1.7 are true for the Markov family  $(u_k, \mathbb{P}_u)$ . This completes the proof of the theorem.

## 3.5 Further results and generalisations

### 3.5.1 Flandoli–Maslowski theorem

The very first result on the uniqueness of a stationary distribution for the 2D Navier–Stokes system perturbed by a random force was obtained by Flandoli and Maslowski [FM95]. In this subsection we formulate their result and give some references.

Let us consider again the Navier–Stokes system (3.84), in which  $h \in H$  is a deterministic function and  $\eta$  is a random process of the form (2.66). In contrast

to the situation studied in Section 3.3, we now assume that *all* coefficients  $b_j$  are non-zero and satisfy the two-sided inequality

$$c\alpha_j^{-\frac{1}{2}} \leq b_j \leq C\alpha_j^{-\frac{3}{8}-\varepsilon} \quad \text{for any } j \geq 1, \quad (3.130)$$

where  $C$ ,  $c$ , and  $\varepsilon$  are positive constants and  $\{\alpha_j\}$  is the non-decreasing sequence of eigenvalues for the Stokes operator  $L$ . Note that, by the well-known spectral asymptotics for  $L$  (e.g., see Chapter 4 in [CF88]), we have  $\alpha_j \sim j$  as  $j \rightarrow \infty$ , whence it follows that condition (2.67) is not satisfied, so that the random force  $\eta(t, x)$  is very rough as a function of the space variable  $x$ . Nevertheless, one can show that the initial-boundary value problem for (3.84) is well posed. We consider the corresponding Markov process in  $H$ . The main result of [FM95] is the following theorem.

**Theorem 3.5.1.** *Assume that the coefficients  $b_j$  satisfy condition (3.130). Then the following assertions hold.*

**Uniqueness:** *The Markov process possesses a unique stationary distribution  $\mu \in \mathcal{P}(H)$ .*

**Mixing:** *For any  $\lambda \in \mathcal{P}(H)$ , we have*

$$\|\mathfrak{P}_t^* \lambda - \mu\|_{\text{var}} \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

**Ergodicity:** *Let  $f : H \rightarrow \mathbb{R}$  be a Borel-measurable function integrable with respect to  $\mu$  and let  $u(t, x)$  be any solution with a deterministic initial datum  $u(0) \in H$ . Then*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(u(t)) dt = \int_H f d\mu \quad \text{almost surely.}$$

Without going into details, let us outline the main ideas of the proof.<sup>7</sup> It is based on Doob's theorem which provides sufficient conditions for uniqueness and mixing; see [Doo48, DZ96, Ste94, Sei97]. According to that result, if the transition function of the Markov RDS  $\Phi$  is *regular* in the sense that the probability measures  $P_t(u, \cdot)$ ,  $u \in H$ , are absolutely continuous with respect to one another for some  $t > 0$ , then the conclusions of Theorem 3.5.1 are valid. The regularity of the transition function is, in turn, a consequence of the following two properties.

**Irreducibility:** There is  $t_1 > 0$  such that

$$P_{t_1}(u, B_H(v, r)) > 0 \quad \text{for all } u, v \in H, r > 0.$$

**Strong Feller property:** There is  $t_2 > 0$  such that the function  $u \mapsto P_{t_2}(u, \Gamma)$  is continuous on  $H$  for any  $\Gamma \in \mathcal{B}(H)$ .

---

<sup>7</sup>The scheme presented below is not entirely accurate. We refer the reader to the original work [FM95] for the precise argument, which is more complicated.

It is not difficult to show that if these properties holds, then the transition probabilities are mutually absolutely continuous for  $t = t_1 + t_2$ . Thus, it suffices to establish the irreducibility and strong Feller property.

The proof of irreducibility can be carried out by repeating the argument used to establish Lemma 3.3.11. The hard point of the Flandoli–Maslowski analysis is the verification of the strong Feller property. The latter is based on an application of the Bismut–Elworthy formula [Bis81, Elw92] to Galerkin’s approximations of the Navier–Stokes system and derivation of a uniform estimate for the derivative of the Markov semigroup.

The Flandoli–Maslowski result was developed and extended by many authors. In particular, Ferrario [Fer99] proved that the non-degeneracy condition (3.130) can be relaxed to allow random forces that are regular with respect to the space variables, Bricmont, Kupiainen, and Lefevere [BKL01] established the exponential convergence to the unique stationary distribution in the total variation norm for the Navier–Stokes system perturbed by a kick force, and Goldys and Maslowski [GM05] obtained a similar result in the case of white noise perturbations. We refer the reader to those papers for further references on this subject.

### 3.5.2 Exponential mixing for the Navier–Stokes system with white noise

Let us recall that Theorem 3.3.1 establishes the uniqueness of a stationary measure and the mixing property for the Navier–Stokes system (3.84) with a spatially regular white noise  $\eta$ . On the other hand, Theorems 3.2.9 and 3.4.1 show that, in the case of a non-degenerate kick force, the corresponding RDS is exponentially mixing. It turns out that a similar result is true for white noise perturbations. Namely, the following theorem was proved in [BKL02, Mat02b, KS02a, Oda08] at various levels of generality; see *Notes and Comments* at the end of this chapter.

**Theorem 3.5.2.** *In the setting of Section 3.3, assume that  $\mathfrak{B} = \sum_j b_j^2 < \infty$ . Then there is an integer  $N \geq 1$  depending on  $\nu$ ,  $\mathfrak{B}$ , and  $|h|_2$  such that the following properties hold, provided that*

$$b_j \neq 0 \quad \text{for } j = 1, \dots, N. \quad (3.131)$$

**Uniqueness:** *The Eq. (2.98) has a unique stationary measure  $\mu \in \mathcal{P}(H)$ .*

**Exponential mixing:** *There are positive constants  $C$  and  $\alpha$  such that*

$$\|\mathfrak{P}_t^* \lambda - \mu\|_L^* \leq C e^{-\alpha t} \left( 1 + \int_H |u|_2^2 \lambda(du) \right), \quad t \geq 0, \quad (3.132)$$

*where  $\lambda \in \mathcal{P}(H)$  is an arbitrary measure for which the right-hand side is finite.*



Let us also mention that this theorem remains true for some classes of multiplicative noises (see [Oda08]). The proof is based on the ideas used to establish the exponential mixing in the case of a kick force, and therefore we confine ourselves to a brief description of the coupling construction. The interested reader will be able to fill in the details repeating the scheme applied for the kick force model (cf. the paper [Shi08] in which the details are carried out for the complex Ginzburg–Landau equation (0.4)).

*Outline of the proof of Theorem 3.5.2.* Let  $(u_t, \mathbb{P}_v)$  be the Markov process associated with (2.98) and let  $\{u(t; v), t \geq 0\}$  be a solution of (2.98) issued from  $v \in H$  at  $t = 0$ . We wish to construct a mixing extension for  $(u_t, \mathbb{P}_v)$ ; cf. Definition 3.1.6. Let us fix a time  $T > 0$ , an integer  $N \geq 1$ , and an arbitrary function  $\chi \in C^\infty(\mathbb{R})$  such that  $\chi(t) = 0$  for  $t \leq 0$  and  $\chi(t) = 1$  for  $t \geq T$ . Recall that  $\mathbf{P}_N$  and  $\mathbf{Q}_N$  stand for the orthogonal projections in  $H$  onto the subspaces  $H_N$  and  $H_N^\perp$ , respectively, where  $H_N$  is the vector span of the first  $N$  eigenfunctions of the Stokes operator. Given two points  $v, v' \in H$ , we denote by  $\lambda(v), \lambda'(v, v') \in \mathcal{P}(C(0, T; H \times H))$  the laws of the random variables

$$((u(t; v), \zeta(t)), t \in [0, T]), \quad ((u(t; v') - \chi(t)\mathbf{P}_N(v' - v), \zeta(t)), t \in [0, T]).$$

Let us define a mapping  $f : C(0, T; H \times H) \rightarrow C(0, T; H_N \times H_N^\perp)$  by the formula

$$f((u(t), \xi(t)), t \in [0, T]) = ((\mathbf{P}_N u(t), \mathbf{Q}_N \xi(t)), t \in [0, T]).$$

In view of Exercise 1.2.30, there is a coupling  $(\mathbf{Y}(v, v'), \mathbf{Y}'(v, v'))$  for the pair of measures  $(\lambda(v), \lambda'(v, v'))$  such that  $(f(\mathbf{Y}(v, v')), f(\mathbf{Y}'(v, v')))$  is a maximal coupling for  $(f_*(\lambda(v)), f_*(\lambda'(v, v')))$ . Thus,  $\mathbf{Y}$  and  $\mathbf{Y}'$  are  $H \times H$ -valued stochastic processes on the interval  $[0, T]$ , and we denote by  $\Upsilon_t(v, v')$  and  $\Upsilon'_t(v, v')$  the restriction of their first components to time  $t$ . We now define coupling operators on the interval  $[0, T]$  by the relations (cf. (3.60))

$$\mathcal{R}_t(v, v') = \Upsilon_t(v, v'), \quad \mathcal{R}'_t(v, v') = \Upsilon'_t(v, v') + \chi(t)\mathbf{P}_N(v' - v), \quad 0 \leq t \leq T.$$

Iteration of  $\mathcal{R}$  and  $\mathcal{R}'$  enables one to construct a Markov RDS  $\Phi_t = (\Phi_t, \Phi'_t)$  in the product space  $H \times H$ ; cf. Section 3.2.2. Namely, we define  $(\Omega, \mathcal{F}, \mathbb{P})$  as the direct product of countably many copies of the probability space for the coupling operators and denote by  $\omega = (\omega^1, \omega^2, \dots)$  points of  $\Omega$ . Given  $v, v' \in H$ , we now set

$$\left. \begin{aligned} \Phi_t(\omega)(v, v') &= \mathcal{R}_t(v, v', \omega^1) \\ \Phi'_t(\omega)(v, v') &= \mathcal{R}'_t(v, v', \omega^1) \end{aligned} \right\} \quad \text{for } 0 \leq t \leq T,$$

$$\left. \begin{aligned} \Phi_t(\omega)(v, v') &= \mathcal{R}_t(\Phi_{(k-1)T}(\omega)(v, v'), \omega^k) \\ \Phi'_t(\omega)(v, v') &= \mathcal{R}'_t(\Phi_{(k-1)T}(\omega)(v, v'), \omega^k) \end{aligned} \right\} \quad \text{for } (k-1)T \leq t \leq kT,$$

where  $k \geq 2$ . It is straightforward to see that the Markov process  $(\mathbf{u}_t, \mathbb{P}_v)$  associated with the RDS  $\Phi_t$  is an extension for  $(u_t, \mathbb{P}_v)$ . It turns out that the recurrence and exponential squeezing properties of Definition 3.1.6 hold for it, so that we have a mixing extension for  $(u_t, \mathbb{P}_v)$ . Application of Theorem 3.1.7 enables one to conclude the proof.  $\square$

As in the case of the Navier–Stokes system perturbed by random kicks, Theorem 3.5.2 implies the exponential convergence of the probabilistic average for an observable to its mean value with respect to the stationary distribution. Namely, we have the following result that can be established by repeating the arguments of the proof of Corollary 3.4.2.

**Corollary 3.5.3.** *Under the hypotheses of Theorem 3.5.2, for any  $\gamma \in (0, 1]$  and sufficiently small  $\varkappa > 0$  there are positive constants  $C$  and  $\beta$  such that, for any  $f \in C^\gamma(H, w)$  with  $w(r) = e^{\varkappa r^2}$ , we have*

$$|\mathbb{E}_u f(u_t) - (f, \mu)| \leq C e^{-\beta t} |f|_{\gamma, w} w(|u|_2), \quad t \geq 0, \quad u \in H. \quad (3.133)$$

All the results of this chapter that concern the Navier–Stokes system were established for a general bounded domain with smooth boundary. In the case of periodic boundary conditions and white noise perturbations, one can go further and prove the uniqueness and exponential mixing for *any*  $\nu > 0$  with a fixed finite-dimensional external force. More precisely, let us consider the Navier–Stokes equation (2.98) on the standard 2D torus. The following result was established by Hairer and Mattingly [HM06, HM08, HM11]; see also the papers [AS05, AS06] by Agrachev and Sarychev dealing with the closely related question of controllability of the Navier–Stokes system by a finite-dimensional external force and the review [Kup10] by Kupiainen.

**Theorem 3.5.4.** *In the setting of Section 3.3, assume that  $x \in \mathbb{T}^2$  and  $h = 0$ . Then there is an integer  $N \geq 1$  not depending on  $\nu > 0$  such that the conclusions of Theorem 3.5.2 are true, provided that condition (3.131) holds.*

The proof of this theorem is based on the Kantorovich functional technique (see Sections 3.1 and 3.2) and some gradient estimates for the Markov semigroup  $\mathfrak{P}_t$ . The latter, in turn, uses a version of Malliavin calculus developed by Mattingly and Pardoux [MP06]. We refer the reader to the above-cited original works for more details of the proof. Let us also mention the paper [AKSS07], which proves that, for a large class stochastic nonlinear systems with finite-dimensional noises, the corresponding solutions are such that projections of their distributions to any finite-dimensional subspace are absolutely continuous with respect to the Lebesgue measure.

### 3.5.3 Convergence for functionals on higher Sobolev spaces

Theorems 3.2.9, 3.3.1, 3.4.1, and 3.5.2 show that if a random perturbation acting on the Navier–Stokes system is sufficiently non-degenerate, then for a large class of continuous functionals on  $H$ , the ensemble average converges, as  $t \rightarrow \infty$ , to the space average with respect to the stationary distribution. At the same time, many important characteristics of fluid flows, such as the enstrophy or correlation tensors, are continuous functions on higher Sobolev spaces, but not on  $H$ . It is therefore desirable to have similar convergence results for that type of functionals. In this subsection, we prove that convergence of functionals defined

on higher Sobolev spaces can easily be obtained from what we have already established, provided that the external force is regular in the space variables. To simplify the presentation, we confine ourselves to the case of a torus with a spatially regular white force.

We thus consider Eq. (2.98) on  $\mathbb{T}^2$  and assume that, for some integer  $k \geq 1$ , we have

$$h \in V^k, \quad \mathfrak{B}_k = \sum_{j=1}^{\infty} \alpha_j^k b_j^2 < \infty. \quad (3.134)$$

For any integer  $m \geq 0$  and  $\varkappa > 0$ , we define weight functions  $w^{(m)}(r) = (1+r)^m$  and  $w_\rho(r) = e^{\rho r^2}$ .

**Theorem 3.5.5.** *In addition to the hypotheses of Theorem 3.5.2, let us assume that (3.134) holds for some integer  $k \geq 1$ . Then for any  $\gamma \in (0, 1]$  and  $\rho > 0$ , any integer  $m \geq 0$ , and any functional  $f \in C^\gamma(V^k, w^{(m)})$  we have*

$$|\mathbb{E}_u f(u_t) - (f, \mu)| \leq C e^{-\beta t} |f|_{\gamma, w^{(m)} w_\rho}(|u|_2), \quad t \geq 0, \quad u \in H, \quad (3.135)$$

where  $C$  and  $\beta$  are positive constants not depending on  $f$ .

For instance, if (3.134) holds with  $k = 1$ , then the result applies to the enstrophy functional  $\Omega(u) = \frac{1}{2} |\operatorname{curl} u|_2^2$ . Furthermore, if (3.134) holds with  $k = 2$ , then convergence (3.135) is valid for the *velocity correlation tensors*

$$f^n(u) = u^{i_1}(x_1) \cdots u^{i_n}(x_n), \quad (3.136)$$

where  $i_l \in \{1, 2\}$  and  $x_1, \dots, x_n \in \mathbb{T}^2$ . Or one can take for  $f$  a structure function  $f(u) = |u(x + \ell) - u(x)|^p$  with  $x \in \mathbb{T}^2$ ,  $\ell \in \mathbb{R}^2$ , and  $p \in \mathbb{N}$ .

To prove Theorem 3.5.5, we shall need the following lemma, which is of independent interest and will also be used in Chapter 4.

**Lemma 3.5.6.** *Suppose that (3.134) holds for some integer  $k \geq 0$ . Then the following assertions hold.*

- (i) *For any  $\gamma \in (0, 1]$  and sufficiently small  $\varkappa > 0$  there exist positive constants  $\alpha \leq \gamma$  and  $\rho$  such that  $\mathfrak{P}_t f \in C^\alpha(H, w_\rho)$  for any  $f \in C^\gamma(H, w_\varkappa)$ , and the norm of the linear operator*

$$\mathfrak{P}_t : C^\gamma(H, w_\varkappa) \rightarrow C^\alpha(H, w_\rho) \quad (3.137)$$

*is bounded uniformly on any compact interval  $[0, T]$ . Moreover, the constants  $\alpha$  and  $\rho$  can be chosen in such a way that  $\rho \rightarrow 0$  as  $\varkappa \rightarrow 0$ .*

- (ii) *For any integer  $m \geq 0$  and any  $\gamma \in (0, 1]$  and  $\rho > 0$ , there is  $\alpha \in (0, \gamma]$  such that  $\mathfrak{P}_t f \in C^\alpha(H, w_\rho)$  for any  $f \in C^\gamma(V^k, w^{(m)})$  and  $t > 0$ . Moreover, the norm of the linear operator*

$$\mathfrak{P}_t : C^\gamma(V^k, w^{(m)}) \rightarrow C^\alpha(H, w_\rho) \quad (3.138)$$

*is bounded uniformly on any compact interval  $[\tau, T]$ , where  $\tau = 0$  for  $k = 0$  and  $\tau > 0$  for  $k \geq 1$ .*

*Proof of Theorem 3.5.5.* Let us note that, by the semigroup property, for  $t \geq 1$  we have

$$\mathbb{E}_u f(u_t) = \mathfrak{P}_t f(u) = (\mathfrak{P}_{t-1}(\mathfrak{P}_1 f))(u).$$

In view of Lemma 3.5.6 (ii), the function  $\mathfrak{P}_1 f$  belongs to  $C^\alpha(H, w_\rho)$ , where  $\rho > 0$  can be chosen arbitrarily small. Applying Corollary 3.5.3 to  $\mathfrak{P}_1 f$ , we see that

$$\begin{aligned} |\mathbb{E}_u f(u_t) - (f, \mu)| &= |(\mathfrak{P}_{t-1}(\mathfrak{P}_1 f))(u) - (\mathfrak{P}_1 f, \mu)| \\ &\leq C e^{-\beta(t-1)} |\mathfrak{P}_1 f|_{\alpha, w_\rho} w_\rho(|u|_2). \end{aligned}$$

Since operator (3.138) is continuous, the right-hand side of this inequality can be estimated by that of (3.135).  $\square$

*Proof of Lemma 3.5.6. (i)* Let us fix arbitrary constants  $\gamma \in (0, 1]$  and  $\varkappa > 0$  and a functional  $f \in C^\gamma(H, w_\varkappa)$  with norm  $|f|_{w_\varkappa, \gamma} \leq 1$ . The continuity of operator (3.137) and the uniform boundedness of its norm will be established if we show that

$$|\mathfrak{P}_t f(u)| \leq C e^{\rho|u|_2^2}, \quad (3.139)$$

$$|\mathfrak{P}_t f(u) - \mathfrak{P}_t f(v)| \leq C |u - v|_2^\alpha e^{bt + \rho(|u|_2^2 + |v|_2^2)}, \quad (3.140)$$

where  $u, v \in H$ ,  $0 \leq t \leq T$ , and  $b, C$  are some positive constants depending only on  $\gamma$  and  $\varkappa$ . Inequality (3.139) with  $\rho = \varkappa$  follows immediately from (2.125). To prove (3.140), let us denote by  $u_t$  and  $v_t$  the trajectories of (2.98) issued from the initial points  $u$  and  $v$ , respectively. Then, by (2.54), the difference  $u_t - v_t$  satisfies the inequality

$$|u_t - v_t|_2 \leq |u - v|_2 \exp\left(C_1 \int_0^t \|u_s\|_1^2 ds\right), \quad t \geq 0. \quad (3.141)$$

Now note that, for any  $\alpha \in (0, \gamma]$ , we have

$$\begin{aligned} |f(u_t) - f(v_t)| &\leq |u_t - v_t|_2^\alpha (w_\varkappa(|u_t|_2) + w_\varkappa(|v_t|_2)) \\ &\leq C_2 |u_t - v_t|_2^\alpha (w_{2\varkappa}(|u_t|_2) + w_{2\varkappa}(|v_t|_2)). \end{aligned} \quad (3.142)$$

Combining inequalities (3.141) and (3.142) and choosing  $\alpha \leq \gamma$  so small that  $\alpha C_1 \leq 2\varkappa\nu$ , we derive

$$|\mathfrak{P}_t f(u) - \mathfrak{P}_t f(v)| \leq \mathbb{E} |f(u_t) - f(v_t)| \leq C_3 |u - v|_2^\alpha \mathbb{E} e^{2\varkappa\xi_t}, \quad (3.143)$$

where we set

$$\xi_t = |u_t|_2^2 + |v_t|_2^2 + \nu \int_0^t \|u_s\|_1^2 ds.$$

Inequality (2.130) implies that

$$\mathbb{P}_u\{\xi_t \geq |u|_2^2 + |v|_2^2 + 2\mathfrak{B}_0 t + z\} \leq e^{-cz}, \quad t \geq 0, \quad z \in \mathbb{R},$$

where  $c > 0$  does not depend on  $t, z$  and  $u, v \in H$ . Therefore, if  $0 < \varepsilon < c$ , then

$$\mathbb{E} e^{\varepsilon \xi_t} \leq C_4 \exp(2c\mathfrak{B}_0 t + c(|u|_2^2 + |v|_2^2)).$$

Hence, if  $2\kappa \leq \varepsilon$ , then

$$\mathbb{E} e^{2\kappa \xi_t} \leq (\mathbb{E} e^{\varepsilon \xi_t})^{2\kappa/\varepsilon} \leq C_4 \exp(bt + \rho(|u|_2^2 + |v|_2^2)),$$

where  $b = 4c\kappa\mathfrak{B}_0/\varepsilon$  and  $\rho = 2\kappa c/\varepsilon$ . Comparing this with (3.143), we arrive at (3.140). The explicit form of the constant  $\rho$  implies that it goes to zero with  $\kappa$ .

(ii) We shall confine ourselves to proving the required properties under a slightly more restrictive condition. Namely, we shall assume that (3.134) holds with  $k$  replaced by  $k + 1$ . The proof in the general case is based on an estimate for the  $V^k$ -norm of the difference between two solutions; cf. (2.55) for the case  $k = 1$ . That estimate can be obtained by taking the scalar product in  $H$  of (2.163) with the function  $2tL^k u$  and carrying out some arguments similar to those in Section 2.6. Since the corresponding estimates are technically rather complicated, we do not present them here, leaving the proof of assertion (ii) in full generality to the reader as an exercise.

We need to show that if  $f \in C^\gamma(V^k, w^{(m)})$  is a functional whose norm does not exceed 1, then inequalities (3.139) and (3.140) hold for  $t \in [\tau, T]$  with some positive constants  $C$  and  $b$ . The first of them follows immediately from (2.138). To prove the second, we use the interpolation inequality (1.7) and the Hölder continuity of  $f$  to write

$$\begin{aligned} |f(u_t) - f(v_t)| &\leq \|u_t - v_t\|_k^\gamma (w^{(m)}(\|u_t\|_k) + w^{(m)}(\|v_t\|_k)) \\ &\leq C_5 |u_t - v_t|_2^\alpha (1 + \|u_t\|_{k+1} + \|v_t\|_{k+1})^{m+1}, \end{aligned}$$

where  $\alpha \in (0, \frac{\gamma}{k+1}]$  is arbitrary, and we keep notation used in the proof of (i). Combining this with (3.141), we derive

$$\begin{aligned} |\mathfrak{P}_t f(u) - \mathfrak{P}_t f(v)| &\leq \mathbb{E} |f(u_t) - f(v_t)| \\ &\leq C_5 |u - v|_2^\alpha \mathbb{E} \left\{ (1 + \|u_t\|_{k+1} + \|v_t\|_{k+1})^{m+1} \exp\left(\alpha C_1 \int_0^t \|u_s\|_1^2 ds\right) \right\}. \end{aligned}$$

Applying Schwarz's inequality to the expectation, using (2.138) with  $k$  and  $m$  replaced by  $k + 1$  and  $m + 1$ , and repeating the argument used in the proof of (i), we see that (3.140) holds for sufficiently small  $\alpha > 0$ . This completes the proof of the lemma.  $\square$

### 3.5.4 Mixing for Navier–Stokes equations perturbed by a compound Poisson process

In this section, we consider the Navier–Stokes system (3.84) in which  $\eta$  is the time derivative of a compound Poisson process (2.68). It is straightforward to check

that a solution of (3.84), (2.68) satisfies a deterministic Navier–Stokes system between two consecutive random times  $t_{k-1}$  and  $t_k$ , while at the point  $t = t_k$  it has a jump of size  $\eta_k$ . Let us set  $\mathcal{T}_\omega = \{t_k(\omega), k \geq 1\}$ . The following result is a consequence of the existence and uniqueness of a solution for the Navier–Stokes system and the well-known property of independence of increments for a compound Poisson process.

*Exercise 3.5.7.* Show that, for any  $u_0 \in H$ , problem (3.84), (2.68) has a unique solution  $u(t, x)$  issued from  $u_0$  whose almost every trajectory belongs to the space  $C(\mathbb{R}_+ \setminus \mathcal{T}_\omega; H) \cap L^2_{\text{loc}}(\mathbb{R}_+; V)$  and can be chosen to be right continuous with left limits existing at the points  $t_k$ . Prove also that the corresponding solutions form a Markov process in  $H$ .

Let us denote by  $\mathfrak{P}_t$  and  $\mathfrak{P}_t^*$  the Markov semigroups associated with (3.84), (2.68). The following result on existence, uniqueness, and polynomial mixing of a stationary measure was established by Nersesyan [Ner08] in the case of the complex Ginzburg–Landau equation, which is technically more complicated than the 2D Navier–Stokes system.

**Theorem 3.5.8.** *Assume that the random kicks  $\eta_k$  satisfy the hypotheses imposed in Theorem 3.2.9. Then for any  $\nu > 0$  there is an integer  $N \geq 1$  such that if  $b_j \neq 0$  for  $j = 1, \dots, N$ , then the following assertions hold.*

**Existence and uniqueness:** *The semigroup  $\mathfrak{P}_t^*$  has a unique stationary distribution  $\mu$ .*

**Polynomial mixing:** *For any  $p \geq 1$  there is a constant  $C_p > 0$  such that for any function  $f \in L_b(H)$  we have*

$$|\mathfrak{P}_t f(u) - (f, \mu)| \leq C_p t^{-p} (1 + |u|_2^2) \|f\|_L, \quad t \geq 0.$$

We confine ourselves to outlining the main idea of the proof, referring the reader to the original work [Ner08] for the details. Let us set  $\tilde{u}_k = u(t_k)$  and note that

$$\tilde{u}_k = S_{\tau_k}(\tilde{u}_{k-1}) + \eta_k, \quad k \geq 1. \quad (3.144)$$

It turns out that the trajectories  $\{\tilde{u}_k\}$  form a Markov chain in  $H$ , whose properties are rather similar to those of the RDS (3.50) with unbounded<sup>8</sup> kicks. Using the ideas of Section 3.4, it is possible to prove that (3.144) has a unique stationary measure  $\tilde{\mu}$ , which is mixing faster than any negative degree of  $k$ . Once this property is established, one can go back to the original problem and prove the uniqueness and polynomial mixing of a stationary measure  $\mu$  for it. Moreover, the measures  $\mu$  and  $\tilde{\mu}$  are linked by the following *Khas'minskii relation*:

$$(f, \mu) = (\mathbb{E} \tau_1)^{-1} \mathbb{E}_{\tilde{\mu}} \int_0^{\tau_1} f(u_t) dt, \quad f \in C_b(H). \quad (3.145)$$

<sup>8</sup>Note that even though the kicks  $\eta_k$  are bounded, in the case when they are non-zero almost all trajectories of (3.144) are unbounded in  $H$ . This is due to the fact that on any finite interval the system can receive arbitrarily large number of kicks.

*Exercise 3.5.9.* Prove that if  $\tilde{\mu}$  is a stationary measure for (3.144), then the measure  $\mu$  defined by (3.145) is stationary for the Markov process associated with (3.84), (2.68).

### 3.5.5 Description of some results on uniqueness and mixing for other PDE's

The methods developed in this chapter for proving the uniqueness of a stationary measure and exponential mixing apply, roughly speaking, to random perturbations of any dissipative PDE with a finite-dimensional attractor. Let us describe briefly some results concerning the complex Ginzburg–Landau equation:

$$\dot{u} - (\nu + i)\Delta u + (\alpha + i\lambda)|u|^{2p}u = h(x) + \eta(t, x), \quad u|_{\partial Q} = 0. \quad (3.146)$$

Here  $\nu > 0$ ,  $\alpha \geq 0$ ,  $\lambda > 0$ , and  $p > 0$  are some parameters,  $Q \subset \mathbb{R}^d$  is a bounded domain with smooth boundary,  $h$  is a deterministic function, and  $\eta$  is a random force from one of the classes described in Section 2.2. We thus consider only the case when the conservative term  $i|u|^{2p}u$  is defocusing. In this case, the Cauchy problem for (3.146) is well posed in  $H_0^1 = H_0^1(Q; \mathbb{C})$  under the following assumptions on the parameter  $p$ :

$$p < \infty \quad \text{for } d = 1, 2, \quad p \leq \frac{2}{d-2} \quad \text{for } d \geq 3. \quad (3.147)$$

The behaviour of solutions for (3.146) heavily depends on whether the parameter  $\alpha$  is positive or not. Roughly speaking, when  $\alpha > 0$ , the dissipative term  $\alpha|u|^{2p}u$  compensates for the strong nonlinear effects due to the conservative term, and the uniqueness of a stationary measure and exponential mixing can be established under essentially the same hypotheses. On the other hand, without strong nonlinear dissipation (i.e., for  $\alpha = 0$ ), these properties are not known to hold, unless either the random perturbation is bounded or the parameter  $p$  satisfies much more restrictive conditions. Let us discuss the latter situation in more details, referring the reader to the papers [Hai02b, Oda08] for the easier case of equations with a strong nonlinear dissipation. Note that Eq. (3.146) with  $\alpha = 0$  is a stochastic PDE with a conservative nonlinearity and a linear damping. So this is an analogue of the main topic of this book—the stochastic Navier–Stokes system.

We assume that  $p$  satisfies the following inequalities (cf. (3.147)):

$$p \leq 1 \quad \text{for } d = 1, \quad p < 1 \quad \text{for } d = 2, \quad p \leq \frac{2}{d} \quad \text{for } d \geq 3. \quad (3.148)$$

Concerning the external force, we suppose that  $h \in H^1(Q; \mathbb{C})$  and  $\eta(t, x)$  is a random process of the form (2.66), where  $\{b_j\} \subset \mathbb{R}$  is a sequence decaying faster than every negative degree of  $j$ ,  $\{e_j\}$  is an eigenbasis for the Dirichlet Laplacian in  $Q$ , and  $\{\beta_j = \beta_j^1 + i\beta_j^2\}$  is a sequence of independent complex Brownian motions. A proof of the following result can be found in [Shi08].

**Theorem 3.5.10.** *Under the above hypotheses, assume that  $\alpha = 0$  and  $b_j \neq 0$  for all  $j \geq 1$ . Then for any  $\nu > 0$  and  $\lambda > 0$  the Markov process associated with (3.146) has a unique stationary measure  $\mu \in \mathcal{P}(H_0^1)$ . Moreover, for any function  $f \in L_b(L^2(Q; \mathbb{R}))$  and any solution  $u(t)$  for (3.146), we have*

$$|\mathbb{E}f(u(t)) - (f, \mu)| \leq Ce^{-\gamma t} \|f\|_L (1 + \mathbb{E}|u(0)|_2^2), \quad t \geq 0,$$

where  $C$  and  $\gamma > 0$  are universal constants.

A brief discussion on other results concerning the uniqueness and mixing for various classes of randomly forced PDE's can be found in the section *Notes and comments* at the end of this chapter.

### 3.5.6 An alternative proof of mixing for kick force models

In Section 3.2, we established the uniqueness of a stationary distribution and the exponential mixing for the Markov chain defined by (3.50). The proof presented there is based on a coupling argument which enabled one to establish a contraction in a space of measures. In this section, we outline an alternative proof of the uniqueness and mixing, which is based on a Lyapunov–Schmidt reduction and a version of the Ruelle–Perron–Frobenius theorem. To simplify the presentation, we impose more restrictive hypotheses on the system than in Section 3.2; however, essentially the same proof works in the general case.

We consider the discrete-time RDS (3.50), in which  $S : H \rightarrow H$  is a locally Lipschitz compact mapping in a Hilbert space  $H$  with a norm  $\|\cdot\|$  and  $\{\eta_k\}$  is a sequence of i.i.d. random variables in  $H$  supported by a compact subset. We assume that

$$\|S(u)\| \leq q \|u\| \quad \text{for all } u \in H, \quad (3.149)$$

where  $q < 1$  does not depend on  $u$ . Then, for a sufficiently large  $R > 0$ , the ball  $B_H(R)$  is an invariant absorbing set for (3.50). It follows that the RDS (3.50) has a compact absorbing set  $X \subset B_H(R)$ , and we can confine our consideration to  $X$ . We assume, in addition, that there is a constant  $\gamma < 1$  and a projection  $P : H \rightarrow H$  with a finite-dimensional range  $E \subset H$  such that inequality (3.55) holds with  $\mathcal{A} = B_H(r)$  and  $r = R/(1 - q)$ , the random variables  $(I - P)\eta_k$  are almost surely zero, and the law of  $P\eta_k$  has a  $C^1$ -smooth density  $\rho$  with respect to the Lebesgue measure on  $E$  such that  $\rho(0) > 0$ . Let us denote by  $P_k(u, \Gamma)$  the transition function of the Markov chain (3.50) restricted to  $X$  and by  $\mathfrak{P}_k$  and  $\mathfrak{P}_k^*$  the corresponding Markov semigroups. Our aim is to establish the following result.

**Theorem 3.5.11.** *Under the above hypotheses, there is a unique stationary distribution  $\mu \in \mathcal{P}(X)$  for  $\mathfrak{P}_k^*$ , and for any  $\lambda \in \mathcal{P}(X)$  we have*

$$\|\mathfrak{P}_k^* \lambda - \mu\|_L^* \rightarrow 0 \quad \text{as } k \rightarrow \infty. \quad (3.150)$$

Existence of a stationary measure is a consequence of the Bogolyubov-Krylov argument (see Section 2.5.2). We now outline the proof of the uniqueness under



the additional hypothesis that  $(I - P)\eta_k = 0$  almost surely. It is divided into three steps.

*Step 1: Lyapunov–Schmidt reduction.* Let us consider the projection of Eq. (3.50) to the orthogonal complement of  $E$ :

$$w_k = QS(v_{k-1} + w_{k-1}). \quad (3.151)$$

Here  $Q = I - P$ ,  $v_k = Pu_k$ , and  $w_k = Qu_k$ . We denote by  $\mathbf{X}$  and  $\mathbf{Y}$  the spaces of sequences  $\mathbf{v} = (v_k, k \in \mathbb{Z}_-)$  and  $\mathbf{w} = (w_k, k \in \mathbb{Z}_-)$  with  $v_k \in B_E(R)$  and  $w_k \in B_{E^\perp}(qr)$ , endowed with the metric

$$d_{\mathbf{X}}(\mathbf{u}^1, \mathbf{u}^2) = \sup_{k \leq 0} (e^{\alpha k} \|u_k^1 - u_k^2\|), \quad \mathbf{u}^i = (u_k^i, k \in \mathbb{Z}_-),$$

where  $\alpha > 0$  is such that  $\gamma e^\alpha < 1$ . Using the contraction mapping principle, it is straightforward to prove that for any  $\mathbf{v} \in \mathbf{X}$  there is a unique  $\mathbf{w} := \mathcal{W}(\mathbf{v}) \in \mathbf{Y}$  satisfying (3.151) for each  $k \leq 0$ . Moreover, the mapping  $\mathcal{W} : \mathbf{X} \rightarrow \mathbf{Y}$  is globally Lipschitz-continuous. (These properties are essentially a consequence the Foias–Prodi estimates discussed in Section 2.1.8.)

Let us define a Markov RDS in  $\mathbf{X}$  by the formula

$$\mathbf{v}^k = (\mathbf{v}^{k-1}, PS(v_0^{k-1} + \mathcal{W}(\mathbf{v}^{k-1})) + \eta_k), \quad (3.152)$$

where  $\mathcal{W} = \mathcal{W}(\mathbf{v})$  stands the zeroth component of  $\mathcal{W}(\mathbf{v})$ , and given two elements  $\mathbf{v} = (v_k, k \in \mathbb{Z}_-) \in \mathbf{X}$  and  $v \in B_E(R)$ , we write  $(\mathbf{v}, v)$  for the sequence  $(\dots, v_{-1}, v_0, v) \in \mathbf{X}$ . Given a stationary measure  $\mu$  for  $\mathfrak{P}_k^*$ , we denote by  $\boldsymbol{\mu} \in \mathcal{P}(\mathbf{X})$  a unique law satisfying the relation

$$\begin{aligned} \boldsymbol{\mu}(\{(v_k, k \in \mathbb{Z}_-) \in \mathbf{X} : (v_{-l}, \dots, v_0) \in \Gamma\}) \\ = \int_{\Gamma} \mu(du_{-l}) P_1(u_{-l}, du_{1-l}) \dots P_1(u_1, du_0) \quad \text{for any } \Gamma \in \mathcal{B}(H^{l+1}). \end{aligned}$$

Here  $P_k(u, \cdot)$  is the transition function of the family of Markov chains corresponding to (3.50), and

$$\tilde{\Gamma} = \{(u_{-l}, \dots, u_0) \in H^{l+1} : (Pu_{-l}, \dots, Pu_0) \in \Gamma\}.$$

In other words,  $\boldsymbol{\mu}$  is the restriction to  $\mathbb{Z}_-$  of the law for the projection to  $E$  of a stationary trajectory of (3.50) with distribution  $\mu$ . The following proposition establishes a relationship between (3.50) and (3.152). Its proof can be found in [KS00].

**Proposition 3.5.12.** *Let  $\mu \in \mathcal{P}(H)$  be a stationary measure for  $\mathfrak{P}_k^*$ . Then  $\boldsymbol{\mu}$  is a stationary measure for (3.152).*

Thus, the uniqueness of a stationary measure for  $\mathfrak{P}_k^*$  will be established if we establish this property for the Markov chain (3.152). In the next step, we formulate an abstract result that gives a sufficient condition for the uniqueness and mixing of a stationary measure for a Markov semigroup on a compact space.

*Step 2: Ruelle–Perron–Frobenius type theorem.* Let  $\mathbf{X}$  be a compact metric space and let  $\mathbf{P}_k(\mathbf{v}, \Gamma)$  be a Markov transition function on  $\mathbf{X}$ . We denote by  $\mathbf{P}_k : C(\mathbf{X}) \rightarrow C(\mathbf{X})$  and  $\mathbf{P}_k^* : \mathcal{P}(\mathbf{X}) \rightarrow \mathcal{P}(\mathbf{X})$  the corresponding Markov semigroups. Recall that a family  $\mathcal{C} \subset C(\mathbf{X})$  is said to be *determining* if any two measures  $\mu, \nu \in \mathcal{P}(\mathbf{X})$  satisfying the relation  $(f, \mu) = (f, \nu)$  for all  $f \in \mathcal{C}$  coincide. The following result gives a sufficient condition for the uniqueness and mixing for a stationary measure.

**Theorem 3.5.13.** *Suppose that the transition function satisfies the following two hypotheses.*

**Uniform Feller property.** *There is a determining family  $\mathcal{C}$  such that for any  $f \in \mathcal{C}$  the sequence  $\{\mathbf{P}_k f, k \geq 0\}$  is uniformly equicontinuous.*

**Uniform recurrence.** *There is  $\hat{\mathbf{v}} \in \mathbf{X}$  such that for any  $r > 0$  one can find an integer  $m \geq 1$  and a constant  $p > 0$  for which*

$$\mathbf{P}_k(\mathbf{u}, B_{\mathbf{X}}(\hat{\mathbf{v}}, r)) \geq p \quad \text{for } k \geq m \text{ and all } \mathbf{u} \in \mathbf{X}. \quad (3.153)$$

*Then  $\mathbf{P}_k^*$  has a unique stationary measure  $\mu \in \mathcal{P}(\mathbf{X})$ , and for any  $f \in C(\mathbf{X})$ , we have*

$$\mathbf{P}_k f \rightarrow (f, \mu) \quad \text{as } k \rightarrow \infty \text{ uniformly on } \mathbf{X}. \quad (3.154)$$

This theorem has an independent interest, and its proof is given at the end of this subsection. We now show how to apply it in our situation.

*Step 3: Verification of the hypotheses.* The uniform recurrence condition is a straightforward consequence of the fact that  $u = 0$  is a globally stable fixed point for the deterministic (nonlinear) semigroup obtained by iterations of  $S$  and the fact that the random variables  $\eta_k$  are small with a positive probability. The latter follows from the inequality  $\rho(0) > 0$ . To prove the uniform Feller property, we note that the transition function  $\mathbf{P}_k(\mathbf{v}, \cdot)$  for the Markov chain associated with (3.152) is such that

$$\mathbf{P}_k(\mathbf{v}, \Gamma) = \int_{\Gamma} \prod_{l=1-k}^0 \rho(z_l - F(\mathbf{v}, z_{1-k}, \dots, z_{l-1})) dz_{1-k} \dots dz_0, \quad (3.155)$$

where  $\Gamma \subset E^{k+1}$  is an arbitrary Borel subset,  $F(\mathbf{v}) = PS(v_0 + \mathcal{W}(\mathbf{v}))$  for  $\mathbf{v} = (v_m, m \in \mathbb{Z}_-)$ , and

$$\Gamma = \{\mathbf{v} = (v_m, m \in \mathbb{Z}_-) : (v_{-k}, \dots, v_0) \in \Gamma\}.$$

Let us denote by  $\mathcal{C}$  the space of functions  $f \in C(\mathbf{X})$  that depend on finitely many components. It follows from (3.155) that if a function  $f \in \mathcal{C}$  depends only on  $v_{-N}, \dots, v_0$ , then

$$\mathbf{P}_k f(\mathbf{v}) = \int_{E^k} \prod_{l=1-k}^0 \rho(z_l - F(\mathbf{v}, z_{1-k}, \dots, z_{l-1})) f(z_{-N}, \dots, z_0) dz_{1-k} \dots dz_0.$$

Since the mapping  $\mathcal{W}$  is Lipschitz continuity, so is  $F$ , and the above formula readily gives uniform equicontinuity of the family  $\{\mathbf{P}_k f, k \geq 0\}$  for any  $f \in \mathcal{C}$ .

Further analysis shows that the original and the reduced systems are equivalent, and therefore the convergence to the stationary measure for (3.152) implies (3.150). We refer the reader to the original work [KS00] for more details on this subject.

*Proof of Theorem 3.5.13.* Existence of a stationary measure follows immediately from the compactness of  $\mathbf{X}$  and the Bogolyubov–Krylov argument. To prove the uniqueness, it suffices to show that if  $\boldsymbol{\mu} \in \mathcal{P}(\mathbf{X})$  is a stationary distribution, then (3.154) holds for any  $f \in \mathcal{C}$ . Indeed, if (3.154) is proved and  $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{P}(\mathbf{X})$  are two stationary measures, then  $(f, \boldsymbol{\mu}) = (f, \boldsymbol{\nu})$  for any  $f \in \mathcal{C}$ . Since  $\mathcal{C}$  is a determining family, we conclude that  $\boldsymbol{\mu} = \boldsymbol{\nu}$ .

Suppose now that  $\boldsymbol{\mu} \in \mathcal{P}(\mathbf{X})$  is a stationary measure. Let  $f \in \mathcal{C}$  be an arbitrary function. We wish to prove (3.154). There is no loss of generality in assuming that  $0 \leq f \leq 1$ . It follows from the uniform Feller property that  $\mathbf{P}_{k_n} f$  converges uniformly to a limit  $g \in C(\mathbf{X})$  for some sequence  $k_n \rightarrow \infty$ . We claim that  $g$  must be constant. If this assertion is proved, then the obvious relation  $(\mathbf{P}_1 h, \boldsymbol{\mu}) = (h, \boldsymbol{\mu})$ , which holds for any  $h \in C(\mathbf{X})$ , implies that the constant must be equal to  $(f, \boldsymbol{\mu})$ , whence we conclude that the whole sequence  $\{\mathbf{P}_k f\}$  converges to  $(f, \boldsymbol{\mu})$ .

Suppose that  $g \not\equiv C$ . We can assume that  $m_n = k_{n+1} - k_n \rightarrow \infty$  as  $n \rightarrow \infty$ . It is easy to see that  $\mathbf{P}_{m_n} g \rightarrow g$  uniformly on  $\mathbf{X}$ . Since  $g$  is not constant, then  $g(\hat{\mathbf{v}})$  (where  $\hat{\mathbf{v}}$  is defined in the uniform recurrence condition) is either smaller than the maximum of  $g$  or greater than the minimum of  $g$ . Assume that the first case is true. Denoting by  $\mathbf{u} \in \mathbf{X}$  a point where  $g$  attains its maximum, we can find positive constants  $\delta$  and  $r$  such that

$$g(\mathbf{u}) \geq g(\mathbf{z}) + \delta \quad \text{for } \mathbf{z} \in B_{\mathbf{X}}(\hat{\mathbf{v}}, r). \quad (3.156)$$

We now write

$$\begin{aligned} \mathbf{P}_{m_n} g(\mathbf{u}) &= \int_{\mathbf{X}} P_{m_n}(\mathbf{u}, d\mathbf{z}) g(\mathbf{z}) \\ &= \int_{B_{\mathbf{X}}(\hat{\mathbf{v}}, r)} P_{m_n}(\mathbf{u}, d\mathbf{z}) g(\mathbf{z}) + \int_{B_{\mathbf{X}}^c(\hat{\mathbf{v}}, r)} P_{m_n}(\mathbf{u}, d\mathbf{z}) g(\mathbf{z}). \end{aligned}$$

Using (3.156), we derive

$$\begin{aligned} \mathbf{P}_{m_n} g(\mathbf{u}) &\leq P_{m_n}(\mathbf{u}, B_{\mathbf{X}}(\hat{\mathbf{v}}, r))(g(\mathbf{u}) - \delta) + P_{m_n}(\mathbf{u}, B_{\mathbf{X}}^c(\hat{\mathbf{v}}, r))g(\mathbf{u}) \\ &\leq g(\mathbf{u}) - \delta P_{m_n}(\mathbf{u}, B_{\mathbf{X}}(\hat{\mathbf{v}}, r)). \end{aligned} \quad (3.157)$$

Since  $m_n \rightarrow \infty$ , inequality (3.153) implies that

$$\mathbf{P}_{m_n} g(\mathbf{u}) \leq g(\mathbf{u}^+) - \delta p \quad \text{for } n \gg 1. \quad (3.158)$$

Passing to the limit as  $n \rightarrow \infty$ , we arrive at a contradiction. This completes the proof of the theorem.  $\square$

### 3.6 Appendix: some technical proofs

In this section, we have compiled some relatively difficult proofs that are not of primary importance. The reader not interested in technical details may safely skip this section.

#### 3.6.1 Proof of Lemma 3.3.11

Compared to the case of  $H^1$ -regular external force, the additional difficulty is that the resolving operator is not continuous as a mapping defined on the space of  $H$ -valued paths. To overcome this problem, we first rewrite the Navier–Stokes system as an equation with random coefficients and then repeat essentially the same scheme, using the continuous dependence of solutions on coefficients. The proof is divided into three steps.

*Step 1.* Let us recall a decomposition of  $u(t; v)$  described in Section 2.4. We write  $u(t; v) = z(t) + \tilde{v}(t)$ , where  $z(t)$  is the solution of Stokes' equation (2.100) with zero initial condition and

$$\tilde{v}(t) = \mathcal{R}_t(v, z)$$

is the solution of problem (2.111), (2.112); see Proposition 2.4.5. Recall that the space  $\mathcal{X}_T = C(0, T; H) \cap L^2(0, T; V)$  and denote by  $\dot{\mathcal{X}}_T$  the space of functions in  $\mathcal{X}_T$  vanishing at  $t = 0$ . We claim that the following two properties hold.

- (a) For any  $\hat{u} \in V$  there is a continuous operator  $Z_{\hat{u}} : H \rightarrow \dot{\mathcal{X}}_1$ ,  $v \mapsto Z_{\hat{u}}(v; t)$ , that is uniformly Lipschitz on bounded subsets and is such that the solution of (2.100) with  $\eta = \partial_t Z_{\hat{u}} + \nu L Z_{\hat{u}}$  equals  $\hat{u}$  at  $t = 1$ :

$$\mathcal{R}_1(v, Z_{\hat{u}}(v)) + Z_{\hat{u}}(v; 1) = \hat{u}. \quad (3.159)$$

- (b) Let  $\mathcal{O} \subset \dot{\mathcal{X}}_1$  be any neighbourhood of the origin and let  $\mathcal{K} \subset \dot{\mathcal{X}}_1$  be a compact set. Then

$$\inf_{\hat{z} \in \mathcal{K}} \mathbb{P}\{z - \hat{z} \in \mathcal{O}\} > 0. \quad (3.160)$$

Taking these properties for granted, let us complete the proof of (3.101). Without loss of generality, we can assume that  $\hat{u} \in V$ . Since the operator  $Z$  is continuous, there is a constant  $\rho > 0$  such that the image of the compact set  $B_V(R) \subset H$  under  $Z$  is contained in the ball  $B_{\dot{\mathcal{X}}_1}(\rho)$ . The Lipschitz property on bounded subsets implies that

$$|\mathcal{R}_1(v, z_1) - \mathcal{R}_1(v, z_2)|_2 \leq C \|z_1 - z_2\|_{\mathcal{X}_1}$$

for any  $v \in B_V(R)$  and  $z_1, z_2 \in B_{\dot{\mathcal{X}}_1}(\rho)$ , where  $C \geq 1$  is a constant depending only on  $R$ . In particular, taking  $z_2 = Z_{\hat{u}}(v)$ , we see that

$$|\mathcal{R}_1(v, z_1) - \mathcal{R}_1(v, Z_{\hat{u}}(v))|_2 \leq \frac{\varepsilon}{2} \quad \text{for } \|z_1 - Z_{\hat{u}}(v)\|_{\mathcal{X}_1} \leq \frac{\varepsilon}{2C}.$$

It follows from property (a) that

$$|u(1; v) - \hat{u}|_2 \leq |z(1) - Z_{\hat{u}}(v; 1)|_2 + |\mathcal{R}_1(v, z_1) - \mathcal{R}_1(v, Z_{\hat{u}}(v))|_2$$

Combining the last two inequalities, we conclude that

$$\mathbb{P}\{|u(1; v) - \hat{u}|_2 < \varepsilon\} \geq \mathbb{P}\left\{\|z - Z_{\hat{u}}(v)\|_{\mathcal{X}_1} < \frac{\varepsilon}{2C}\right\} \quad \text{for any } v \in B_V(R). \quad (3.161)$$

Applying property (b) in which  $\mathcal{K} = Z_{\hat{u}}(B_V(R))$  and  $\mathcal{O} \subset \dot{\mathcal{X}}_1$  is the open ball of radius  $\frac{\varepsilon}{2C}$  centred at zero, we see that the right-hand side of (3.161) is separated from zero uniformly in  $v \in B_V(R)$ . Thus, to complete the proof of (3.101), it suffices to establish properties (a) and (b).

*Step 2: Proof of (a).* Let  $\tilde{w} \in \mathcal{H}$  be the solution of the problem

$$\partial_t \tilde{w} + \nu L \tilde{w} = 0, \quad \tilde{w}(0) = v.$$

For any  $v \in H$ , define a function  $w \in \mathcal{H}$  by the relation  $w(t) = \chi(t)\tilde{w}(t) + t\hat{u}$ , where  $\chi \in C^\infty(\mathbb{R})$  is an arbitrary function such that  $\chi(t) = 1$  for  $t \leq 0$  and  $\chi(t) = 0$  for  $t \geq 1$ . Then  $w(0) = v$ ,  $w(1) = \hat{u}$ , and

$$\eta := \partial_t w + \nu L w + B(w) - h \in L^2(0, 1; V^*).$$

Moreover,  $\eta$  is uniformly Lipschitz continuous in  $v \in H$  on bounded subsets as a function with range in  $L^2(0, 1; V^*)$ . Let us define  $Z_{\hat{u}}(v) \in \dot{\mathcal{X}}_1$  as the unique solution of the problem

$$\dot{Z} + \nu LZ = \eta(t), \quad Z(0) = 0.$$

Then

$$w(t) = Z_{\hat{u}}(v; t) + \mathcal{R}_t(v, Z_{\hat{u}}(v)),$$

whence we conclude that (3.159) holds. The Lipschitz continuity of  $Z_{\hat{u}}(v)$  follows from a similar property for the mapping  $v \mapsto \eta$ .

*Step 3: Proof of (b).* Let us note that the process  $z(t)$  can be written in the form (cf. (2.101))

$$z(t) = \sum_{j=1}^{\infty} b_j \int_0^t e^{-\alpha_j \nu(t-s)} d\beta_j(s) e_j(x), \quad 0 \leq t \leq 1,$$

where  $\{e_j\}$  stands for the family of normalised eigenfunctions of the Stokes operator,  $\alpha_j$  denotes the eigenvalue corresponding to  $e_j$ , and the series above converges in  $L^2(\Omega; \dot{\mathcal{X}}_1)$ . We claim that the support of the law for the process  $\{z(t), 0 \leq t \leq 1\}$  regarded as a random variable in  $\dot{\mathcal{X}}_1$  coincides with the entire space. To prove this, it suffices to show that, for any  $\hat{z} \in C(0, 1; V)$  vanishing at zero and any  $\varepsilon > 0$ , we have

$$\mathbb{P}\{\|z - \hat{z}\|_{\mathcal{X}_1} < \varepsilon\} > 0. \quad (3.162)$$

Let us fix a function  $\hat{z}$  and write it in the form

$$\hat{z}(t) = \sum_{j=1}^{\infty} \hat{z}_j(t) e_j(x),$$

where  $\hat{z}_j(t) = \langle z(t), e_j \rangle$ , and the series converges in  $V$  uniformly in  $t \in [0, 1]$ . For any integer  $N \geq 1$ , we have

$$z(t) - \hat{z}(t) = D_N(t) + r_N(t) + \hat{r}_N(t), \quad (3.163)$$

where we set

$$\begin{aligned} D_N(t) &= \sum_{j=1}^N \left( b_j \int_0^t e^{-\alpha_j \nu(t-s)} d\beta_j(s) - \hat{z}_j(t) \right) e_j(x), \\ r_N(t) &= \sum_{j=N+1}^{\infty} b_j \int_0^t e^{-\alpha_j \nu(t-s)} d\beta_j(s) e_j(x), \\ \hat{r}_N(t) &= \sum_{j=N+1}^{\infty} \hat{z}_j(t) e_j(x). \end{aligned}$$

Let us estimate each term on the right-hand side of (3.163). Since  $\{e_j\}$  is an orthogonal system with respect to the scalar product  $\langle Lu, u \rangle^{1/2}$  and  $\|e_j\|_V^2 = \alpha_j$  for any  $j \geq 1$ , we have

$$\|D_N(t)\|_V \leq \sqrt{\alpha_N} \max_{1 \leq j \leq N} \left( \sup_{0 \leq t \leq 1} b_j \left| \int_0^t e^{-\alpha_j \nu(t-s)} d\beta_j(s) - \hat{z}_j(t) \right| \right).$$

Since  $\hat{z}_j$  is a continuous function on  $[0, 1]$  vanishing at zero, combining this inequality with Proposition 7.4.2, we see that

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq 1} \|D_N(t)\|_V < \frac{\varepsilon}{3} \right\} > 0 \quad \text{for any } N \geq 1. \quad (3.164)$$

Now let  $\mathbf{P}_N : H \rightarrow H$  be the orthogonal projection onto the vector span of  $e_1, \dots, e_N$ . Then one can find an integer  $N_1(\varepsilon) \geq 1$  such that

$$\sup_{0 \leq t \leq 1} \|\hat{r}_N(t)\|_V = \sup_{0 \leq t \leq 1} \|(I - \mathbf{P}_N)\hat{z}(t)\|_V \leq \frac{\varepsilon}{3} \quad \text{for } N \geq N_1(\varepsilon).$$

Furthermore, the technique developed in the proof of Proposition 2.4.2 (see derivation of (2.109)) enables one to show that

$$\mathbb{E} \|r_N\|_{\mathcal{X}_1}^2 = \mathbb{E} \|(I - \mathbf{P}_N)z\|_{\mathcal{X}_1}^2 \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

Therefore there is an integer  $N_2(\varepsilon) \geq 1$  such that

$$\mathbb{P} \left\{ \|r_N\|_{\mathcal{X}_1} \leq \frac{\varepsilon}{3} \right\} > 0 \quad \text{for } N \geq N_2(\varepsilon).$$

Combining the above estimates and noting that the random variables  $D_N$  and  $r_N$  are independent, we arrive at the required inequality (3.162).

We are now ready to prove (3.160). Since  $\text{supp } \mathcal{D}(z) = \mathcal{X}_1$ , for any neighbourhood of zero  $\mathcal{O} \subset \dot{\mathcal{X}}_1$  the function  $p(\hat{z}) = \mathbb{P}\{z - \hat{z} \in \mathcal{O}\}$  is positive at any point of  $\dot{\mathcal{X}}_1$ . Furthermore, in view of a well-known property of the weak convergence of measures (e.g., see Theorem 11.1.1 in [Dud02]),  $p$  is lower semi-continuous. It follows that  $p$  is separated from zero on any compact subset of  $\dot{\mathcal{X}}_1$ . This completes the proof of Lemma 3.3.11.

### 3.6.2 Recurrence for Navier–Stokes equations with unbounded kicks

This section is devoted to the proof of Proposition 3.4.3. For any  $d > 0$ , we introduce the stopping time

$$\tau_d = \inf\{k \geq 0 : \mathbf{u}_k \in \mathbf{B}_d\}.$$

We wish to show that, for any  $\mathbf{u} \in \mathbf{H}$  and  $d > 0$ , the random variable  $\tau_d$  is  $\mathbb{P}_{\mathbf{u}}$ -almost surely finite, and there are positive constants  $\delta$  and  $C > 0$  such that

$$\mathbb{E}_{\mathbf{u}} \exp(\delta \tau_d) \leq C(1 + |\mathbf{u}|), \quad (3.165)$$

where  $|\mathbf{u}| = |u|_2 + |u'|_2$ .

As in the case of the Navier–Stokes system with spatially regular white noise, inequality (3.165) is a consequence of the following two lemmas (cf. Section 3.3.2).

**Lemma 3.6.1.** *There is  $R_* > 0$  such that*

$$\mathbb{P}_{\mathbf{u}}\{\tau_R < \infty\} = 1 \quad \text{for any } \mathbf{u} \in \mathbf{H}, R \geq R_*. \quad (3.166)$$

*Moreover, there are positive constants  $K$  and  $\gamma$  such that*

$$\mathbb{E}_{\mathbf{u}} \exp(\gamma \tau_R) \leq 1 + KR^{-1}|\mathbf{u}| \quad \text{for any } \mathbf{u} \in \mathbf{H}, R \geq R_*. \quad (3.167)$$

**Lemma 3.6.2.** *For any positive constants  $\mathfrak{B}$ ,  $R$ ,  $d$ , and  $\nu$  there is  $p > 0$  and integers  $N_0 \geq 1$  and  $m \geq 1$  such that if (3.121) holds, then*

$$\mathbb{P}_{\mathbf{u}}\{\mathbf{u}_m \in \mathbf{B}_d\} \geq p \quad \text{for any } \mathbf{u} \in \mathbf{B}_R. \quad (3.168)$$

Once these lemmas are established, inequality (3.165) will follow by exactly the same argument as in Section 3.3.2. Thus, we shall confine ourselves to the proof of the above lemmas.

*Proof of Lemma 3.6.1.* We use a well-known argument based on the existence of a Lyapunov function. It follows from Proposition 2.3.3 that

$$\mathbb{E}_{\mathbf{u}} |\mathbf{u}_k| \leq q(|\mathbf{u}| \vee R_*) \quad \text{for any } \mathbf{u} \in \mathbf{H}, \quad (3.169)$$

where  $q < 1$  and  $R_* > 0$  are some constants not depending on  $\mathbf{u}$ . Let us fix any  $R \geq R_*$  and set

$$p_k(\mathbf{u}) = \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\{\tau_R > k\}} | \mathbf{u}_k|).$$

The Markov property and inequality (3.169) imply that

$$\begin{aligned} p_{k+1}(\mathbf{u}) &\leq \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\{\tau_R > k\}} (\mathbb{E}_{\mathbf{u}} |\mathbf{u}_{k+1}| | \mathcal{F}_k)) = \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\{\tau_R > k\}} (\mathbb{E}_{\mathbf{v}} |\mathbf{u}_1|) |_{\mathbf{v}=\mathbf{u}_k}) \\ &\leq \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\{\tau_R > k\}} q(|\mathbf{u}_k| \vee R_*)) = q p_k(\mathbf{u}), \end{aligned}$$

where we used the fact that  $|\mathbf{u}_k| > R \geq R_*$  on the set  $\{\tau_R > k\}$ . Iterating the above inequality and noting that  $p_0(\mathbf{u}) \leq |\mathbf{u}|$ , we obtain

$$p_k(\mathbf{u}) \leq q^k |\mathbf{u}| \quad \text{for all } k \geq 0, \mathbf{u} \in \mathbf{H}.$$

It follows that

$$\mathbb{P}_{\mathbf{u}}\{\tau_R > k\} \leq R^{-1} \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\{\tau_R > k\}} |\mathbf{u}_k|) \leq R^{-1} q^k |\mathbf{u}|, \quad (3.170)$$

whence we conclude that (3.166) holds.

We now prove (3.167). Let  $\gamma > 0$  be so small that  $e^{\gamma} q < 1$ . Then, using (3.166) and (3.170), we derive

$$\begin{aligned} \mathbb{E}_{\mathbf{u}} \exp(\gamma \tau_R) &= \sum_{k=0}^{\infty} \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\{\tau_R = k\}} \exp(\gamma \tau_R)) \\ &\leq 1 + \sum_{k=1}^{\infty} e^{\gamma k} \mathbb{P}_{\mathbf{u}}\{\tau_R > k-1\} \\ &\leq 1 + R^{-1} |\mathbf{u}| \sum_{k=1}^{\infty} e^{\gamma k} q^{k-1}. \end{aligned}$$

The series on the right-hand side of this inequality converges in view of the choice of  $\gamma$ , and we obtain (3.167).  $\square$

*Proof of Lemma 3.6.2. Step 1.* Inequality (3.168) will be established if for any constants  $R > d > 0$  we find  $\varepsilon > 0$  and  $q < 1$  such that

$$\mathbb{P}\{|\mathcal{R}(v, v')|_2 \vee |\mathcal{R}'(v, v')|_2 \leq (q(|v| \vee |v'|)) \vee d\} \geq \varepsilon \quad \text{for } (v, v') \in \mathbf{B}_R. \quad (3.171)$$

Indeed, let us denote by  $\mathbf{G}(v, v')$  the event on the left-hand side of this inequality and choose an integer  $m \geq 1$  so that  $qd < q^m R \leq d$ . Denoting  $\mathbf{B}^{(k)} = \mathbf{B}_{q^k R}$ ,  $k \geq 0$ , we have  $\mathbf{B}^{(m)} \subset \mathbf{B}_d$ . Combining (3.171) with the Markov property, for  $1 \leq k \leq m-1$  we derive

$$\begin{aligned} \mathbb{P}_{\mathbf{u}}\{\mathbf{u}_k \in \mathbf{B}^{(k)}\} &\geq \mathbb{P}_{\mathbf{u}}(\mathbf{G}(\mathbf{u}_{k-1}) \cap \{\mathbf{u}_{k-1} \in \mathbf{B}^{(k-1)}\}) \\ &= \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\mathbf{B}^{(k-1)}}(\mathbf{u}_{k-1}) \mathbb{P}_{\mathbf{u}}\{\mathbf{G}(\mathbf{u}_{k-1}) | \mathcal{F}_{k-1}\}) \\ &= \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\mathbf{B}^{(k-1)}}(\mathbf{u}_{k-1}) \mathbb{P}(\mathbf{G}(v, v')) |_{(v, v')=\mathbf{u}_{k-1}}) \\ &\geq \varepsilon \mathbb{P}_{\mathbf{u}}\{\mathbf{u}_{k-1} \in \mathbf{B}^{(k-1)}\}. \end{aligned}$$



A similar argument shows that

$$\mathbb{P}_{\mathbf{u}}\{\mathbf{u}_m \in \mathbf{B}_d\} \geq \varepsilon \mathbb{P}_{\mathbf{u}}\{\mathbf{u}_{m-1} \in \mathbf{B}^{(m-1)}\}$$

Combining the above inequalities, we see that

$$\mathbb{P}_{\mathbf{u}}\{\mathbf{u}_m \in \mathbf{B}_d\} \geq \varepsilon^m \mathbb{P}_{\mathbf{u}}\{\mathbf{u}_0 \in \mathbf{B}^{(0)}\} = \varepsilon^m.$$

This implies inequality (3.168) with  $p = \varepsilon^m$ .

*Step 2.* We now prove (3.171). For  $\delta > 0$ , we set

$$A_\delta = \{\mathbf{P}_N \mathcal{R}(v, v') = \mathbf{P}_N \mathcal{R}'(v, v'), |\mathbf{P}_N \mathcal{R}(v, v')|_2 \leq \delta, |\mathbf{Q}_N \zeta(v, v')|_2 \leq \delta\},$$

where  $\zeta(v, v') = \mathcal{R}(v, v') - S(v)$ . Suppose that for any  $\delta > 0$  and a constant  $\varepsilon' = \varepsilon'(\delta) > 0$  we have

$$\mathbb{P}(A_\delta) \geq \varepsilon' \quad \text{for } (v, v') \in \mathbf{B}_R. \quad (3.172)$$

In this case, using the regularising property of  $S$  and the Foias–Prodi estimate (see Theorems 2.1.18 and 2.1.30), for any  $\omega \in A_\delta$  we derive

$$\begin{aligned} |\mathcal{R}(v, v')|_2 &\leq |\mathbf{P}_N \mathcal{R}(v, v')|_2 + |\mathbf{Q}_N \mathcal{R}(v, v')|_2 \\ &\leq \delta + C_1 \alpha_N^{-1/2} \|S(v)\|_1 + |\mathbf{Q}_N \zeta(v, v')|_2 \\ &\leq 2\delta + C_2 \alpha_N^{-1/2} |v|_2, \\ |\mathcal{R}'(v, v')|_2 &\leq |\mathcal{R}(v, v')|_2 + |\mathcal{R}'(v, v') - \mathcal{R}(v, v')|_2 \\ &\leq 2\delta + C_2 \alpha_N^{-1/2} |v|_2 + C_3(R) \alpha_N^{-1/2} |v - v'|_2. \end{aligned}$$

Combining these two inequalities, we obtain

$$|\mathcal{R}(v, v')|_2 \vee |\mathcal{R}'(v, v')|_2 \leq 2\delta + C_4(R) \alpha_N^{-1/2} (|v|_2 \vee |v'|_2).$$

Choosing  $N \geq 1$  and  $\delta > 0$  so that  $C_4(R) \alpha_N^{-1/2} \leq 1/2$  and  $\delta = d/6$ , we see that the right-hand side of this inequality does not exceed  $(\frac{3}{4}(|v|_2 \vee |v'|_2)) \vee d$ . We thus arrive at (3.171) with  $q = \frac{3}{4}$ .

*Step 3.* It remains to establish (3.172). The construction of  $(\mathcal{R}, \mathcal{R}')$  implies that  $\mathbf{Q}_N \zeta$  is independent of  $(\mathbf{P}_N \mathcal{R}, \mathbf{P}_N \mathcal{R}')$ . Hence,

$$\mathbb{P}(A_\delta) \geq \mathbb{P}(A_\delta^{(1)}) \mathbb{P}(A_\delta^{(2)}), \quad (3.173)$$

where we set

$$\begin{aligned} A_\delta^{(1)} &= \{|\mathbf{Q}_N \zeta(v, v')|_2 \leq \delta\}, \\ A_\delta^{(2)} &= \{|\mathbf{P}_N \mathcal{R}(v, v')|_2 \leq \delta, \mathbf{P}_N \mathcal{R}(v, v') = \mathbf{P}_N \mathcal{R}'(v, v')\}. \end{aligned}$$

Since  $\mathcal{D}(\zeta(v, v')) = \mathcal{D}(\eta_1)$ , the conditions imposed on  $b_j$  and  $\mathcal{D}(\xi_{jk})$  imply that

$$\mathbb{P}(A_\delta^{(1)}) \geq \varepsilon_1(\delta) > 0 \quad \text{for any } \delta > 0.$$

Furthermore, recalling that the pair  $(\mathbf{P}_N \mathcal{R}(v, v'), \mathbf{P}_N \mathcal{R}'(v, v'))$  is a maximal coupling for  $(\chi_N^v, \chi_N^{v'})$ , where  $\chi_N^u = (\mathbf{P}_N)_* \mathcal{D}(S(u) + \eta_1)$ , and using Lemma 1.2.26, we derive

$$\mathbb{P}(A_\delta^{(2)}) = (\chi_N^v \wedge \chi_N^{v'})(B_{H(N)}(\delta)).$$

Now note that  $\chi_N^u(dx) = p(x - \mathbf{P}_N S(u)) dx$ , where  $p$  stands for the density of the law for  $\mathbf{P}_N \eta_1$  with respect to the Lebesgue measure on  $H(N)$  (see (3.83)). Since  $p(x) > 0$  almost everywhere, we conclude that

$$\inf_{(v, v') \in \mathbf{B}_R} (\chi_N^v \wedge \chi_N^{v'})(B_{H(N)}(\delta)) \geq \varepsilon_2(\delta) > 0 \quad \text{for any } \delta > 0.$$

The required inequality (3.172) follows now from (3.173).  $\square$

### 3.6.3 Exponential squeezing for Navier–Stokes equations with unbounded kicks

In this section, we establish Proposition 3.4.4. Its proof is divided into three steps.

*Step 1.* For any  $M > 0$ , we define the stopping times

$$\begin{aligned} \tau(M) &= \min \left\{ k \geq 1 : \sum_{j=0}^k \int_0^T (\|S_t(u_j)\|_1^2 + \|S_t(u'_j)\|_1^2) dt > M(k+1) \right\}, \\ \sigma(M) &= \tau(M) \wedge \min \{ k \geq 1 : \mathbf{P}_N u_k \neq \mathbf{P}_N u'_k \}. \end{aligned}$$

We claim that for any  $M > 0$  there is an integer  $N \geq 1$  such that, for any  $d > 0$  and an appropriate choice of the parameter  $c > 0$  entering the definition of  $\sigma_0$ , relation (3.129) holds, that is,  $\sigma \leq \sigma_0$  almost surely. Indeed, by the Foias–Prodi estimate (2.62), on the set  $\{\sigma(M) > k\}$  we have

$$|u_k - u'_k|_2 \leq (C\alpha_N^{-1/2})^k \exp(C(M+1)k) |u_0 - u'_0|_2 \leq 2d(Ce^{C(M+1)}\alpha_N^{-1/2})^k.$$

Choosing  $c = 3d$  and  $N \geq 1$  so large that  $Ce^{C(M+1)}\alpha_N^{-1/2} \leq e^{-1}$ , we see that (3.129) is true.

*Step 2.* We now prove that, for sufficiently large  $M > 0$  and  $N \geq 1$ , the stopping time  $\sigma(M)$  satisfies (3.32) – (3.34). To this end, it suffices to show that

$$\mathbb{P}_{\mathbf{u}}\{\sigma(M) = k\} \leq 2e^{-2k} \quad \text{for } k \geq 1, \mathbf{u} \in \mathbf{B}_d. \quad (3.174)$$

Indeed, if these relations are established, then

$$\begin{aligned} \mathbb{P}_{\mathbf{u}}\{\sigma(M) = \infty\} &= 1 - \sum_{k=1}^{\infty} \mathbb{P}_{\mathbf{u}}\{\sigma(M) = k\} \geq 1 - 2 \sum_{k=1}^{\infty} e^{-2k} > 0, \\ \mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\{\sigma(M) < \infty\}} e^{\sigma(M)}) &= \sum_{k=1}^{\infty} e^k \mathbb{P}_{\mathbf{u}}\{\sigma(M) = k\} = 2 \sum_{k=1}^{\infty} e^{-k} < \infty. \end{aligned}$$

Furthermore, by inequality (2.82), we have

$$\mathbb{E}_{\mathbf{u}}(\mathbb{I}_{\{\sigma(M) < \infty\}} |\mathbf{u}_{\sigma(M)}|^2) \leq \mathbb{E}_{\mathbf{u}} |\mathbf{u}|^2 + C_1 \leq 2d^2 + C_1, \quad \mathbf{u} \in \mathbf{B}_d,$$

where  $C_1 > 0$  is a constant depending on  $\mathbb{E} |\eta_1|_2^4$  and  $\sum_k (\mathbb{P}_{\mathbf{u}} \{\sigma(M) = k\})^{1/2}$ . Thus, it remains to prove (3.174).

*Step 3.* Let us introduce the events

$$A(k) = \{\mathbb{P}_N u_k = \mathbb{P}_N u'_k\}, \quad \bar{A}(k) = \bigcap_{l=1}^k A(l),$$

$$D(k) = \left\{ \sum_{j=0}^k \int_0^T (\|S_t(u_j)\|_1^2 + \|S_t(u'_j)\|_1^2) dt \leq M(k+1) \right\}.$$

We argue by induction on  $k \geq 1$ . For  $k = 1$ , we have

$$\{\sigma(M) = 1\} \subset \{\tau(M) = 1\} \cup A(1)^c. \quad (3.175)$$

Proposition 2.3.8 implies that if  $M > 0$  is sufficiently large, then

$$\mathbb{P}_{\mathbf{u}} \{\tau(M) = 1\} \leq e^{-2}. \quad (3.176)$$

Furthermore, since  $S : H \rightarrow H$  is Lipschitz-continuous on bounded subsets (see Proposition 2.1.25), we have

$$\|S(v) - S(v')\|_2 \leq C_2 \|v - v'\| \leq 2C_2 d \quad \text{for } (v, v') \in \mathbf{B}_d.$$

Assertion (iii) of Lemma 3.2.6 now implies that

$$\mathbb{P}_{\mathbf{u}} \{A(1)^c\} = \mathbb{P}_{\mathbf{u}} \{\mathbb{P}_N u_k \neq \mathbb{P}_N u'_k\} \leq C_3(N)d \leq e^{-2} \quad \text{for } d \ll 1.$$

Combining this with (3.175) and (3.176), we obtain (3.174) with  $k = 1$ .

We now assume that  $k = m \geq 2$  and for  $1 \leq k \leq m-1$  inequality (3.174) is already established. It follows from the definition of  $\sigma(M)$  that (cf. (3.175))

$$\{\sigma(M) = m\} \subset \{\tau(M) = m\} \cup B(m), \quad (3.177)$$

where  $B(m) = \bar{A}(m-1) \cap A(m)^c \cap \{\tau(M) > m\}$ . Let us estimate the probability of the two events on the right-hand side of (3.177). Proposition 2.3.8 implies that, for  $M \gg 1$ , we have

$$\mathbb{P}_{\mathbf{u}} \{\tau(M) = m\} \leq C_4 Q^m e^{-\delta M(m+1)} \leq e^{-2m}. \quad (3.178)$$

Furthermore, using (3.174) with  $1 \leq k \leq m-1$ , for any  $\mathbf{u} \in \mathbf{B}_d$  we derive

$$\mathbb{P}_{\mathbf{u}} \{\bar{A}(m-1) \cap D(m-1)\} \geq \mathbb{P}_{\mathbf{u}} \{\sigma(M) \geq m\} \geq 1 - 2 \sum_{k=1}^{m-1} e^{-2k} \geq \frac{1}{2}. \quad (3.179)$$

Therefore, since  $\{\tau(M) > m\} \subset D(m-1)$ , we can write

$$\begin{aligned} \mathbb{P}_{\mathbf{u}}(B(m)) &\leq \mathbb{P}_{\mathbf{u}}\{\bar{A}(m-1) \cap A(m)^c \cap D(m-1)\} \\ &\leq 2 \mathbb{P}_{\mathbf{u}}\{A(m)^c \mid \bar{A}(m-1) \cap D(m-1)\}, \end{aligned} \quad (3.180)$$

where we used (3.179) to get the second inequality. By the Markov property, the conditional probability  $\mathbb{P}_{\mathbf{u}}\{A(m)^c \mid \mathcal{F}_{m-1}\}$  depends only on  $\mathbf{u}_{m-1}$ . The construction of coupling operators and Lemma 3.2.6 imply that

$$\mathbb{P}_{\mathbf{u}}\{A(m)^c \mid \mathcal{F}_{m-1}\} \leq C_5(N) |S(u_{m-1}) - S(u'_{m-1})|_2.$$

It follows from Theorem 2.1.30 and Remark 2.1.31 that, for  $N \gg 1$  and  $\mathbb{P}_{\mathbf{u}}$ -a.e.  $\omega \in \bar{A}(m-1) \cap D(m-1)$ ,

$$|S(u_{m-1}) - S(u'_{m-1})|_2 \leq 2d (C\alpha_N^{-1/2})^m e^{C(M+1)m} \leq C_6 d e^{-2m}.$$

Combining these two inequalities with (3.180), we see that

$$\mathbb{P}_{\mathbf{u}}(B(m)) \leq C_7(N) d e^{-2m}.$$

Choosing  $d \ll 1$  and evoking (3.177) and (3.178), we get the required inequality (3.174) with  $k = m$ . The proof of Proposition 3.4.4 is complete.

### 3.7 Relevance of the results for physics

The uniqueness of a stationary distribution and the convergence to this distribution are very important for the statistical hydrodynamics. In the physical literature, these two properties are usually postulated. For instance, in the classical book of G. K. Batchelor [Bat82], it is taken for granted that *dynamical systems with a large number of degrees of freedom, and with coupling between these degrees of freedom, approach a statistical state which is independent (partially, if not wholly) of the initial condition*; see there pp. 6–7. The results of this chapter rigorously prove this postulate for the periodic 2D turbulent flows (with zero space average) and for flows in bounded 2D domains, driven by a non-degenerate random forces.

Namely, we show that if in Eq. (0.1) the random force  $f(t, x)$  is a kick force, or a white in time force, or a compound Poisson process (see Introduction and Section 2.2) and if it is non-degenerate, then Eq. (0.1) has a unique stationary distribution  $\mu$ , which is a measure in the space of divergence-free vector fields  $\{u(x)\}$ . The distribution of every solution  $u(t, x)$  for (0.1) at time  $t$  (which is a measure in the same function space) converges to  $\mu$  exponentially fast when  $t \rightarrow \infty$ . The measure  $\mu$  is one-smoother than the force  $f(t, x)$ , regarded as a function of the space variables, and any functional  $g(u)$  is integrable with respect to  $\mu(du)$ , unless  $g(u)$  grows at infinity very fast (i.e., faster than  $\exp(\delta|u|_2^2)$  with some fixed positive  $\delta$ ). So practically for any observable  $g(u)$  and for any solution  $u(t)$  of (0.1) our results imply that the averaged in ensemble

observable  $\mathbb{E} g(u(t))$  approaches  $\int g(u) \mu(du)$  exponentially fast when the time  $t$  goes to infinity, provided that the random force  $f$  is as above and is sufficiently smooth in the  $x$ -variable. In particular, this is true if the observable  $g(u)$  is the energy of a flow  $u$ , or its enstrophy, or the correlation tensor  $u^i(x)u^j(y)$  (where  $i$  and  $j$  are 1 or 2, and  $x, y$  are any fixed points).

In the case of periodic boundary conditions when the flows  $u$  have zero space-average, the stationary measure  $\mu$  is space-homogeneous if so is the force  $f$ . In this situation, we have

$$\int g(u) \mu(du) = \int \langle g \rangle(u) \mu(du),$$

where  $\langle g \rangle$  is the space-averaged observable defined by

$$\langle g \rangle(u) = (2\pi)^{-2} \int_{\mathbb{T}^2} g(u_y(\cdot)) dy, \quad u_y(x) = u(x + y)$$

(we assumed that both periods equal  $2\pi$ ). In particular, if  $g(u)$  is a linear functional, then  $\langle g \rangle(u) = g\langle u \rangle = 0$ , so  $\int g(u) \mu(du) = 0$ . For example, choosing  $g(u) = u^j(x)$  we get

$$\int u^j(x) \mu(du) = 0 \quad \text{for } j = 1, 2 \text{ and for any } x.$$

That is, for the homogeneous turbulence the mean value of velocity, evaluated at any fixed point, vanishes.

## Notes and comments

The investigation of uniqueness of a stationary distribution for the Markov process generated by the randomly forced 2D Navier–Stokes equation began in 1995 by the pioneering article of Flandoli and Maslowski [FM95]. The main result of their paper is Theorem 3.5.1 of this section on uniqueness of a stationary distribution and the mixing property in total variation norm. The proof in [FM95] is obtained by adjusting Doob’s argument [Doo48] to the infinite-dimensional system defined by the Navier–Stokes equations (3.84), under the crucial assumption that the noise is a rough function of  $x$  (i.e. the coefficients  $b_j$  decay to zero very slowly). The Flandoli–Maslowski theorem was generalised in various directions. Ferrario [Fer97, Fer99] gave a simpler proof of Theorem 3.5.1, allowing for more regular noises. Bricmont, Kupiainen, and Lefevre [BKL01] established the uniqueness of a stationary measure and exponential mixing for the Navier–Stokes system perturbed by a rough kick force. Goldys and Maslowski [GM05] proved a similar result in the case of white noise force with low spatial regularity. Da Prato, Debussche, Tubaro [DDT05], Barbu, Da Prato [BD07], and Da Prato, Röckner, Rozovskii, Wang [DRRW06] studied the Burgers, magneto-hydrodynamics and porous media equations in analogous settings. Eckmann,

Hairer [EH01] and Hairer [Hai02a] considered a one-dimensional Ginzburg–Landau equation with a mildly degenerate rough forcing and proved the exponential convergence to a unique stationary measure in the total variation norm.

From the physical point of view, the roughness condition imposed on the noise is not very natural, and much efforts were spent to remove it. In the context of randomly forced PDE's, a first result in this direction was obtained by Sinai [Sin91]. He studied the Burgers equation on the real line, perturbed by a white in time space-periodic random force. Sinai proved that there exists a space-periodic stationary measure which attracts distributions of a large class of solutions (these solutions are not space-periodic, and the uniqueness of a space-periodic stationary measure was not established). A few years later Mattingly [Mat99], then a PhD student of Sinai, established that the 2D Navier–Stokes system considered on the torus  $\mathbb{T}^2$ , with a large viscosity and spatially regular white force, has a unique stationary measure, which attracts distributions of all solutions.<sup>9</sup>

A first result on the uniqueness of a stationary measure for Navier–Stokes equations with any positive viscosity and a smooth in  $x$  random force was obtained by Kuksin and Shirikyan [KS00] (see also [KS01b, KS02b] for some further developments). They proved Theorem 3.5.11, applicable to a large class of dissipative PDE's perturbed by a smooth random kick force which is non-zero in sufficiently many modes. The proof is based on two key ingredients: a Lyapunov–Schmidt reduction and a version of the Ruelle–Perron–Frobenius theorem. The first of them is a well-known tool for studying dissipative PDE's and was used, for instance, in the theory of inertial manifolds. It is a consequence of the Foias–Prodi estimates and enables one to reduce the Navier–Stokes equations to a finite-dimensional system with memory (which may be regarded as an abstract Gibbs system in the sense of Bowen [Bow75]). The second ingredient provides a sufficient condition for the uniqueness of a stationary measure for Markov semigroups. Various versions of this result were known earlier (see [Rue68, LY94, Sza97]), however, the uniform Feller property introduced in [KS00] was a crucial point and turned out to be useful in other situations; e.g., see [LS06]. The simple proof of the Ruelle–Perron–Frobenius theorem given in Section 3.5.6 seems to be new. Next, Weinan E, Mattingly, Sinai [EMS01] and Bricmont, Kupiainen, Lefevre [BKL02] studied the stochastic Navier–Stokes system in the case when the space variables  $x$  belong to the torus and the right-hand side is white noise in time. They showed that there is at most one stationary measure. Moreover, it was established in [BKL02] that the distributions of solutions converge to the stationary measure exponentially fast. These two works are similar to [KS00]: they also are based on the Lyapunov–Schmidt reduction, followed by a study of the resulting finite-dimensional system with memory. The papers [EMS01, BKL02] treat the case when the random force

---

<sup>9</sup>The large viscosity case is much simpler since, in this situation, inequality (2.130) implies that any two solutions with non-random initial data converge exponentially fast almost surely; cf. Exercise 2.5.9.

has a large, but finite number of excited modes (so it is a random trigonometric polynomial in  $x$ ). Important contribution of [EMS01] to this field of research is the introduction of Girsanov's theorem as a tool for studying ergodicity of stochastic PDE's with spatially regular white noise.

A different approach for investigating mixing behaviour of randomly forced PDE's was developed by Kuksin and Shirikyan [KS01a], Mattingly [Mat02b], Masmoudi and Young [MY02], and Hairer [Hai02b]. It is based on the concept of coupling and enables one to obtain stronger results with shorter proofs. Namely, it was established in [KS01a, KPS02, Kuk02b] that the Navier–Stokes system perturbed by a bounded kick force is exponentially mixing. The approach presented in Sections 3.1.1 and 3.2 is taken from these articles. In particular, the method of Kantorovich functional as tool to study mixing for stochastic Navier–Stokes system was introduced in [KPS02, Kuk02b], while the Lipschitz–dual distance as a suitable metric for convergence to a stationary distribution in the context of randomly forced PDE's was suggested in [KS01a]. These two tools are now commonly used in works of this field. A different, but closely related method was developed independently by Masmoudi and Young [MY02]. Mattingly [Mat02b] combined the coupling with a stopping time technique to establish the exponential mixing for the Navier–Stokes system with periodic boundary conditions and a random force which is a white noise in time and a trigonometric polynomial in space of a large finite degree. Similar techniques were used by Hairer [Hai02b] to study some parabolic PDE's with nonlinear dissipation. Various versions of the coupling argument were applied by Kuksin and Shirikyan [KS02a] to prove the exponential mixing for the Navier–Stokes system in a bounded domain with a spatially regular white noise of infinite dimension, by Shirikyan [Shi04] to study the Navier–Stokes system perturbed by an unbounded kick force, and by Debussche and Odasso [DO05] to investigate a 1D Schrödinger equation with a linear damping. Further results on the problem of mixing for various types of randomly forced PDE's can be found in [Shi05a, Shi06b, Ner08, Oda08, Shi08]. The presentation of Sections 3.1.2 and 3.1.3 follows essentially [Shi04, Shi05a, Shi08], and the proof of the stability property of Section 3.3.3 is based on some ideas taken from [EMS01, Oda08]. We refer the reader to the review papers [ES00, Kuk02a, Bri02, Mat03, Hai05, Shi05b] for further discussions of the results on the uniqueness and mixing for stochastic PDE's with a sufficiently non-degenerate smooth noise.

All the above results concern the case when all determining modes of the problem in question are perturbed by a random force. In particular, when these results are applied to the Navier–Stokes system, in order to have a unique stationary measure, one has to assume that the dimension of the random perturbation goes to infinity when the viscosity goes to zero; e.g., every mode of the random force is excited. An important question is whether one can prove the uniqueness and mixing for finite-dimensional noises for any positive value of the viscosity. A progress in this direction was achieved by Hairer and Mattingly in 2006. Combining the method of the Kantorovich functional with an infinite-dimensional Malliavin calculus (developed in [MP06]), they proved

in [HM06, HM08, HM11] that if the problem is studied on the standard torus and the deterministic part of the noise is zero, then the uniqueness of a stationary measure is true for a four-dimensional white-noise force and the exponential convergence of other solutions to the stationary measure holds in the dual-Lipschitz metric; see Theorem 3.5.4. Two key questions remain open in this context. Namely, does a similar result hold when: (a) the deterministic part of the external force is non-zero; (b) the random perturbation is not white in time. Note that the answers to both questions is positive for a different type of degenerate noises. Namely, it was proved recently by Shirikyan [Shi11a] that the Navier–Stokes equations with a random perturbation supported by a given subregion of the space domain possess the exponential mixing property under some mild non-degeneracy assumptions. We also mention the work [AKSS07], in which it is proved that, for a large class of low-dimensional random forces, the support of the law  $\mathcal{D}(u(t))$  of a solution of the 2D Navier–Stokes system is infinite-dimensional for any  $t > 0$ ; see Section 6.3.1 for more details.



## Chapter 4

# Ergodicity and limiting theorems

In this chapter, we study limiting theorems for the 2D Navier–Stokes system with random perturbations. To simplify the presentation, we shall confine ourselves to the case of spatially regular white noise; however, all the results remain true for random kick forces. The first section is devoted to derivation of the strong law of large numbers (SLLN), the law of iterated logarithm (LIL), and the central limit theorem (CLT). Our approach is based on the reduction of the problem to similar questions for martingales and an application of some general results on SLLN, LIL, and CLT. In Section 4.2, we study the relationship between stationary distribution and random attractors. Roughly speaking, it is proved that the support of the random probability measure obtained by the disintegration of the unique stationary distribution is a random point attractor for the RDS in question. The third section deals with the stationary distributions for the Navier–Stokes system perturbed by a random force depending on a parameter. We first prove that the stationary measures continuously depend on spatially regular white noise. We next consider high-frequency random kicks and show that, under suitable normalisation, the corresponding family of stationary measures converges weakly to the unique stationary distribution corresponding to the white noise perturbation. Finally, in Section 4.4, we discuss the physical relevance of the results of this chapter.

### 4.1 Ergodic theorems

#### 4.1.1 Strong law of large numbers

Let us consider the Navier–Stokes system (3.84) in which  $h \in H$  is a given function and  $\eta$  is a spatially regular white noise of the form (2.66). As was established in Theorem 3.3.1, if sufficiently many first coefficients  $b_j$  are non-zero, then Eq. (3.84) has a unique stationary distribution. Moreover, by Theorem 3.5.2,

the Markov process  $(u_t, \mathbb{P}_u)$  defined by the equation is exponentially mixing, and by Corollary 3.5.3, for any functional  $f \in C^\gamma(H, w_\varkappa)$ , where  $\gamma \in (0, 1]$  and  $w_\varkappa(r) = e^{\varkappa r^2}$  with  $0 < \varkappa \ll 1$ , and any initial datum  $v \in H$  the mean value of the corresponding trajectory  $\mathbb{E}_v f(u_t)$  converges to  $(f, \mu)$  as  $t \rightarrow \infty$  exponentially fast (see inequality (3.133)). The following theorem shows that the mean value with respect to the probability ensemble can be replaced by time average. In this case, however, the convergence to the mean value holds only at an algebraical rate.

**Theorem 4.1.1.** *Under the hypotheses of Theorem 3.5.2, there is a constant  $\varkappa > 0$  such that for any  $\varepsilon > 0$ ,  $\gamma \in (0, 1]$ ,  $v \in H$ , and  $f \in C^\gamma(H, w_\varkappa)$ , with  $\mathbb{P}_v$ -probability 1, we have*

$$\lim_{t \rightarrow \infty} t^{\frac{1}{2} - \varepsilon} \left( t^{-1} \int_0^t f(u_s) ds - (f, \mu) \right) = 0.$$

*Proof.* We first outline the main idea. Without loss of generality, we can assume that  $(f, \mu) = 0$ . We need to prove that, for any  $v \in H$ ,

$$t^{-\frac{1}{2} - \varepsilon} \int_0^t f(u_s) ds \rightarrow 0 \quad \text{with } \mathbb{P}_v\text{-probability 1.} \quad (4.1)$$

Let us denote by  $\mathcal{F}_t$  the filtration corresponding to the Markov process  $(u_t, \mathbb{P}_v)$  (see Section 1.3.1) and consider Gordin's *martingale approximation*:

$$M_t = \int_0^\infty (\mathbb{E}_v(f(u_s) | \mathcal{F}_t) - \mathbb{E}_v(f(u_s) | \mathcal{F}_0)) ds, \quad v \in H, \quad t \geq 0. \quad (4.2)$$

By Proposition 7.13.3, the process  $M_t$  is a martingale with respect to the filtration  $\{\mathcal{F}_t\}$  and the probability  $\mathbb{P}_v$ . Furthermore, relation (7.70) implies that

$$\int_0^t f(u_s) ds = \int_0^{\hat{t}} f(u_s) ds + \int_{\hat{t}}^t f(u_s) ds = M_{\hat{t}} - g(u_{\hat{t}}) + g(u_0) + d_t, \quad (4.3)$$

where  $\hat{t}$  stands for the integer part of  $t \geq 0$ ,

$$d_t = \int_{\hat{t}}^t f(u_s) ds, \quad g(u) = \int_0^\infty \mathfrak{P}_s f(u) ds. \quad (4.4)$$

Using representation (4.3) and a priori estimates for solutions of the Navier–Stokes system, we show that

$$\lim_{t \rightarrow \infty} \left| t^{-\frac{1}{2}} \int_0^t f(u_s) ds - \hat{t}^{-\frac{1}{2}} M_{\hat{t}} \right| = 0 \quad \mathbb{P}_v\text{-almost surely.} \quad (4.5)$$

Hence, to prove (4.1), it suffices to establish the SLLN for the discrete-time martingale  $M_k$ . This will be done with the help of Theorem 7.12.1. We now turn to the accurate proof divided into two steps.

*Step 1.* We first prove (4.5). In view of (4.3), it suffices to show that

$$\mathbb{P}_v \left\{ \lim_{k \rightarrow \infty} k^{-\frac{1}{2}} \left( \sup_{k \leq t \leq k+1} |d_t| + |g(u_k)| \right) = 0 \right\} = 1. \quad (4.6)$$

In what follows, we denote by  $C_i$  positive constants that may depend on  $f$ , but not on  $u$ . Let us set

$$U_k = \sup_{k \leq t \leq k+1} |u_t|_2^2, \quad D_k = \sup_{k \leq t \leq k+1} |d_t|.$$

By inequality (2.130), we can find positive constants  $C_1$  and  $\gamma$  such that

$$\mathbb{P}_v \{U_0 \geq \rho\} \leq C_1 e^{\gamma|v|_2^2 - \gamma\rho} \quad \text{for } v \in H, \rho \geq 0.$$

Combining this with the Markov property and inequality (2.125), we obtain

$$\mathbb{P}_v \{U_k \geq \rho\} = \mathbb{E}_v \mathbb{P}_{u_k} \{U_0 \geq \rho\} \leq C_1 e^{-\gamma\rho} \mathbb{E}_v e^{\gamma|u_k|_2^2} \leq C_2 (e^{\gamma|v|_2^2} + 1) e^{-\gamma\rho}. \quad (4.7)$$

On the other hand, in view of Corollary 3.5.3, we have

$$|g(v)| \leq \int_0^\infty |\mathfrak{P}_s f(v)| ds \leq C_3 e^{\varkappa|v|_2^2}, \quad v \in H. \quad (4.8)$$

Furthermore, the definition of  $d_t$  implies that

$$D_k \leq C_4 e^{\varkappa U_k} \quad \text{for } k \geq 0. \quad (4.9)$$

Combining (4.7) – (4.9), for  $\varkappa < \frac{\gamma}{4}$  we derive

$$\begin{aligned} \sum_{k=1}^{\infty} \mathbb{P}_v \{D_k + |g(u_k)| \geq k^{1/4}\} &\leq \sum_{k=1}^{\infty} \mathbb{P}_v \{C_4 e^{\varkappa U_k} + C_3 e^{\varkappa|u_k|_2^2} \geq k^{1/4}\} \\ &\leq \sum_{k=1}^{\infty} \mathbb{P}_v \{C_5 e^{\varkappa U_k} \geq k^{1/4}\} \\ &= \sum_{k=1}^{\infty} \mathbb{P}_v \{U_k \geq \frac{1}{4\varkappa} \ln k - C_6\} \\ &\leq C_7 (e^{\gamma|v|_2^2} + 1) \sum_{k=1}^{\infty} k^{-\gamma/4\varkappa} < \infty. \end{aligned}$$

Hence, by the Borel–Cantelli lemma, there is a  $\mathbb{P}_v$ -almost surely finite random integer  $k_0 \geq 1$  such that

$$\sup_{k \leq t \leq k+1} |d_t| + |g(u_k)| \leq k^{1/4} \quad \text{for } k \geq k_0.$$

This implies the required relation (4.6).

*Step 2.* We now prove that

$$\mathbb{P}_v \left\{ \lim_{k \rightarrow \infty} k^{-\frac{1}{2}-\varepsilon} M_k = 0 \right\} = 1 \quad \text{for any } v \in H.$$

In view of Theorem 7.12.1, this claim will be established if we check that the martingale differences  $X_k = M_k - M_{k-1}$  satisfies the condition

$$\sum_{k=1}^{\infty} k^{-1-\delta} \mathbb{E}_v X_k^2 < \infty \quad \text{for any } v \in H, \delta > 0. \quad (4.10)$$

It follows from (4.3) that

$$X_k = g(u_k) - g(u_{k-1}) + d_k^-, \quad d_k^- = \int_{k-1}^k f(u_s) ds. \quad (4.11)$$

By the Markov property, we have

$$\mathbb{E}_v X_k^2 = \mathbb{E}_v \mathbb{E}_v (X_k^2 | \mathcal{F}_{k-1}) = \mathbb{E}_v \varphi(u_{k-1}) = (\mathfrak{P}_{k-1} \varphi)(v), \quad (4.12)$$

where  $\varphi(v) = \mathbb{E}_v X_1^2$ . We see that

$$\sum_{k=1}^{\infty} k^{-1-\delta} \mathbb{E}_v X_k^2 = \sum_{k=1}^{\infty} k^{-1-\delta} (\mathfrak{P}_{k-1} \varphi)(v). \quad (4.13)$$

To estimate the right-hand side of this relation, let us derive an explicit formula for  $\varphi$ . It follows from (4.11) that

$$\begin{aligned} \varphi(v) &= \mathbb{E}_v g^2(u_1) + \mathbb{E}_v g^2(u_0) + \mathbb{E}_v (d_1^-)^2 \\ &\quad - 2 \mathbb{E}_v (g(u_1)g(u_0)) + 2 \mathbb{E}_v (d_1^- g(u_1)) - 2 \mathbb{E}_v (d_1^- g(u_0)). \end{aligned} \quad (4.14)$$

Now note that

$$\begin{aligned} \mathbb{E}_v g^2(u_0) &= g^2(v), & \mathbb{E}_v g^2(u_1) &= \mathfrak{P}_1 g^2(v), \\ \mathbb{E}_v (g(u_0)g(u_1)) &= g(v) \mathfrak{P}_1 g(v), & \mathbb{E}_v (d_1^- g(u_0)) &= g(v) \int_0^1 \mathfrak{P}_s f(v) ds. \end{aligned}$$

Furthermore, using the Markov property, we write

$$\begin{aligned} \mathbb{E}_v (d_1^-)^2 &= \mathbb{E}_v \left( \int_0^1 f(u_s) ds \right)^2 = \int_0^1 \int_0^1 \mathbb{E}_v (f(u_s) f(u_t)) ds dt \\ &= 2 \int_0^1 \int_0^t \mathfrak{P}_s (f \mathfrak{P}_{t-s} f)(v) ds dt, \\ \mathbb{E}_v (d_1^- g(u_1)) &= \mathbb{E}_v \left( g(u_1) \int_0^1 f(u_s) ds \right) = \int_0^1 \mathfrak{P}_s (f \mathfrak{P}_{1-s} g)(v) ds. \end{aligned}$$

We now need the following lemma, whose proof is given at the end of this subsection.

**Lemma 4.1.2.** *Under the hypotheses of Theorem 4.1.1, for any  $\gamma \in (0, 1]$  and sufficiently small  $\varkappa > 0$  there exist positive constants  $\alpha \leq \gamma$  and  $\rho$  such that the function  $g$  defined by (4.4) belongs to  $C^\alpha(H, w_\rho)$  for any  $f \in C^\gamma(H, w_\varkappa)$  satisfying the condition  $(f, \mu) = 0$ , and the linear operator taking  $f$  to  $g$  is bounded in the corresponding spaces. Moreover, the constants  $\alpha$  and  $\rho$  can be chosen in such a way that  $\rho \rightarrow 0$  as  $\varkappa \rightarrow 0$ .*

Lemmas 3.5.6 and 4.1.2 show<sup>1</sup> that all the functions on the right-hand side of (4.14) belong to the space  $C^\alpha(H, w_\rho)$  with suitable constants  $\alpha$  and  $\rho$ , and their norms are bounded by  $C|f|_{w_\varkappa, \gamma}$ . Therefore we have  $\varphi \in C^\alpha(H, w_\rho)$  and

$$|\varphi|_{w_\rho, \alpha} \leq C_7 |f|_{w_\varkappa, \gamma}. \quad (4.15)$$

Furthermore, since  $\rho \rightarrow 0$  as  $\varkappa \rightarrow 0$ , Corollary 3.5.3 imply that if  $\varkappa > 0$  is sufficiently small, then

$$|(\mathfrak{P}_k \varphi)(v) - (\varphi, \mu)| \leq C_8 (|v|_2) e^{-\beta k}, \quad k \geq 0.$$

Combining this with (4.13), we see that condition (4.10) is satisfied. This completes the proof of the theorem.  $\square$

*Proof of Lemma 4.1.2.* Since  $g$  satisfies (4.8), to prove that  $g \in C^\alpha(H, w_\rho)$ , it suffices to show that if  $|f|_{\gamma, w_\varkappa} \leq 1$  and  $(f, \mu) = 0$ , then

$$|g(u) - g(v)| \leq C_2 |u - v|^\alpha e^{\rho(|u|_2^2 + |v|_2^2)}, \quad u, v \in H. \quad (4.16)$$

By Corollary 3.5.3, for any such function  $f \in C^\gamma(H, w_\varkappa)$ , we have

$$|\mathfrak{P}_t f(u)| \leq C \exp(-\beta t + \varkappa |u|_2^2), \quad t \geq 0, \quad u \in H. \quad (4.17)$$

Combining this with (3.140), for any  $u, v \in H$  and  $T > 0$ , we derive

$$\begin{aligned} |g(u) - g(v)| &\leq \int_0^T |\mathfrak{P}_t f(u) - \mathfrak{P}_t f(v)| dt + \int_T^\infty (|\mathfrak{P}_t f(u)| + |\mathfrak{P}_t f(v)|) dt \\ &\leq C_3 |u - v|^\alpha e^{bT + \rho(|u|_2^2 + |v|_2^2)} + C_4 e^{-\beta T} (e^{\varkappa |u|_2^2} + e^{\varkappa |v|_2^2}). \end{aligned}$$

Choosing  $T = \varepsilon \varkappa (|u|_2^2 + |v|_2^2 + \ln(1 \vee |u - v|_2^{-1}))$ , where  $0 < \varepsilon \ll 1$ , we arrive at inequality (4.16) in which  $\alpha$  and  $\rho$  are replaced with  $(\alpha - \varepsilon \varkappa) \wedge (\beta \varepsilon \varkappa)$  and  $(\rho + b \varepsilon \varkappa) \vee \varkappa$ , respectively.

The fact that the linear operator taking  $f$  to  $g$  is continuous follows from inequalities (4.8) and (4.16), which are true for any  $f \in C^\gamma(H, w_\varkappa)$  such that  $|f|_{w_\varkappa, \gamma} \leq 1$  and  $(f, \mu) = 0$ . Finally, the explicit form of the constant  $\rho$  implies that it goes to zero with  $\varkappa$ .  $\square$

<sup>1</sup>Lemma 3.5.6 was proved for the Navier–Stokes system on the torus. However, when  $k = 0$ , the assertions remain true in the case of a bounded domain.

*Exercise 4.1.3.* Under the hypotheses of Theorem 4.1.1, prove that, for any function  $f \in C^\gamma(H, w_\varkappa)$  with a sufficiently small  $\varkappa > 0$  and any measure  $\lambda$  satisfying the condition  $(w_\varkappa(|u|_2), \lambda) < \infty$ , we have

$$\mathbb{P}_\lambda \left\{ k^{-1} \sum_{l=0}^{k-1} f(u_l) \rightarrow (f, \mu) \text{ as } k \rightarrow \infty \right\} = 1.$$

*Hint:* Use the same scheme as for the continuous time. In this case, the martingale approximation is defined by

$$M_k = \sum_{l=0}^{\infty} (\mathbb{E}_\lambda(f(u_l) | \mathcal{F}_k) - \mathbb{E}_\lambda(f(u_l) | \mathcal{F}_0)).$$

### 4.1.2 Law of iterated logarithm

The law of large numbers established in the previous subsection can be strengthened to the law of iterated logarithm. To simplify the presentation, we shall confine ourselves to the case of stationary solutions. The general case is briefly outlined in Exercise 4.1.6.

As before, we consider the Navier–Stokes system (3.84), in which  $\eta$  is spatially regular white noise of the form (2.66). Let  $\{\mathcal{F}_t, t \geq 0\}$  be the filtration associated with the corresponding Markov process  $(u_t, \mathbb{P}_v)$ . In what follows, we assume that the hypotheses of Theorem 3.5.2 are satisfied and denote by  $\mu \in \mathcal{P}(H)$  the unique stationary measure. Recall that, given a function  $f \in C^\gamma(H, w_\varkappa)$  with  $\varkappa \ll 1$  such that  $(f, \mu) = 0$ , we defined  $g(u) = \int_0^\infty \mathfrak{P}_s f(u) ds$ ; see (4.4). Before formulating the LIL, we establish the following auxiliary result.

**Proposition 4.1.4.** *Under the above hypotheses, we have*

$$\lim_{t \rightarrow \infty} \mathbb{E}_\mu \left( \frac{1}{\sqrt{t}} \int_0^t f(u_s) ds \right)^2 = 2(fg, \mu) =: \sigma_f^2, \quad \sigma_f \geq 0. \quad (4.18)$$

Moreover,  $\sigma_f > 0$  if  $f \not\equiv 0$  and the constants  $b_j$  entering (2.66) are all non-zero.

*Proof.* Using the Markov property and the stationarity of  $\mu$ , we write

$$\begin{aligned} \mathbb{E}_\mu \left( \int_0^t f(u_s) ds \right)^2 &= \mathbb{E}_\mu \int_0^t \int_0^t f(u_r) f(u_s) dr ds \\ &= 2 \int_0^t dr \int_r^t \mathbb{E}_\mu (f(u_r) \mathbb{E}_\mu (f(u_s) | \mathcal{F}_r)) ds \\ &= 2 \int_0^t dr \int_r^t \mathbb{E}_\mu (f(u_r) (\mathfrak{P}_{s-r} f)(u_r)) ds \\ &= 2 \int_0^t dr \int_r^t (f \mathfrak{P}_{s-r} f, \mu) ds = 2 \int_0^t (t-s) (f \mathfrak{P}_s f, \mu) ds. \end{aligned}$$

Dividing both sides of this relation by  $t$  and using (4.17) together with Lebesgue's theorem to pass to the limit as  $t \rightarrow \infty$ , we arrive at (4.18).

We now assume that  $b_j \neq 0$  for all  $j \geq 1$  and prove that  $\sigma_f > 0$ , provided that  $f \neq 0$ . Suppose, by contradiction, that  $\sigma_f = 0$ . We first show that

$$\mathbb{P}_\mu\{M_k = 0 \text{ for all integers } k \geq 0\} = 1. \quad (4.19)$$

Indeed, it follows from relation (4.3), which is true with  $\mathbb{P}_\mu$ -probability 1, that

$$\int_0^k f(u_s) ds = M_k - g(u_k) + g(u_0), \quad k \geq 0. \quad (4.20)$$

In view of the Markov property, we have

$$\begin{aligned} \mathbb{E}_\mu M_k^2 &\leq 3 \mathbb{E}_\mu \left( \int_0^k f(u_s) ds \right)^2 + 3 \mathbb{E}_\mu |g(u_k)|^2 + 3 \mathbb{E}_\mu |g(u_0)|^2 \\ &= 3 \mathbb{E}_\mu \left( \int_0^k f(u_s) ds \right)^2 + 3 \int_H (\mathfrak{P}_k g^2(u) + g^2(u)) \mu(du). \end{aligned}$$

Combining this with Corollary 3.5.3 and Lemma 4.1.2, we see that

$$\lim_{k \rightarrow \infty} k^{-1} \mathbb{E}_\mu M_k^2 \leq 3 \mathbb{E}_\mu \left( \int_0^k f(u_s) ds \right)^2 = 0. \quad (4.21)$$

On the other hand, since  $M_k$  is a martingale with stationary differences  $X_k$ , we have

$$\mathbb{E}_\mu M_k^2 = \mathbb{E}_\mu \left( \sum_{l=1}^k X_l \right)^2 = \mathbb{E}_\mu \sum_{l=1}^k X_l^2 = k \mathbb{E}_\mu X_1^2.$$

Substituting this into (4.21), we conclude that  $\mathbb{E}_\mu X_k^2 = 0$  for any  $k \geq 0$ . It follows that  $M_k = 0$  almost surely for any  $k \geq 0$ , and therefore (4.19) holds.

We now prove that relation (4.20) with  $M_k = 0$  cannot hold for all  $k \geq 0$ , unless  $f \equiv 0$ . Indeed, suppose there is a non-degenerate ball  $B \subset H$  and a constant  $\varepsilon > 0$  such that  $f(u) \geq \varepsilon$  for  $u \in B$ . Using the fact that  $b_j \neq 0$  for all  $j \geq 1$ , it is not difficult to show that (cf. proof of Lemma 3.3.11)

$$\mathbb{P}_\mu\{u_t \in B \text{ for } 0 \leq t \leq k\} > 0 \quad \text{for any } k \geq 1.$$

It follows that

$$\mathbb{P}_\mu \left\{ \int_0^k f(u_s) ds \geq k\varepsilon \right\} > 0 \quad \text{for any } k \geq 1. \quad (4.22)$$

On the other hand, since  $g$  is bounded on bounded subsets of  $H$ , we have  $g(v) \leq C$  for  $v \in B$ . This implies that if  $u_0, u_k \in B$ , then  $|g(u_k) - g(u_0)| \leq 2C$ . Choosing  $k \geq 1$  so large that  $k\varepsilon > 2C$  and recalling (4.22), we see that (4.20) cannot hold with probability 1. The contradiction obtained completes the proof of the proposition.  $\square$

**Theorem 4.1.5.** *Under the hypotheses of Theorem 3.5.2, assume that<sup>2</sup>  $\sigma_f > 0$ . Then for any  $\gamma \in (0, 1]$ , sufficiently small  $\varkappa > 0$ , and any function  $f \in C^\gamma(H, w_\varkappa)$  such that  $(f, \mu) = 0$  we have*

$$\mathbb{P}_\mu \left\{ \limsup_{t \rightarrow \infty} \frac{1}{\omega(t)} \int_0^t f(u_s) ds = \sigma_f \right\} = 1, \quad (4.23)$$

$$\mathbb{P}_\mu \left\{ \liminf_{t \rightarrow \infty} \frac{1}{\omega(t)} \int_0^t f(u_s) ds = -\sigma_f \right\} = 1, \quad (4.24)$$

where  $\omega(t) = (2t \ln \ln t)^{1/2}$ .

*Proof.* In view of (4.5), it suffices to establish the LIL for  $M_k$ . To this end, we shall apply Theorem 7.12.2. Let us define the conditional variance  $V_k^2$  by relation (7.58) in which  $\mathbb{E}$  is replaced by  $\mathbb{E}_\mu$ . The required result will be established if we show that

$$\mathbb{P}_\mu \{k^{-1}V_k^2 \rightarrow \sigma_f^2 \text{ as } k \rightarrow \infty\} = 1. \quad (4.25)$$

Repeating the calculations used in Step 2 of the proof of Theorem 4.1.1, we see that

$$k^{-1}V_k^2 = k^{-1} \sum_{l=1}^k \varphi(u_{l-1}),$$

where  $\varphi(v) = \mathbb{E}_v M_1^2$ . As was mentioned in the proof of Theorem 4.1.1 (see inequality (4.15)), the function  $\varphi$  belongs to  $C^\alpha(H, w_\rho)$ , where  $\rho \rightarrow 0$  as  $\varkappa \rightarrow 0$ . Therefore, by Exercise 4.1.3 with  $\lambda = \mu$  and  $f = \varphi$ , the required relation (4.25) will be established if we show that  $(\varphi, \mu) = \sigma_f^2$ .

In view of relation (4.14) and the formulas for the terms in its right-hand side, we have

$$\begin{aligned} (\varphi, \mu) &= (\mathfrak{P}_1 g^2, \mu) + (g^2, \mu) + 2 \int_0^1 \int_0^t (\mathfrak{P}_s (f \mathfrak{P}_{t-s} f), \mu) ds dt \\ &\quad - 2(g \mathfrak{P}_1 g, \mu) + 2 \int_0^1 (\mathfrak{P}_s (f \mathfrak{P}_{1-s} g), \mu) ds - 2 \int_0^1 (g \mathfrak{P}_s f, \mu) ds. \end{aligned} \quad (4.26)$$

Let us denote by  $A_m$  the  $m^{\text{th}}$  term on the right-hand side of this relation. Then, using the stationarity of  $\mu$ , we derive

$$A_1 = (g^2, \mu), \quad A_3 = 2 \int_0^1 (1-s)(f \mathfrak{P}_s f, \mu) ds, \quad A_5 = 2 \int_0^1 (f \mathfrak{P}_s g, \mu) ds.$$

Substituting these expressions into (4.26) and using the simple relation

$$\mathfrak{P}_s g = \int_0^\infty \mathfrak{P}_{s+t} f dt = g - \int_0^s \mathfrak{P}_t f dt,$$

we obtain the required result. The proof is complete.  $\square$

<sup>2</sup>We recall that, by Proposition 4.1.4,  $\sigma_f > 0$  if  $b_j \neq 0$  for all  $j \geq 1$ .



*Exercise 4.1.6.* Prove that the assertions of Theorem 4.1.5 remain valid for all solutions. *Hint:* There are two possibilities: (a) one can apply Theorem 4.7 of [HH80] to prove that the martingale  $M_k$  satisfies the LIL; (b) using the existence of a mixing extension, one can couple a given solution with a stationary one. See the paper [Kuk02a] in which the latter idea is applied to prove a CLT.

*Exercise 4.1.7.* Under the hypotheses of Theorem 3.5.2, assume that  $\sigma_f = 0$ . Show that, for any  $\gamma \in (0, 1]$ , sufficiently small  $\varkappa > 0$ , and any  $f \in C^\gamma(H, w_\varkappa)$  such that  $(f, \mu) = 0$ , we have

$$\mathbb{P}_\mu \left\{ \lim_{t \rightarrow \infty} \frac{1}{\sqrt{t}} \int_0^t f(u_s) ds = 0 \right\} = 1. \quad (4.27)$$

*Hint:* Use (4.19) and (4.5).

### 4.1.3 Central limit theorem

Let us denote by  $N(0, \sigma)$  the centred normal law with a variance  $\sigma^2 \geq 0$  and by  $\Phi_\sigma$  its distribution function (see Section 7.4). Recall that the constant  $\sigma_f \geq 0$  is defined by (4.18). The following theorem establishes the CLT for solutions of (3.84), (2.66).

**Theorem 4.1.8.** *Under the hypotheses of Theorem 3.5.2, the following convergence holds for an arbitrary  $\gamma \in (0, 1]$ , sufficiently small  $\varkappa > 0$ , any function  $f \in C^\gamma(H, w_\varkappa)$  with  $(f, \mu) = 0$ , and any  $v \in H$ :*

$$\mathcal{D}_v \left( \frac{1}{\sqrt{t}} \int_0^t f(u_s) ds \right) \rightarrow N(0, \sigma_f) \quad \text{as } t \rightarrow \infty, \quad (4.28)$$

where  $\mathcal{D}_v$  denote the distribution of a random variable under the law  $\mathbb{P}_v$ .

Note that, in view of Lemma 1.2.16, the weak convergence (4.28) is equivalent to

$$\mathbb{P}_v \left\{ \frac{1}{\sqrt{t}} \int_0^t f(u_s) ds \leq x \right\} \rightarrow \Phi_{\sigma_f}(x) \quad \text{as } t \rightarrow \infty \text{ for any } x \in \mathbb{R}. \quad (4.29)$$

A more detailed analysis enables one to prove that convergence (4.29) holds with the rate  $t^{-1/4-\varepsilon}$  for any  $\varepsilon > 0$ ; see [Shi06c]. This rate is the best possible in the case of martingales (see [Bol82]), but it is not known if it is optimal in our context.

*Proof of Theorem 4.1.8.* We confine ourselves to the more complicated case  $\sigma_f > 0$ . The main idea of the proof is the same as in the previous two subsections: we use the martingale approximation to reduce the problem to a similar question for martingales and apply a general result on CLT for martingales.

Let us note that, for any  $v \in H$ , we have the relation (cf. (4.5))

$$\lim_{t \rightarrow \infty} \mathbb{E}_v \left| t^{-1/2} \int_0^t f(u_s) ds - \hat{t}^{-1/2} M_{\hat{t}} \right|^2 = 0,$$

which follows from (4.3) and the inequality

$$\mathbb{E}_v |d_t - g(u_t) + g(u_0)|^2 \leq C_1 \quad \text{for all } t \geq 0,$$

with a constant  $C_1 > 0$  depending only on  $|v|_2$ . Thus, it suffices to prove that

$$\mathcal{D}_v(k^{-1/2}M_k) \rightarrow N(0, \sigma_f) \quad \text{as } k \rightarrow \infty.$$

By Theorem 7.12.3, this convergence will be established if we check that the martingale  $M_k$  satisfies Lindeberg's condition (7.64). To this end, it suffices to show that the martingale differences  $X_k$  satisfy the inequality

$$\mathbb{E}_v X_k^4 \leq C \exp(c|v|_2^2) \quad \text{for all } k \geq 0; \quad (4.30)$$

see the remark preceding Theorem 7.12.3. To prove (4.30), using the Markov property, we write (cf. (4.12))

$$\mathbb{E}_v X_k^4 = \mathbb{E}_v \mathbb{E}_v (X_k^4 | \mathcal{F}_{k-1}) = \mathbb{E}_v \mathbb{E}_{u_{k-1}} X_1^4 = \mathbb{E}_v \psi(u_{k-1}),$$

where we set  $\psi(v) = \mathbb{E}_v X_1^4$ . Since

$$\psi(v) \leq C_3 \left( \mathbb{E}_v g^4(u_1) + g(v) + \int_0^1 \mathbb{E}_v f^4(u_s) ds \right),$$

$f \in C^\gamma(H, w_\varkappa)$ , and  $g \in C^\alpha(H, w_\rho)$ , where  $\rho \rightarrow 0$  as  $\varkappa \rightarrow 0$ , we see from Lemmas 3.5.6 and 4.1.2 that  $\psi(v) \leq C_4 \exp(c|v|_2^2)$  for  $v \in H$  and some small  $c > 0$ . Applying again Lemma 3.5.6, we arrive at the required estimate (4.30). This completes the proof of Theorem 4.1.8.  $\square$

*Exercise 4.1.9.* Prove analogues of Theorems 4.1.1, 4.1.5, and 4.1.8 for the Navier–Stokes system perturbed by a kick force of the form (2.66). *Hint:* The martingale approximation described in Section 7.13 remains true for discrete-time RDS, and the same schemes apply. Note that, in this case the constant  $\sigma_f$  entering the LIL and CLT is defined by the relation  $\sigma_f^2 = 2(fg, \mu) - (f^2, \mu)$ , where  $g(u) = \sum_{k=0}^{\infty} \mathfrak{F}_k f(u)$ .

*Exercise 4.1.10.* Prove the SLLN, LIL, and CLT for Hölder continuous functionals defined on higher Sobolev spaces. *Hint:* Combine Theorem 3.5.5 with scheme used above.

## 4.2 Random attractors and stationary distributions

In this section, we study the relationship between stationary distributions for Markov RDS and random attractors. We first establish a sufficient condition ensuring the existence of a minimal random point attractor in the sense of almost sure convergence. We next prove the Ledrappier–Le Jan–Crauel theorem which

shows that there is a one-to-one correspondence between stationary distributions and Markov invariant measures. Finally, we establish the main result of this section claiming that the support of the disintegration of a Markov invariant measure is a minimal random point attractor in the sense of convergence in probability. We shall confine ourselves to the case of continuous-time RDS, which is technically more complicated.

### 4.2.1 Random point attractors

Let  $X$  be a Polish space with a metric  $\text{dist}_X$ . For any subset  $K \subset H$ , we define the function

$$\text{dist}_X(u, K) = \inf_{v \in K} \text{dist}_X(u, v),$$

where the infimum over an empty set is equal to  $+\infty$ . Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space and let  $\mathbf{K} = \{K_\omega, \omega \in \Omega\}$  be a family of closed subsets in  $X$ .

**Definition 4.2.1.** The family  $\mathbf{K}$  is said to be *measurable* if for any  $u \in X$  the real-valued function  $\omega \mapsto \text{dist}_X(u, K_\omega)$  is  $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$ -measurable. In this case,  $\mathbf{K}$  will be called a *closed random set*. If, in addition,  $K_\omega$  is compact for almost every  $\omega$ , then  $\mathbf{K}$  is called a *compact random set*.

*Exercise 4.2.2.* Let  $\mathbf{K}$  be a closed random set in  $X$ . Prove the following properties.

- (i) If  $f : \Omega \rightarrow X$  is an  $(\mathcal{F}, \mathcal{B}(X))$ -measurable function, then the function  $\omega \mapsto \text{dist}_X(f(\omega), K_\omega)$  is measurable.
- (ii) The function  $(\omega, u) \mapsto \text{dist}_X(u, K_\omega)$  is  $(\mathcal{F} \otimes \mathcal{B}(X), \mathcal{B}(\mathbb{R}))$ -measurable.

*Hint:* Use approximation by functions taking at most countably many values.

Let us assume that a group of measure-preserving transformations  $\boldsymbol{\theta} = \{\theta_t, t \in \mathbb{R}\}$  is defined on  $\Omega$  and consider an RDS  $\boldsymbol{\Phi} = \{\varphi_t^\omega, t \in \mathbb{R}_+\}$  over  $\boldsymbol{\theta}$  (see Section 1.3.2). We shall always assume that the trajectories  $\varphi_t^\omega(u)$  are continuous in time for any  $u \in X$  and  $\omega \in \Omega$ .

**Definition 4.2.3.** A compact random set  $\mathbf{A} = \{A_\omega, \omega \in \Omega\}$  in  $H$  is called a *random point attractor for  $\boldsymbol{\Phi}$  in the sense of almost sure convergence*<sup>3</sup> if it satisfies the following properties:

- (i) *Invariance.* For almost every  $\omega \in \Omega$ , we have

$$\varphi_t^\omega A_\omega = A_{\theta_t \omega}, \quad t \geq 0. \quad (4.31)$$

- (ii) *Attraction.* For any  $u \in X$ , we have

$$\mathbb{P}\{\text{dist}_X(\varphi_t^{\theta^{-t}\omega}(u), A_\omega) \rightarrow 0 \text{ as } t \rightarrow \infty\} = 1. \quad (4.32)$$

<sup>3</sup>In what follows, if there is no confusion, we shall simply say a *random attractor*.

**Definition 4.2.4.** A random point attractor  $\mathbf{A}$  is said to be minimal if for any other random point attractor  $\mathbf{A}' = \{A'_\omega, \omega \in \Omega\}$  we have

$$A_\omega \subset A'_\omega \quad \text{for almost every } \omega \in \Omega.$$

The following theorem provides a sufficient condition for the existence of a minimal random point attractor in the sense of almost sure convergence.

**Theorem 4.2.5.** Let  $\Phi = \{\varphi_t, t \geq 0\}$  be an RDS over  $\theta$  in a Polish space  $X$ . Suppose that there is a compact random set  $\mathbf{K} = \{K_\omega\}$  and a subset  $\Omega_* \in \mathcal{F}$  of full measure such that  $\theta_t(\Omega_*) = \Omega_*$  for all  $t \in \mathbb{R}$  and

$$\text{dist}_X(\varphi_t^{\theta-t\omega}(u), K_\omega) \rightarrow 0 \quad \text{as } t \rightarrow \infty \text{ for any } \omega \in \Omega_*, u \in X. \quad (4.33)$$

Then  $\Phi$  possesses a minimal random attractor.

*Proof. Step 1.* For any  $u \in X$  and  $\omega \in \Omega_*$ , we set

$$A_\omega(u, n) = \overline{\bigcup_{t \geq n} \varphi_t^{\theta-t\omega}(u)}, \quad (4.34)$$

where  $\overline{C}$  denotes the closure of  $C$  in  $X$ . Note that  $A_\omega(u, n)$  is the union of the continuous curve  $\{\varphi_t^{\theta-t\omega}(u), t \geq n\}$  and of the set of its limit points. The latter is a closed set that belongs to  $K_\omega$  in view of (4.33). The sequence  $\{A_\omega(u, n)\}$  is a nested family of compact sets, and therefore

$$A_\omega(u) = \bigcap_{n \geq 0} A_\omega(u, n)$$

is a non-empty compact subset of  $X$ . Furthermore,

$$A_\omega(u) \text{ is the set of all limiting points for the curve } \{\varphi_t^{\theta-t\omega}(u), t \geq 0\}. \quad (4.35)$$

Thus, by (4.33),

$$A_\omega(u) \subset K_\omega \quad \text{for all } u \in X, \omega \in \Omega_*. \quad (4.36)$$

We now define a family of subsets  $\mathbf{A} = \{A_\omega, \omega \in \Omega\}$  by the rule

$$A_\omega = \begin{cases} \overline{\bigcup_{u \in X} A_\omega(u)} & \text{for } \omega \in \Omega_*, \\ \emptyset & \text{for } \omega \in \Omega \setminus \Omega_*. \end{cases} \quad (4.37)$$

Due to (4.36),  $A_\omega$  is a compact set for each  $\omega \in \Omega$ . We shall show that  $\mathbf{A}$  is a minimal random point attractor.

*Step 2.* Let us show that  $\mathbf{A}$  is a compact random set. To this end, we need to show that the function  $\omega \mapsto \text{dist}_X(v, A_\omega)$  is  $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$ -measurable for any  $v \in X$ .

Using the compactness of  $A_\omega(u, n)$ , it is easy to show that

$$\text{dist}_X(v, A_\omega(u)) = \lim_{n \rightarrow \infty} \text{dist}_X(v, A_\omega(u, n)) = \liminf_{k \rightarrow \infty} d_X(v, \varphi_t^{\theta^{-t}\omega} u)$$

for any  $u, v \in X$  and  $\omega \in \Omega_*$ . Since the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  is complete and  $\mathbb{P}(\Omega_*) = 1$ , we conclude that the function

$$d : X \times X \times \Omega \rightarrow \mathbb{R}, \quad d(u, v, \omega) = \text{dist}_X(v, A_\omega(u)),$$

is measurable. By Corollary 7.3.2 of the projection theorem (see Section 7.3), for any  $v \in X$  the function

$$\omega \mapsto \inf_{u \in X} d(u, v, \omega) = \text{dist}_X(v, A_\omega)$$

is universally measurable. Using again that  $(\Omega, \mathcal{F}, \mathbb{P})$  is complete, we arrive at the required result.

*Step 3.* Let us prove that  $\mathbf{A}$  satisfies the invariance property. In what follows, we shall sometimes write  $\varphi_t(\omega)$  instead of  $\varphi_t^\omega$  to avoid complicated expressions in superscripts.

The invariance property is a consequence of the relation

$$\varphi_t^\omega A_\omega(u) = A_{\theta_t \omega}(u) \quad \text{for any } u \in X, \omega \in \Omega_*. \quad (4.38)$$

Indeed, if (4.38) is proved, then  $\varphi_t^\omega A_\omega(u) \subset A_{\theta_t \omega}(u)$ , whence it follows that  $\varphi_t^\omega A_\omega \subset A_{\theta_t \omega}$ . To establish the converse inclusion, note that  $\varphi_t^\omega A_\omega$  is a compact set containing  $A_{\theta_t \omega}(u)$  for any  $u \in X$ . Therefore, it must contain the closure of  $\cup_u A_{\theta_t \omega}(u)$ , which coincides with  $A_{\theta_t \omega}$ .

To prove (4.38), note that, in view of (4.35), a point  $v \in X$  belongs to  $A_\omega(u)$  if and only if there is a sequence  $t_j \rightarrow \infty$  such that

$$\varphi_{t_j}(\theta_{-t_j} \omega) u \rightarrow v \quad \text{as } j \rightarrow \infty.$$

Applying the continuous operator  $\varphi_t^\omega$  and using the cocycle property, we see that

$$\varphi_{t+t_j}(\theta_{-(t+t_j)}(\theta_t \omega)) u \rightarrow \varphi_t^\omega v \quad \text{as } j \rightarrow \infty.$$

This implies that  $\varphi_t^\omega v \in A_{\theta_t \omega}(u)$  for any  $v \in A_\omega(u)$ , that is,  $\varphi_t^\omega A_\omega(u) \subset A_{\theta_t \omega}(u)$ . Conversely, suppose that  $w \in A_{\theta_t \omega}(u)$  and choose a sequence  $t_j \rightarrow \infty$  such that

$$\varphi_{t_j}(\theta_{-t_j}(\theta_t \omega)) u \rightarrow w \quad \text{as } j \rightarrow \infty. \quad (4.39)$$

Now note that

$$\varphi_{t_j}(\theta_{-t_j}(\theta_t \omega)) = \varphi_t^\omega \circ \varphi_{t_j-k}(\theta_{-(t_j-k)} \omega) \quad \text{for } t_j \geq k. \quad (4.40)$$

It follows from (4.35) that the sequence  $\varphi_{t_j-k}(\theta_{-(t_j-k)} \omega) u$  has a limit point  $v$ , which belongs to  $A_\omega(u)$ . Since  $\varphi_t^\omega$  is continuous, we conclude from (4.40) that

$$\varphi_{t_j}(\theta_{-t_j}(\theta_t \omega)) u \rightarrow \varphi_t^\omega v \quad \text{as } j \rightarrow \infty.$$

Comparing this with (4.39), we obtain  $\varphi_t^\omega v = w$ . Since  $w \in A_{\theta_t \omega}$  was arbitrary, we have proved that  $\varphi_t^\omega A_\omega(u) \supset A_{\theta_t \omega}(u)$ .

*Step 4.* We now prove that  $\mathbf{A}$  possesses the attraction property. Let  $\omega \in \Omega_*$  and let  $u \in X$ . If  $\varphi_t^{\theta_{-t}\omega} u$  does not tend to  $A_\omega$ , then there is a constant  $\varepsilon > 0$  and a sequence  $t_j \rightarrow \infty$  such that

$$\text{dist}_X(\varphi_{t_j}(\theta_{-t_j}\omega)u, A_\omega) \geq \varepsilon \quad \text{for all } j \geq 1. \quad (4.41)$$

In view of (4.33) and the compactness of  $K_\omega$ , there is no loss of generality in assuming that the sequence  $\{\varphi_{t_j}(\theta_{-t_j}\omega)u\}$  converges to a point in  $X$ . This point belongs to  $A_\omega$ , which contradicts (4.41).

*Step 5.* It remains to show that  $\mathbf{A}$  is a minimal random attractor. We claim that if  $\mathbf{A}' = \{A'_\omega\}$  is another random attractor and  $\Omega'_*$  is the set of full measure on which the attraction property holds for  $\mathbf{A}'$ , then

$$A_\omega \subset A'_\omega \quad \text{for } \omega \in \Omega_* \cap \Omega'_*. \quad (4.42)$$

Indeed, for any  $\omega \in \Omega_* \cap \Omega'_*$  and  $u \in X$  the set  $A'_\omega$  must contain all the limit points of the curve  $\{\varphi_t(\theta_{-t}\omega)u\}$ . It follows that

$$A_\omega(u) \subset A'_\omega \quad \text{for any } u \in X.$$

Since  $A'_\omega$  is closed, we conclude that (4.42) holds. To complete the proof of Theorem 4.34, it remains to note that  $\Omega_* \cap \Omega'_*$  is a set of full measure.  $\square$

In what follows, we shall need another concept of random attractors for which the attraction property holds in a weaker sense.

**Definition 4.2.6.** A compact random set  $\mathbf{A} = \{A_\omega, \omega \in \Omega\}$  in  $X$  is called a *random point attractor for  $\Phi$  in the sense of convergence in probability* (or a *weak random attractor*) if it possesses the following properties (cf. Definition 4.2.3):

- (i) *Invariance.* For any  $t \geq 0$  and almost every  $\omega \in \Omega$ , we have  $\varphi_t^\omega A_\omega = A_{\theta_t \omega}$ .
- (ii) *Attraction.* For any  $u \in X$ , the function  $\omega \mapsto \text{dist}_X(\varphi_t^{\theta_{-t}\omega} u, A_\omega)$  converges to zero in probability as  $k \rightarrow +\infty$ , that is, for any  $\varepsilon > 0$ , we have

$$\mathbb{P}\{\text{dist}_X(\varphi_t(\theta_{-t}\omega)u, A_\omega) \geq \varepsilon\} \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (4.43)$$

Note that the convergence “from the past” (4.43) is equivalent to the convergence “in the future”

$$\mathbb{P}\{\text{dist}_X(\varphi_t^\omega u, A_{\theta_t \omega}) \geq \varepsilon\} \rightarrow 0 \quad \text{as } t \rightarrow \infty, \quad (4.44)$$

because  $\theta_t$  preserves  $\mathbb{P}$ , and for discrete-time RDS, the invariance property of Definitions 4.2.6 and 4.2.3 are equivalent. Since almost sure convergence implies convergence in probability, any random attractor (in the sense of Definition 4.2.3) is also a weak random attractor. However, it does not need to be minimal (cf. Definition 4.2.4).

A natural question is whether (weak) random attractors contain “all relevant information” on the system as time goes to  $+\infty$ . In particular, do they support disintegrations of invariant measures? In general, the answer is negative (see [Bax89, Cra01]). There is, however, an important class of invariant measures for which this property is true. They are called *Markov invariant measures* and are studied in the next section.

### 4.2.2 Ledrappier–Le Jan–Crauel theorem

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space, let  $X$  be a Polish space, and let  $\Phi = \{\varphi_t^\omega, t \geq 0\}$  be a Markov RDS over a measure-preserving group of transformations  $\theta = \{\theta_t, t \in \mathbb{R}\}$ . Recall that the concepts of an invariant measure for  $\Phi$  and of its disintegration, as well as the class  $\mathcal{P}(\Omega \times X, \mathbb{P})$ , were introduced in Section 1.3.4.

**Definition 4.2.7.** An invariant measure  $\mathfrak{M} \in \mathcal{P}(\Omega \times X, \mathbb{P})$  for  $\Phi$  is said to be *Markov* if its disintegration  $\{\mu_\omega\}$  is  $\mathcal{F}_0$ -measurable, that is, for any  $\Gamma \in \mathcal{B}(X)$  the function  $\omega \mapsto \mu_\omega(\Gamma)$  is  $(\mathcal{F}_0, \mathcal{B}(\mathbb{R}))$ -measurable.

The following proposition shows that the time  $t = 0$  does not play any particular role for a Markov invariant measure.

**Proposition 4.2.8.** *Let  $\mathfrak{M} \in \mathcal{P}(\Omega \times H, \mathbb{P})$  be an invariant measure for a Markov RDS  $\Phi$ . Then following properties are equivalent:*

- (a)  $\mathfrak{M}$  is Markov.
- (b)  $\mu_\omega$  is  $\mathcal{F}_t$ -measurable for some  $t \in \mathbb{R}$ .
- (c)  $\mu_\omega$  is  $\mathcal{F}_t$ -measurable for any  $t \in \mathbb{R}$ .

*Proof.* We shall only show that (b) implies (a), because the proof of the implication (a)  $\Rightarrow$  (c) is similar, and (c) trivially implies (b).

Suppose that  $\mu_\omega$  is  $\mathcal{F}_t$ -measurable for some  $t > 0$ . Then the random measure  $\mu_{\theta_{-t}\omega}$  is  $\mathcal{F}_0$ -measurable and the mapping  $\omega \mapsto \varphi_t^{\theta_{-t}\omega} u$  acting from  $\Omega$  to  $X$  is  $(\mathcal{F}_0, \mathcal{B}(X))$ -measurable. We claim that the random measure  $(\varphi_t^{\theta_{-t}\omega})_* \mu_{\theta_{-t}\omega}$  is  $\mathcal{F}_0$ -measurable. If this is proved, then the  $\mathcal{F}_0$ -measurability of  $\mu_\omega$  follows from relation (1.72) and the fact that  $\mathcal{F}_0$  contains all subsets of  $\mathcal{F}$  of zero measure.

We need to show that the real-valued function  $f(\omega) = \mu_{\theta_{-t}\omega}((\varphi_t^{\theta_{-t}\omega})^{-1}(\Gamma))$  is  $\mathcal{F}_0$ -measurable for any  $\Gamma \in \mathcal{B}(X)$ . We have

$$f(\omega) = \int_X h(\omega, u) \mu_{\theta_{-t}\omega}(du), \quad (4.45)$$

where  $h(\omega, u) = \mathbb{I}_\Gamma(\varphi_t^{\theta_{-t}\omega} u)$ . The function  $h$  is  $(\mathcal{F}_0 \otimes \mathcal{B}(X), \mathcal{B}(\mathbb{R}))$ -measurable. If it was the product of two functions  $h_1(\omega)$  and  $h_2(u)$ , then the integral (4.45) could be written in the form

$$h_1(\omega) \int_X h_2(u) \mu_{\theta_{-t}\omega}(du),$$

and the required measurability would follow immediately. The  $\mathcal{F}_0$ -measurability of (4.45) in the general case can be obtained with the help of the monotone class technique. We leave the details to the reader as an exercise.  $\square$

We now turn to the main result of this section, which shows that there is a one-to-one correspondence between the Markov invariant measures and the stationary distributions for  $\Phi$ .

**Theorem 4.2.9.** *Let  $\Phi = \{\varphi_t^\omega, t \geq 0\}$  be a Markov RDS in a Polish space  $X$ . Then the following assertions hold.*

- (i) *Let  $\mathfrak{M} \in \mathcal{P}(\Omega \times X, \mathbb{P})$  be a Markov invariant measure and let  $\{\mu_\omega\}$  be its disintegration. Then  $\mathbb{E}\mu_\cdot$  is a stationary measure for  $\Phi$ . Moreover, if  $\mathfrak{M}' \in \mathcal{P}(\Omega \times X, \mathbb{P})$  is another Markov invariant measure with disintegration  $\{\mu'_\omega\}$  such that  $\mathbb{E}\mu' = \mathbb{E}\mu_\cdot$ , then  $\mathfrak{M}' = \mathfrak{M}$ .*
- (ii) *Let  $\mu \in \mathcal{P}(X)$  be a stationary measure for  $\Phi$ . Then for any sequence  $t_k \rightarrow +\infty$  there is a set  $\tilde{\Omega} \in \mathcal{F}$  of full measure such that there is a weak\* limit*

$$\mu_\omega = \lim_{k \rightarrow \infty} \varphi_{t_k}(\theta_{-t_k}\omega)_*(\mu) \quad (4.46)$$

for any  $\omega \in \tilde{\Omega}$ . Moreover, if  $\{t'_k\}$  is another sequence going to  $+\infty$ , then the corresponding family of measures  $\{\mu'_\omega\}$  coincides with  $\{\mu_\omega\}$  almost surely. Finally, the measure  $\mathfrak{M} \in \mathcal{P}(\Omega \times X, \mathbb{P})$  defined by its disintegration<sup>4</sup>  $\{\mu_\omega\}$  is a Markov invariant measure for  $\Phi$ , and  $\mathbb{E}\mu_\cdot = \mu$ .

Thus, there is a one-to-one correspondence between the stationary measures and Markov invariant measures for  $\Phi$ .

*Proof.* (i) In view of Proposition 1.3.27, we have

$$(f(\varphi_t u), \mu_\omega) = (f, \mu_{\theta_t \omega}) \quad \mathbb{P}\text{-almost surely,} \quad (4.47)$$

where  $f : X \rightarrow \mathbb{R}$  is an arbitrary bounded measurable function. We wish to take the mean value. Using the fact that  $\mu_\omega$  is  $\mathcal{F}_0$ -measurable and  $\varphi_t u$  is  $\mathcal{F}_t^+$ -measurable and recalling that  $\mathcal{F}_0$  and  $\mathcal{F}_t^+$  are independent, we can write

$$\mathbb{E}(f(\varphi_t u), \mu_\omega) = (\mathbb{E}f(\varphi_t u), \mathbb{E}\mu_\omega) = (\mathfrak{P}_t f, \mu), \quad (4.48)$$

where we set  $\mu = \mathbb{E}\mu_\cdot$ . On the other hand, since  $\mathbb{P}$  is invariant with respect to  $\theta$ , we have

$$\mathbb{E}(f, \mu_{\theta_t \omega}) = (f, \mu). \quad (4.49)$$

Combining (4.47) – (4.49), we obtain

$$(f, \mathfrak{P}_t^* \mu) = (f, \mu) \quad \text{for any bounded measurable function } f,$$

whence it follows that  $\mu$  is a stationary measure for  $\Phi$ .

<sup>4</sup>We set  $\mu_\omega = \gamma$  for  $\omega \in \Omega \setminus \tilde{\Omega}$ , where  $\gamma \in \mathcal{P}(X)$  is an arbitrary measure.



*Exercise 4.2.10.* Justify the first equality in (4.48) and relation (4.49). *Hint:* Use the monotone class technique.

We now prove the second part of (i). Let us fix any increasing sequence  $\{t_k\}$  going to  $+\infty$ . In view of Proposition 1.3.27, with probability 1 we have

$$\varphi_{t_k}(\theta_{-t_k}\omega)\mu_{\theta_{-t_k}\omega} = \mu_\omega \quad \text{for all } k \geq 1. \quad (4.50)$$

It follows that

$$\int_X f(\varphi_{t_k}(\theta_{-t_k}\omega)u)\mu_{\theta_{-t_k}\omega}(du) = \int_X f(u)\mu_\omega(du) \quad \text{for any } k \geq 1, f \in C_b(X). \quad (4.51)$$

Let us set  $\mathcal{G}_k = \mathcal{F}_{[-t_k, 0]}$  for  $k \geq 1$  and note that  $f(\varphi_{t_k}(\theta_{-t_k}\omega)u)$  is  $\mathcal{G}_k$ -measurable, while  $\mu_{\theta_{-t_k}\omega}$  is independent of  $\mathcal{G}_k$ . Therefore, taking the conditional expectation of both sides of (4.51) given  $\mathcal{G}_k$ , with probability 1 we derive

$$\int_X f(\varphi_{t_k}(\theta_{-t_k}\omega)u)\mu(du) = \mathbb{E}((f, \mu_\omega) | \mathcal{G}_k). \quad (4.52)$$

Since  $\mu_\omega$  is  $\mathcal{F}_0$ -measurable and  $\mathcal{F}_0 = \sigma\{\mathcal{G}_k, k \geq 1\}$ , we see that the right-hand side of (4.52) is a right-closable martingale with respect to the filtration  $\{\mathcal{G}_k\}$ . Hence, by Doob's theorem on convergence of right-closed martingale sequences, we have

$$\lim_{k \rightarrow \infty} (f, \varphi_{t_k}(\theta_{-t_k}\omega)\mu) = \lim_{k \rightarrow \infty} \int_X f(\varphi_{t_k}(\theta_{-t_k}\omega)u)\mu(du) = (f, \mu_\omega)$$

for any  $f \in C_b(X)$  and almost every  $\omega \in \Omega$ . Theorem 7.5.2 now implies that there is a set of full measure  $\tilde{\Omega} \in \mathcal{F}$  such that (4.46) holds, where the convergence of measures is understood in the weak\* topology. It follows that  $\mu_\omega$  is uniquely defined by  $\mu$  on the set  $\tilde{\Omega}$ , and therefore  $\mathfrak{M}' = \mathfrak{M}$  for any Markov invariant measure  $\mathfrak{M}'$  for which the associated stationary measure coincides with  $\mu$ .

(ii) Let us take an arbitrary sequence  $\{t_k\}$  going to  $+\infty$ . Without loss of generality, we can assume that  $\{t_k\}$  is increasing. Given a bounded continuous function  $f : X \rightarrow \mathbb{R}$ , consider the sequence  $\xi_k(\omega) = (f, \varphi_{t_k}(\theta_{-t_k}\omega)\mu)$ . Suppose we have shown that  $\{\xi_k\}$  is a martingale with respect to the filtration  $\mathcal{G}_k = \mathcal{F}_{[-t_k, 0]}$ . Since  $\{\xi_k\}$  is bounded uniformly in  $k$  and  $\omega$ , Doob's martingale convergence theorem implies that  $\xi_k(\omega)$  converges almost surely. In particular, for any  $f \in C_b(X)$  the sequence  $(f, \mu_\omega^k)$  converges for almost every  $\omega \in \Omega$ . Therefore, by Theorem 7.5.2, there is a random probability measure  $\{\mu_\omega\} \subset \mathcal{P}(X)$  and a set of full measure  $\tilde{\Omega} \in \mathcal{F}$  such that (4.46) holds.

We now prove that  $\{\xi_k\}$  is a martingale. Setting  $g_k(\omega, v) = f(\varphi_{t_k}(\theta_{-t_k}\omega)v)$ , we can write

$$\xi_k = (g_k(\omega, \cdot), \mu) = (g_k(\omega, v), \mu(dv)).$$

The cocycle and Markov properties imply that

$$\begin{aligned}
\mathbb{E} \{ \xi_{k+1} | \mathcal{G}_k \} &= \mathbb{E} \{ (g_{k+1}(\omega, \cdot), \mu) | \mathcal{G}_k \} \\
&= \mathbb{E} \{ (g_k(\omega, \varphi_{s_k}(\theta_{-t_{k+1}}\omega)v), \mu(dv)) | \mathcal{G}_k \} \\
&= \mathbb{E}_{\omega'} \{ (g_k(\omega, \varphi_{s_k}(\theta_{-t_{k+1}}\omega')v), \mu(dv)) \} \\
&= \mathbb{E}_{\omega'} \{ (g_k(\omega, \varphi_{s_k}(\omega')v), \mu(dv)) \} \\
&= (\mathfrak{P}_{s_k} g_k(\omega, \cdot), \mu) = (g_k(\omega, \cdot), \mathfrak{P}_{s_k}^* \mu) = \xi_k,
\end{aligned}$$

where  $s_k = t_{k+1} - t_k$ , and we used that  $\mu$  is a stationary measure for  $\Phi$ .

Let us show that the limit in (4.46) does not depend on  $\{t_k\}$ . Indeed, let  $\{t'_k\}$  be another sequence going to  $+\infty$  and let  $\mu'_\omega$  be the corresponding limit. Let us consider the altered sequence  $\{s_k\} = \{t_1, t'_1, t_2, t'_2, \dots\}$ . Since the limit (4.46) with  $t_k = s_k$  also exists almost surely, we conclude that  $\mu_\omega = \mu'_\omega$  for a.a.  $\omega \in \Omega$ .

We now set

$$\mathfrak{M}(d\omega, du) = \mu_\omega(du) \mathbb{P}(d\omega) \quad (4.53)$$

and prove that is a Markov invariant measure for  $\Phi$ . Indeed, the definition of  $\mathfrak{M}$  implies that its disintegration is  $\mathcal{F}_0$ -measurable. To prove that  $\mathfrak{M}$  is invariant, it suffices to show that

$$\varphi_t^\omega \mu_\omega = \mu_{\theta_t \omega} \quad \text{for any } t > 0 \text{ and a.e. } \omega \in \Omega. \quad (4.54)$$

Let us set  $t_k = kt$  and denote by  $\tilde{\Omega}$  the set of convergence in (4.46). Choose any  $\omega \in \tilde{\Omega}$  such that  $\theta_t \omega \in \tilde{\Omega}$ . Then convergence (4.46) and the cocycle property imply that

$$\begin{aligned}
(f, \varphi_t^\omega \mu_\omega) &= \int_X f(\varphi_t^\omega u) \mu_\omega(du) = \lim_{k \rightarrow \infty} \int_X f(\varphi_t^\omega u) (\varphi_{t_k}(\theta_{-t_k} \omega) \mu)(du) \\
&= \lim_{k \rightarrow \infty} \int_X f(\varphi_t^\omega \circ \varphi_{t_k}(\theta_{-t_k} \omega) u) \mu(du) \\
&= \lim_{k \rightarrow \infty} \int_X f(\varphi_{t_{k+1}}(\theta_{-t_{k+1}}(\theta_t \omega)) u) \mu(du) = (f, \mu_{\theta_t \omega}).
\end{aligned}$$

Since  $f \in C_b(X)$  is arbitrary, it remains to note that the set  $\{\omega \in \tilde{\Omega} : \theta_t \omega \in \tilde{\Omega}\}$  is of full measure.

It remains to show that  $\mathbb{E} \mu_\cdot = \mu$ . To this end, take any function  $f \in C_b(X)$  and note that, in view of (4.46), we have

$$(f, \mu_\omega) = \lim_{k \rightarrow \infty} (f, \varphi_k(\theta_{-k} \omega) \mu) \quad \text{for almost every } \omega \in \Omega.$$

Taking the mean value of both sides and using the Lebesgue theorem on dominated convergence and the stationarity of  $\mu$ , we obtain

$$\begin{aligned}
\mathbb{E}(f, \mu_\cdot) &= \lim_{k \rightarrow \infty} \mathbb{E}(f, \varphi_k(\theta_{-k} \cdot) \mu) \\
&= \lim_{k \rightarrow \infty} \mathbb{E} \int_X f(\varphi_k u) \mu(du) = \lim_{k \rightarrow \infty} (\mathfrak{P}_k f, \mu) = (f, \mu).
\end{aligned}$$

The proof of Theorem 4.2.9 is complete.  $\square$

*Exercise 4.2.11.* In the setting of Theorem 4.2.9, prove that, for any  $\Gamma \in \mathcal{B}(X)$  and any sequence  $\{t_k\}$  going to  $+\infty$ , we have

$$(\varphi_{t_k}(\theta_{-t_k}\omega)\mu)(\Gamma) \rightarrow \mu_\omega(\Gamma) \quad \text{as } k \rightarrow \infty \text{ for almost every } \omega \in \Omega.$$

*Hint:* The fact that  $\{\xi_k\}$  is a martingale is true for any bounded measurable function  $f$ .

We now study the relationship between the support of the disintegration for a Markov invariant measure and weak random attractors.

**Proposition 4.2.12.** *Let  $\Phi$  be a Markov RDS, let  $\mathfrak{M} \in \mathcal{P}(\Omega \times \mathbb{P}, \mathbb{P})$  be a Markov invariant measure for  $\Phi$  with disintegration  $\{\mu_\omega\}$ , and let  $\mathbf{A} = \{A_\omega\}$  be a weak random attractor for  $\Phi$ . Then*

$$\text{supp } \mu_\omega \subset A_\omega \quad \text{for almost all } \omega \in \Omega. \quad (4.55)$$

*Proof. Step 1.* We first show that if  $f(\omega, u)$  is a real-valued bounded measurable function on  $\Omega \times H$  that is continuous in  $u$  and  $\mu \in \mathcal{P}(H)$  is the stationary measure associated with  $\mathfrak{M}$  (see Theorem 4.2.9), then

$$\iint_{\Omega \times H} f(\omega, \varphi_k^{\theta_{-k}\omega} u) \mu(du) \mathbb{P}(d\omega) \rightarrow \iint_{\Omega \times H} f(\omega, u) \mu_\omega(du) \mathbb{P}(d\omega) \quad (4.56)$$

as  $k \rightarrow \infty$ . Indeed, by (4.46), for almost every  $\omega \in \Omega$ , we have

$$\int_H f(\omega, \varphi_k^{\theta_{-k}\omega} u) \mu(du) = \int_H f(\omega, u) (\varphi_k^{\theta_{-k}\omega} \mu)(du) \rightarrow \int_H f(\omega, u) \mu_\omega(du)$$

as  $k \rightarrow \infty$ . Taking the mean value with respect to  $\omega$  and using the Lebesgue theorem on dominated convergence, we obtain (4.56).

*Step 2.* We now fix any integer  $n \geq 1$  and set

$$\begin{aligned} f_n(\omega, u) &= 1 - (n \text{dist}_X(u, A_\omega)) \wedge 1, \\ A_\omega^n &= \{u \in H : \text{dist}_X(u, A_\omega) \leq 1/n\}. \end{aligned}$$

Applying (4.56) to  $f_n$ , we obtain

$$\begin{aligned} \lim_{k \rightarrow \infty} \iint_{\Omega \times H} f_n(\omega, \varphi_k^{\theta_{-k}\omega} u) \mu(du) \mathbb{P}(d\omega) &= \iint_{\Omega \times H} f_n(\omega, u) \mu_\omega(du) \mathbb{P}(d\omega) \\ &\leq \int_\Omega \mu_\omega(A_\omega^n) \mathbb{P}(d\omega). \end{aligned} \quad (4.57)$$

On the other hand, it follows from (4.43) that

$$\lim_{k \rightarrow \infty} \int_\Omega f_n(\omega, \varphi_k^{\theta_{-k}\omega} u) \mathbb{P}(d\omega) = 1 \quad \text{for any } u \in H.$$

Integrating this relation with respect to  $\mu$  and using again the Lebesgue theorem, we derive

$$\lim_{k \rightarrow \infty} \iint_{\Omega \times H} f_n(\omega, \varphi_k^{\theta_{-k}\omega} u) \mu(du) \mathbb{P}(d\omega) = 1. \quad (4.58)$$

Comparing (4.57) and (4.58), we see that

$$\int_{\Omega} \mu_{\omega}(A_{\omega}^n) \mathbb{P}(d\omega) = 1.$$

It follows that

$$\mathbb{P}\{\omega \in \Omega : \mu_{\omega}(A_{\omega}^n) = 1\} = 1 \quad \text{for any integer } n \geq 1.$$

Since  $\bigcap_n A_{\omega}^n = A_{\omega}$ , we obtain the relation

$$\mathbb{P}\{\mu_{\omega}(A_{\omega}) = 1\} = 1,$$

which is equivalent to (4.55).  $\square$

### 4.2.3 Ergodic RDS and minimal attractors

We now study the connection between invariant measures and random attractors for ergodic RDS. As before, we denote by  $\Phi = \{\varphi_t^{\omega}, t \geq 0\}$  an RDS in a Polish space  $X$  over a group of measure-preserving transformations  $\theta = \{\theta_t, t \in \mathbb{R}\}$ . We shall assume that  $\Phi$  satisfies the two hypotheses below.

**Condition 4.2.13.** *Compactness:* There is a compact random set  $\mathbf{K} = \{K_{\omega}\}$  attracting trajectories of  $\varphi_t^{\omega}$  (in the sense specified in Theorem 4.2.5). Moreover, for any  $\varepsilon > 0$  there is  $\Omega_{\varepsilon} \in \mathcal{F}$  and a compact set  $\mathcal{C}_{\varepsilon} \subset X$  such that  $\mathbb{P}(\Omega_{\varepsilon}) \geq 1 - \varepsilon$  and

$$\varphi_t^{\theta_{-t}\omega} u \in \mathcal{C}_{\varepsilon} \quad \text{for } \omega \in \Omega_{\varepsilon}, u \in X, t \geq t_{\varepsilon}(u), \quad (4.59)$$

where the time  $t_{\varepsilon}(u) > 0$  depends only on  $u$  and  $\varepsilon$ .

**Condition 4.2.14.** *Mixing:* The RDS  $\Phi$  is of mixing type in the following sense: it has a unique stationary measure  $\mu$ , and for any  $f \in L_b(X)$  and any initial point  $u \in X$  we have

$$\mathfrak{P}_t f(u) = \mathbb{E} f(\varphi_t u) \rightarrow (f, \mu) = \int_X f(u) \mu(du) \quad \text{as } t \rightarrow \infty. \quad (4.60)$$

*Exercise 4.2.15.* Prove that if an RDS  $\Phi$  satisfies Condition 4.2.13, then there is a compact random set  $\mathbf{A} = \{A_{\omega}\}$  that satisfies the invariance property of Definition 4.2.6 and the attraction property of Definition 4.2.3. In particular,  $\mathbf{A}$  is a weak random attractor for  $\Phi$ . *Hint:* Repeat the scheme used in the proof of Theorem 4.2.5.

Recall that the concept of the semigroup associated with an RDS was introduced in Section 1.3.4. The following theorem shows that if a Markov RDS  $\Phi$  satisfies Conditions 4.2.13 and 4.2.14, then the associated semigroup  $\Theta_t$  also possesses a mixing property. For any sub- $\sigma$ -algebra  $\mathcal{G} \subset \mathcal{F}$ , denote by  $\mathbb{L}(X, \mathcal{G})$  the set of measurable functions  $F(\omega, u): \Omega \times X \rightarrow \mathbb{R}$  that are  $\mathcal{G}$ -measurable in  $\omega$  for any fixed  $u \in X$  and satisfy the condition

$$\text{ess sup}_{\omega \in \Omega} \|F(\omega, \cdot)\|_L < \infty. \tag{4.61}$$

**Theorem 4.2.16.** *Let us assume that  $\Phi$  satisfies Conditions 4.2.13 and 4.2.14. Let  $\mu \in \mathcal{P}(X)$  be the unique stationary measure for  $\Phi$  and let  $\mathfrak{M} \in \mathcal{P}(\Omega \times X, \mathbb{P})$  be the corresponding Markov invariant measure. Then for any  $F \in \mathbb{L}(X, \mathcal{F}^-)$  we have*

$$\mathbb{E} F(\Theta_t(\cdot, u)) \rightarrow (F, \mathfrak{M}) = \int_{\Omega} \int_X F(\omega, u) \mu_{\omega}(du) \mathbb{P}(d\omega) \quad \text{as } t \rightarrow \infty, \tag{4.62}$$

where  $u \in X$  is an arbitrary initial point.

Let us denote by  $\mu_{\omega}$  the disintegration of  $\mathfrak{M}$  and set

$$A_{\omega} = \begin{cases} \text{supp } \mu_{\omega}, & \omega \in \tilde{\Omega}, \\ \emptyset, & \omega \notin \tilde{\Omega}, \end{cases} \tag{4.63}$$

where  $\tilde{\Omega} \in \mathcal{F}$  is a set of full measure on which the limit (4.46) exists. By Corollary 1.6.5 in [Arn98],  $A_{\omega}$  is a closed random set. Moreover, it follows from (4.46) that  $A_{\omega}$  is measurable with respect to  $\mathcal{F}^-$  in the sense that the function  $\omega \mapsto \text{dist}_X(u, A_{\omega})$  is  $(\mathcal{F}^-, \mathcal{B}(\mathbb{R}))$ -measurable for any  $u \in X$ .

**Theorem 4.2.17.** *Let us assume that  $\Phi$  satisfies Conditions 4.2.13 and 4.2.14. Then the random set  $\{A_{\omega}\}$  defined by (4.63) is almost surely compact and forms a minimal weak random attractor for  $\Phi$ .*

*Proof of Theorem 4.2.16. Step 1.* We first assume that  $F(\omega, u) \in \mathbb{L}(X, \mathcal{F}_{[-\ell, 0]})$  for some  $\ell \geq 0$ . Since  $\theta_t$  preserves  $\mathbb{P}$ , for any  $t \geq s \geq 0$  we have

$$p_t(u) := \mathbb{E} F(\theta_t \omega, \varphi_t^{\omega} u) = \mathbb{E} F(\omega, \varphi_t^{\theta_{-t} \omega} u) = \mathbb{E} \mathbb{E} \{ F(\omega, \varphi_t^{\theta_{-t} \omega} u) \mid \mathcal{F}_{[-s, 0]} \}.$$

By the cocycle property (see (1.62)),

$$\varphi_t^{\theta_{-t} \omega} = \varphi_s^{\theta_{-s} \omega} \circ \varphi_{t-s}^{\theta_{-t} \omega}, \quad s \leq t.$$

Hence, setting  $F_s(\omega, u) = F(\omega, \varphi_s^{\theta_{-s} \omega} u)$ , for any  $s \leq t$  we derive

$$p_t(u) = \mathbb{E} \mathbb{E} \{ F_s(\omega, \varphi_{t-s}^{\theta_{-t} \omega} u) \mid \mathcal{F}_{[-s, 0]} \}. \tag{4.64}$$

We now note that  $F_s \in \mathbb{L}(X, \mathcal{F}_{[-s, 0]})$  if  $s \geq \ell$ . Since  $\varphi_{t-s}^{\theta_{-t} \omega} u$  is measurable with respect to  $\mathcal{F}_{[-t, -s]}$  and since the  $\sigma$ -algebras  $\mathcal{F}_{[-s, 0]}$  and  $\mathcal{F}_{[-t, -s]}$  are independent, it follows from (4.64) that

$$p_t(u) = \mathbb{E} \mathbb{E}' \{ F_s(\omega, \varphi_{t-s}^{\theta_{-t} \omega'} u) \} = \mathbb{E} (\mathfrak{F}_{t-s} F_s)(\omega, u), \tag{4.65}$$

where  $\ell \leq s \leq t$  and  $\mathbb{E}'$  denotes the expectation with respect to  $\omega'$ . In view of Condition 4.2.14 and the Lebesgue theorem, for any  $s \geq \ell$ , the right-hand side of (4.65) tends to  $\mathbb{E}(F_s(\omega, \cdot), \mu)$  as  $t \rightarrow +\infty$ . Recalling the definition of  $F_s$ , we see that

$$(F_s(\omega, \cdot), \mu) = (F(\omega, \cdot), \varphi_s^{\theta-s\omega} \mu) \rightarrow (F(\omega, \cdot), \mu_\omega) \quad \text{as } s \rightarrow \infty,$$

where we used Proposition 1.3.27. What has been said implies the relation

$$\lim_{t \rightarrow +\infty} p_t(u) = \mathbb{E}(F(\omega, \cdot), \mu_\omega),$$

which coincides with (4.62).

*Step 2.* We now show that (4.62) holds for functions of the form  $F(\omega, u) = f(u)g(\omega)$ , where  $f \in L_b(X)$  and  $g$  is a bounded  $\mathcal{F}^-$ -measurable function. To this end, we use a version of the monotone class theorem.

Let us fix  $f \in L_b(X)$  and denote by  $\mathcal{H}$  the set of bounded  $\mathcal{F}^-$ -measurable functions  $g$  for which convergence (4.62) with  $F = fg$  holds. It is clear that  $\mathcal{H}$  is a vector space containing the constant functions. Moreover, as was shown in Step 1, it contains all bounded functions measurable with respect to  $\mathcal{F}_{[-\ell, 0]}$  for some  $\ell \geq 0$ . Since the union of  $\mathcal{F}_{[-\ell, 0]}$ ,  $\ell \geq 0$ , generates  $\mathcal{F}^-$ , the required assertion will be proved as soon as we establish the following property: if  $g_n \in \mathcal{H}$  is an increasing sequence of non-negative functions such that  $g = \sup g_n$  is bounded, then  $g \in \mathcal{H}$ .

Suppose that a sequence  $\{g_n\} \subset \mathcal{H}$  satisfies the above conditions. Without loss of generality, we shall assume that  $0 \leq g, g_n \leq 1$ . By Egorov's theorem, for any  $\varepsilon > 0$  there is  $\Omega_\varepsilon \in \mathcal{F}$  such that  $\mathbb{P}(\Omega_\varepsilon) \geq 1 - \varepsilon$  and

$$\lim_{n \rightarrow +\infty} \sup_{\omega \in \Omega_\varepsilon} |g_n(\omega) - g(\omega)| = 0.$$

It follows that for any  $\varepsilon > 0$  there is an integer  $n_\varepsilon \geq 1$  such that  $n_\varepsilon \rightarrow +\infty$  as  $\varepsilon \rightarrow 0$  and

$$g_{n_\varepsilon}(\omega) \leq g(\omega) \leq g_{n_\varepsilon}(\omega) + \varepsilon + \mathbb{1}_{\Omega_\varepsilon^c}(\omega) \quad \text{for all } \omega \in \Omega.$$

Multiplying this inequality by  $f(\varphi_t^{\theta-t\omega} u)$ , taking the expectation, passing to the limit as  $t \rightarrow +\infty$ , and using the estimate  $\mathbb{P}(\Omega_\varepsilon^c) \leq \varepsilon$ , we derive

$$\begin{aligned} \mathbb{E}\{(f, \mu_\omega)g_{n_\varepsilon}(\omega)\} &\leq \liminf_{k \rightarrow +\infty} \mathbb{E}\{f(\varphi_t^{\theta-t\omega} u)g(\omega)\} \\ &\leq \limsup_{k \rightarrow +\infty} \mathbb{E}\{f(\varphi_t^{\theta-t\omega} u)g(\omega)\} \leq \mathbb{E}\{(f, \mu_\omega)g_{n_\varepsilon}(\omega)\} + 2\varepsilon. \end{aligned}$$

Since  $\varepsilon > 0$  is arbitrary and  $\mathbb{E}\{(f, \mu_\omega)g_{n_\varepsilon}(\omega)\} \rightarrow \mathbb{E}\{(f, \mu_\omega)g(\omega)\}$  as  $\varepsilon \rightarrow 0$  (by the monotone convergence theorem), we conclude that

$$\mathbb{E}\{f(\varphi_t^\omega u)g(\theta_t \omega)\} = \mathbb{E}\{f(\varphi_t^{\theta-t\omega} u)g(\omega)\} \xrightarrow{t \rightarrow +\infty} \mathbb{E}\{(f, \mu_\omega)g(\omega)\},$$

which means that  $g \in \mathcal{H}$ . This completes the proof of (4.62) in the case when  $F(\omega, u) = f(u)g(\omega)$ .

*Step 3.* Now we consider the general case. Let  $F \in \mathbb{L}(X, \mathcal{F}^-)$  be an arbitrary function such that  $\|F(\omega, \cdot)\|_{L(X)} \leq 1$  for a.e.  $\omega \in \Omega$ . For any  $u \in X$  and  $\varepsilon > 0$ , we choose  $t_\varepsilon(u) \geq 1$ ,  $\Omega_\varepsilon \in \mathcal{F}$ , and  $\mathcal{C}_\varepsilon \Subset X$  for which (4.59) holds. By the Arzelà–Ascoli theorem, the unit ball  $B_\varepsilon = \{f \in L_b(\mathcal{C}_\varepsilon) : \|f\|_L \leq 1\}$  is compact in the space  $C_b(\mathcal{C}_\varepsilon)$ , and therefore there is a finite set  $\{h_j\} \subset B_\varepsilon$  whose  $\varepsilon$ -neighbourhood contains  $B_\varepsilon$ . It follows that  $B_\varepsilon$  can be covered by non-intersecting Borel sets  $U_j \ni h_j$ ,  $j = 1, \dots, N$ , whose diameters do not exceed  $2\varepsilon$ . Let us denote by  $f_j \in L(X)$  arbitrary extensions of  $h_j$  to  $X$  such that  $\|f_j\|_L \leq 2$ . For instance, we can take

$$f_j(u) = \inf_{v \in \mathcal{C}_\varepsilon} (h_j(v) + \text{dist}_X(u, v) \wedge 1).$$

Let us consider the following approximation of  $F$ :

$$G_\varepsilon(\omega, u) = \sum_{j=1}^N f_j(u)g_j(\omega), \quad g_j(\omega) = \mathbb{I}_{U_j}(F_{\mathcal{C}_\varepsilon}(\omega, \cdot)),$$

where  $F_{\mathcal{C}_\varepsilon}(\omega, u)$  is the restriction of  $F$  to  $\Omega \times \mathcal{C}_\varepsilon$ . Since only one of the functions  $g_j$  can be non-zero, we have  $\|G_\varepsilon(\omega, \cdot)\|_\infty \leq 2$ . Therefore, for any  $u \in X$  and a.e.  $\omega \in \Omega$ , we derive

$$\begin{aligned} |G_\varepsilon(\omega, u) - F(\omega, u)| &\leq 2\varepsilon + \mathbb{I}_{\mathcal{C}_\varepsilon}(u) (\|G_\varepsilon(\omega, \cdot)\|_\infty + \|F(\omega, \cdot)\|_\infty) \\ &\leq 2\varepsilon + 3\mathbb{I}_{\mathcal{C}_\varepsilon}(u), \end{aligned} \quad (4.66)$$

where we used the inequality  $\|F(\omega, \cdot)\|_\infty \leq 1$ . Let us set

$$p_t(u) = \mathbb{E} F(\theta_t \omega, \varphi_t^\omega u), \quad p_t(u, \varepsilon) = \mathbb{E} G_\varepsilon(\theta_t \omega, \varphi_t^\omega u).$$

It is clear that

$$|p_t(u) - (F, \mathfrak{M})| \leq |p_t(u) - p_t(u, \varepsilon)| + |p_t(u, \varepsilon) - (G_\varepsilon, \mathfrak{M})| + |(G_\varepsilon - F, \mathfrak{M})|. \quad (4.67)$$

Let us estimate each term on the right-hand side of (4.67). Combining (4.59) and (4.66), for  $t \geq t_\varepsilon(u)$  we derive

$$\begin{aligned} |p_t(u) - p_t(u, \varepsilon)| &\leq |\mathbb{E}\{F(\omega, \varphi_t^{\theta-t\omega} u) - G_\varepsilon(\omega, \varphi_t^{\theta-t\omega} u)\}| \\ &\leq 2\varepsilon + 3\mathbb{P}\{\varphi_t^{\theta-t\omega} u \notin \mathcal{C}_\varepsilon\} \leq 2\varepsilon + 3\mathbb{P}(\Omega_\varepsilon^c) \leq 5\varepsilon. \end{aligned} \quad (4.68)$$

Furthermore, the functions  $g_j$  are  $\mathcal{F}^-$ -measurable, and hence, by Step 2, for any fixed  $\varepsilon > 0$ ,

$$p_t(u, \varepsilon) \rightarrow (G_\varepsilon, \mathfrak{M}) \quad \text{as } t \rightarrow +\infty. \quad (4.69)$$

Finally, inequality (4.66) implies that

$$|(G_\varepsilon - F, \mathfrak{M})| \leq 2\varepsilon + 3(\mathbb{I}_{\mathcal{C}_\varepsilon}, \mathfrak{M}) = 2\varepsilon + 3\mu(\mathcal{C}_\varepsilon). \quad (4.70)$$

Since  $\varepsilon > 0$  is arbitrary, it follows from (4.67) – (4.70) that the required convergence (4.62) will be established if we show that  $\mu(\mathcal{C}_\varepsilon^c) \rightarrow 0$  as  $\varepsilon \rightarrow 0$ .

To this end, we note that

$$\mu(\mathcal{C}_\varepsilon^c) = \int_X \mathbb{P}\{\varphi_t^\omega u \notin \mathcal{C}_\varepsilon\} \mu(du). \quad (4.71)$$

It follows from Condition 4.2.13 that, for any fixed  $u \in X$ ,

$$\limsup_{t \rightarrow +\infty} \mathbb{P}\{\varphi_t^\omega u \notin \mathcal{C}_\varepsilon\} = \limsup_{t \rightarrow +\infty} \mathbb{P}\{\varphi_t^{\theta^{-t}\omega} u \notin \mathcal{C}_\varepsilon\} \leq \varepsilon.$$

Passing to the limit  $t \rightarrow +\infty$  in (4.71), we conclude that  $\mu(\mathcal{C}_\varepsilon^c) \leq \varepsilon$  for any  $\varepsilon > 0$ . This completes the proof of Theorem 4.2.16.  $\square$

*Proof of Theorem 4.2.17.* We first show that  $A_\omega$  is compact for almost every  $\omega \in \Omega$ . To this end, note that, by Exercise 4.2.15, the RDS possesses a (compact) weak random attractor  $\{A'_\omega\}$ . Furthermore, by Proposition 4.2.12, we have  $A_\omega \subset A'_\omega$  for almost every  $\omega$ . Since  $A_\omega$  is closed, we obtain the required result.

We now prove that the compact random set  $\{A_\omega\}$  is a random attractor. Let us fix  $\delta \in (0, 1)$  and consider the function

$$F(\omega, u) = 1 - \frac{\text{dist}_X(u, A_\omega)}{\delta} \wedge 1, \quad u \in X, \quad \omega \in \Omega.$$

We claim that  $F \in \mathbb{L}(X, \mathcal{F}^-)$ . Indeed, the definition implies that  $F(\omega, u)$  is a bounded measurable function and that

$$|F(\omega, u) - F(\omega, v)| \leq \frac{\text{dist}_X(u, v)}{\delta} \quad \text{for all } u, v \in X, \omega \in \Omega.$$

Thus,  $F$  satisfies (4.61). Since  $A_\omega$  is a compact random set measurable with respect to  $\mathcal{F}^-$ , we conclude that so is the function  $F$ . This proves the required properties of  $F$ .

Since  $F(\omega, u) = 1$  for  $u \in A_\omega$ , we see that  $(F, \mathfrak{M}) = 1$ . So, applying Theorem 4.2.16, we get

$$\mathbb{E} F(\theta_t \omega, \varphi_t^\omega u) = 1 - \mathbb{E} \left( \frac{\text{dist}_X(\varphi_t^\omega u, A_{\theta_t \omega})}{\delta} \wedge 1 \right) \rightarrow (F, \mathfrak{M}) = 1.$$

That is,

$$p_t(u) := \mathbb{E} \left( \frac{\text{dist}_X(\varphi_t^\omega u, A_{\theta_t \omega})}{\delta} \wedge 1 \right) \rightarrow 0. \quad (4.72)$$

We now note that, by Chebyshev's inequality,

$$\mathbb{P}\{\text{dist}_X(\varphi_t^\omega u, A_{\theta_t \omega}) > \delta\} \leq \frac{p_t(u)}{\delta}.$$

In view of (4.72), the right-hand side of this inequality goes to zero as  $t \rightarrow +\infty$ . This completes the proof of the fact that  $A_\omega$  is a random attractor.

To show that  $A_\omega$  is a minimal random attractor, it suffices to note that, by Proposition 4.2.12, the Markov invariant measure  $\mathfrak{M}$  is supported by any random attractor  $\{A'_\omega\}$ , and therefore  $\text{supp } \mu_\omega \subset A'_\omega$  for a.e.  $\omega \in \Omega$ .  $\square$



#### 4.2.4 Application to the Navier–Stokes system

In this subsection, we show that Theorems 4.2.16 and 4.2.17 can be applied to the Navier–Stokes system (3.84) perturbed by a white noise force (2.66). To this end, it suffices to check that Conditions 4.2.13 and 4.2.14 are fulfilled. As was proved in Theorem 3.3.1, the second condition holds if the coefficients  $b_j$  are all positive and satisfy (2.67). Thus, we shall concentrate on Condition 4.2.13 and prove that it holds for any sequence  $b_j$  satisfying the hypothesis

$$\mathfrak{B}_1 = \sum_{j=1}^{\infty} \alpha_j b_j^2 < \infty. \quad (4.73)$$

Recall that a Markov RDS associated with the Navier–Stokes system perturbed by a spatially regular white noise was constructed in Section 2.4.4. We denote by  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  the corresponding filtered probability space and by  $\theta_t : \Omega \rightarrow \Omega$  the group of shift operators on  $\Omega$ . We shall show that the required properties are implied by the following proposition.

**Proposition 4.2.18.** *There is an increasing sequence of subsets  $\Omega_k \in \mathcal{F}$  satisfying the following properties.*

- (i) *The union  $\Omega_* = \cup_k \Omega_k$  is invariant under the group  $\theta_t$  and has the full measure.*
- (ii) *For any  $k \geq 1$  there is an almost surely finite random constant  $R_k$  such that*

$$\|\varphi_t^{\theta^{-t}\omega}(w)\|_1 \leq R_k(\omega) \quad \text{for } \omega \in \Omega_k, w \in H, t \geq T_k(w), \quad (4.74)$$

where  $T_k(w) \geq 0$  is a deterministic constant.

Taking this result for granted, let us check Condition 4.2.13. We first prove the existence of a compact random set attracting the trajectories. Let us set  $\Omega_*$  be the set defined in Proposition 4.2.18 and let

$$K_\omega = \begin{cases} B_V(R_k(\omega)) & \text{for } \omega \in \Omega_k \setminus \Omega_{k-1}, \\ \emptyset, & \text{for } \omega \notin \Omega_*, \end{cases}$$

where  $\Omega_0 = \emptyset$ . Then  $\mathbf{K} = \{K_\omega\}$  is a compact random set in  $H$ . We claim that it is an absorbing set. Namely, there is a measurable function  $T : H \times \Omega_* \rightarrow \mathbb{R}_+$  such that

$$\varphi_t^{\theta^{-t}\omega}(w) \in K_\omega \quad \text{for any } w \in H, \omega \in \Omega_*, t \geq T(w, \omega). \quad (4.75)$$

Indeed, if we set  $T(w, \omega) = T_k(w)$  for  $\omega \in \Omega_k \setminus \Omega_{k-1}$ , then (4.75) follows immediately from (4.74). Furthermore, since  $T_k$  is almost surely finite, so is  $T$ .

We now prove (4.59). Let us fix  $\varepsilon > 0$  and choose  $k_\varepsilon \geq 1$  so large that  $\mathbb{P}(\Omega_{k_\varepsilon}) \geq 1 - \frac{\varepsilon}{2}$ . We next find  $\rho_\varepsilon > 0$  such that  $\mathbb{P}\{R_{k_\varepsilon} > \rho_\varepsilon\} \leq \frac{\varepsilon}{2}$ . It is

straightforward to see that (4.59) holds with  $\mathcal{C}_\varepsilon = B_{\rho_\varepsilon}(V)$  and  $t_\varepsilon(u) = T_{k_\varepsilon}(u)$ . Thus, it suffices to establish Proposition 4.2.18. Before doing so, we summarise the properties resulting from the application of Theorems 4.2.16 and 4.2.17 to the Navier–Stokes system (3.84).

**Theorem 4.2.19.** *Let us assume that the random part  $\eta$  of the external force in (3.84) is such that  $b_j \neq 0$  for all  $j \geq 1$  and  $\mathfrak{B}_1 < \infty$ . Let  $\mu$  be the corresponding unique stationary measure and let  $\{\mu_\omega\}$  be the Markov disintegration of the invariant measure associated with  $\mu$ . Then there is a full measure event  $\tilde{\Omega} \subset \Omega$  such that the random set  $\mathcal{A}_\omega$  defined by (4.63) is a minimal weak random attractor for (3.84).*

Let us mention that similar results hold for the Navier–Stokes system perturbed by a random kick force. Furthermore, the weak random attractor described in Theorem 4.2.19 is contained in the *global random attractor*, which may be even larger than the one constructed in Theorem 4.2.5. A minimal global random attractor  $\{\mathcal{A}_\omega\}$  is uniquely defined (in the sense described in Definition 4.2.4), and for many stochastic PDE’s (including the 2D Navier–Stokes system), it is known that the Hausdorff dimension of almost every set  $\mathcal{A}_\omega$  can be estimated by a deterministic constant; see [Deb97, Deb98, LR06].

*Proof of Proposition 4.2.18.* We shall need the following auxiliary result established at the end of this subsection.

**Lemma 4.2.20.** *Suppose that (4.73) holds and define a process  $\zeta(t)$  as in (2.98). Then for any  $\alpha > 0$ , the stochastic integral*

$$z_\alpha(t) = \int_{-\infty}^t e^{-\alpha(t-s)L} d\zeta(s), \quad t \in \mathbb{R}, \quad (4.76)$$

*defines an  $H$ -valued stationary process possessing the following properties.*

- (i) *The process  $z_\alpha$  is adapted to the filtration  $\{\mathcal{F}_t, t \in \mathbb{R}\}$ , and its trajectories belong to the space  $C(\mathbb{R}; V) \cap L_{\text{loc}}^2(\mathbb{R}; V^2)$  with probability 1 and satisfy the equation*

$$\dot{z} + \alpha Lz = \partial_t \zeta(t), \quad t \in \mathbb{R}. \quad (4.77)$$

- (ii) *For any  $\gamma > 0$  and almost every  $\omega \in \Omega$ , we have*

$$\sup_{-1 \leq t \leq 0} \|z_\alpha(t)\|_1 < \infty, \quad (4.78)$$

$$\lim_{t \rightarrow -\infty} |t|^{-\gamma} \|z_\alpha(t)\|_2^2 = 0, \quad (4.79)$$

$$\lim_{t \rightarrow -\infty} \frac{1}{|t|} \int_t^0 |\nabla z_\alpha(s)|_2^2 ds = \frac{\mathfrak{B}}{2\alpha}, \quad (4.80)$$

$$\lim_{t \rightarrow -\infty} \frac{1}{|t|} \int_t^0 |Lz_\alpha(s)|_2^2 ds = \frac{\mathfrak{B}_1}{2\alpha}. \quad (4.81)$$

We wish to construct an increasing sequence  $\{\Omega_k\} \subset \mathcal{F}$  satisfying (i) and random constants  $R_k \geq 0$  such that the solution  $u$  of Eq. (3.84) issued from  $w$  at time  $t = t_0$  satisfies the inequality

$$\|u(0)\|_1 \leq R_k(\omega) \quad \text{for } \omega \in \Omega_k, w \in H, t_0 \leq -T_k(w). \quad (4.82)$$

This will imply the required inequality (4.74).

Let  $u(t)$  be a solution (3.84) satisfying the initial condition

$$u(t_0) = w.$$

We can write it in the form  $u = z_\alpha + v$ , where the (large) constant  $\alpha > 0$  will be chosen later. Then  $v$  must satisfy the equations

$$\dot{v} + \nu Lv + B(v + z_\alpha) = h + (\alpha - \nu)Lz_\alpha, \quad v(t_0) = w - z_\alpha(t_0). \quad (4.83)$$

For fixed  $M > 0$  and  $k \in \mathbb{Z}_+$ , let  $\Omega_k$  be the set of those  $\omega \in \Omega$  for which

$$\frac{1}{|t|} \int_t^0 |Lz_\alpha(s)|_2^2 ds \leq \frac{M}{\alpha}, \quad e^{\nu\alpha_1 t} |z_\alpha(t)|_2^2 \leq 1 \quad \text{for } t \leq -k. \quad (4.84)$$

By (4.81), for a sufficiently large  $M > 0$  we have  $\mathbb{P}(\Omega_k) \rightarrow 1$  as  $k \rightarrow \infty$ , whence it follows that the union  $\Omega_* = \cup_k \Omega_k$  is a set of full measure. Moreover, it is easy to see that it is invariant under the shifts. The proof of (4.82) is divided into two steps.

*Step 1.* We first show that

$$\sup_{-1 \leq t \leq 0} |v(t)|_2 \leq r_k(\omega) \quad \text{for } \omega \in \Omega_k, w \in H, t_0 \leq -T_k(w), \quad (4.85)$$

where  $r_k$  is an almost surely finite random constant and  $T_k$  depends only on  $w$ . Taking the scalar product of (4.83) with  $v$  and carrying out some transformations based on Schwarz's inequality and standard estimates of the nonlinear term, we derive

$$\partial_t |v|_2^2 + \nu \|v\|_1^2 \leq C_1 g(t) + C_1 |\nabla z_\alpha|_2^2 |v|_2^2, \quad (4.86)$$

where  $g = |h|_2^2 + \alpha^2 |\nabla z_\alpha|_2^2 + |z_\alpha|_2 |Lz_\alpha|_2$ . Since  $\|v\|_1 \geq \alpha_1 |v|_2^2$ , applying Gronwall's inequality, we obtain

$$|v(t)|_2^2 \leq e^{-A(t,t_0)} |w - z_\alpha(t_0)|_2^2 + C_1 \int_{t_0}^t e^{-A(t,s)} g(s) ds, \quad (4.87)$$

where we set

$$A(t, s) = \int_s^t (\nu\alpha_1 - C_1 |\nabla z_\alpha(r)|_2^2) dr.$$

Choosing  $\alpha > 0$  sufficiently large, we see from the first inequality in (4.84) that

$$A(t, s) \geq \nu\alpha_1(t - s - 1) \quad \text{for } -1 \leq t \leq 0, s \leq -k, \omega \in \Omega_k.$$

Substituting this into (4.87) and using the second relation in (4.84), for  $t_0 \leq -k$ ,  $-1 \leq t \leq 0$ , and  $\omega \in \Omega_k$  we obtain

$$|v(t)|_2^2 \leq C_2(1 + e^{\nu\alpha_1 t_0} |w|_2^2) + C_2 \left( \int_{-\infty}^{-k} e^{-\nu\alpha_1(t-s-1)} g(s) ds + \int_{-k}^t e^{-A(t,s)} g(s) ds \right).$$

This implies the required inequality (4.85), in which

$$r_k(\omega) = C_2 \left( 2 + \int_{-\infty}^{-k} e^{\nu\alpha_1(s-2)} g(s) ds + \sup_{-1 \leq t \leq 0} \int_{-k}^t e^{-A(t,s)} g(s) ds \right).$$

*Step 2.* We now use the regularising property of the Navier–Stokes dynamics to prove that

$$\|v(0)\|_1 \leq \tilde{r}_k(\omega) \quad \text{for } \omega \in \Omega_k, w \in H, t_0 \leq -T_k(w), \quad (4.88)$$

where  $\tilde{r}_k$  are random constants. Once this inequality is established, the required estimate (4.82) with  $R_k(\omega) = \|z_\alpha(0)\|_1 + \tilde{r}_k(\omega)$  will follow from the representation  $u(0) = z_\alpha(0) + v(0)$ .

Integrating (4.86) in  $t \in (-1, 0)$  and using (4.85) and the first inequality in (4.84), we get

$$\begin{aligned} \nu \int_0^1 \|v(t)\|_1^2 dt &\leq |v(-1)|_2^2 + C_1 \int_{-1}^0 g(t) dt + C_1 \int_{-1}^0 |\nabla z_\alpha|_2^2 |v|_2^2 dt \\ &\leq r_k(\omega) \left( 1 + C_1 \int_{-k}^0 |\nabla z_\alpha|_2^2 dt \right) + C_3 \int_{-k}^0 (1 + |Lz_\alpha|_2^2) dt \\ &\leq C_4 k r_k(\omega). \end{aligned} \quad (4.89)$$

We now take the scalar product in  $H$  of (4.83) with  $2(t+1)Lv$ . After some transformations, we obtain

$$\begin{aligned} \partial_t((t+1)|L^{1/2}v|_2^2) + \nu(t+1)|Lv|_2^2 &\leq \|v\|_1^2 + C_5(t+1)(|h|_2^2 + \|z_\alpha\|_2^2) \\ &\quad + C_5(t+1)(|v+z_\alpha|_2|L(v+z_\alpha)|_2)^{1/2} \|v+z_\alpha\|_1 |Lv|_2. \end{aligned}$$

It follows from (4.85) and (4.78) that the last term on the right-hand side of this inequality can be estimated by

$$C_6(\omega)(t+1)(|Lv|_2^{3/2} + |Lz_\alpha|_2^{3/2})(|L^{1/2}v|_2 + 1),$$

where  $C_6$  is a random constant. Therefore, the function  $\psi(t) = (t+1)|L^{1/2}v(t)|_2$  satisfies the differential inequality

$$\psi'(t) \leq C_7(\omega)\psi(t) + C_7(\omega)(|Lz_\alpha|_2^2 + \|v\|_1^2 + 1), \quad -1 \leq t \leq 0.$$

By Gronwall's inequality, we obtain

$$\begin{aligned} \|v(0)\|_1^2 &\leq e^{C_7(\omega)} \int_{-1}^0 (|Lz_\alpha|_2^2 + \|v\|_1^2 + 1) dt \\ &\leq e^{C_7(\omega)} \left( \int_{-k}^0 |Lz_\alpha|_2^2 dt + \int_{-1}^0 \|v\|_1^2 dt + 1 \right). \end{aligned}$$

Recalling (4.84) and (4.89), we arrive at the required inequality (4.88). This completes the proof of the proposition.  $\square$

*Proof of Lemma 4.2.20.* The fact that  $z_\alpha$  is a well-defined  $H$ -valued stationary process possessing properties mentioned in (i) follows easily by finite-dimensional approximations (cf. proof of Proposition 2.4.2). Therefore we shall confine ourselves to the proof of assertion (ii).

An analogue of Proposition 2.4.10 is true for the process  $z_\alpha$ . By stationarity, it follows that

$$\mathbb{E} \left( \sup_{T \leq t \leq T+1} \exp(\sigma \nu \|z_\alpha\|_1^2) \right) = C_1 < \infty \quad \text{for all } T \in \mathbb{R}, \quad (4.90)$$

where  $\sigma > 0$  does not depend on  $\nu$  and  $T$ . Inequality (4.90) implies, in particular, that (4.78) holds. Let us introduce the events

$$\Gamma_k = \left\{ \sup_{-k \leq t \leq 1-k} |z_\alpha(t)|_2 > k^{-\gamma/4} \right\}, \quad k \geq 1.$$

Then (4.90) implies that  $\sum_k \mathbb{P}(\Gamma_k) < \infty$ , whence it follows that  $|z_\alpha(t)|_2 \leq t^{-\gamma/4}$  for  $t \leq -t_0$ , where  $t_0 > 0$  is an almost surely finite random constant. We thus obtain (4.79).

To prove (4.80), we apply Itô's formula to the process  $|z_\alpha(t)|_2^2$ . This results in

$$|z_\alpha(0)|_2^2 + 2\alpha \int_t^0 |\nabla z_\alpha(s)|_2^2 ds = |z_\alpha(t)|_2^2 + \mathfrak{B}|t| + 2 \int_t^0 \langle z_\alpha(s), d\zeta(s) \rangle.$$

Dividing this relation by  $2\alpha|t|$ , we derive

$$\frac{1}{|t|} \int_t^0 |\nabla z_\alpha(s)|_2^2 ds = \frac{\mathfrak{B}}{2\alpha} + \frac{1}{2\alpha|t|} \left( |z_\alpha(t)|_2^2 - |z_\alpha(0)|_2^2 + 2 \int_t^0 \langle z_\alpha, d\zeta \rangle \right). \quad (4.91)$$

As was proved above,

$$\frac{1}{|t|} (|z_\alpha(0)|_2^2 + |z_\alpha(t)|_2^2) \rightarrow 0 \quad \text{as } t \rightarrow -\infty \text{ almost surely.}$$

On the other hand, using Theorem 7.12.1 and inequality (4.90), one can show

$$\frac{1}{|t|} \left| \int_t^0 \langle z_\alpha, d\zeta \rangle \right| \rightarrow 0 \quad \text{as } t \rightarrow -\infty \text{ almost surely.}$$

Combining this with (4.91), we obtain (4.80). Finally, the proof of (4.81) is based on the application of Itô's formula to the functional  $\|z\|_1^2$  and can be carried out by a similar argument.  $\square$

### 4.3 Dependence of a stationary measure on the random force

In this section, we discuss some continuity properties of stationary measures with respect to parameters. We shall study only the dependence of a stationary measure on the random component of an external force. However, similar results remain true when one varies the deterministic component of an external force or the viscosity. We begin with the case in which the random force  $\eta$  entering Eq. (3.84) depends continuously on a parameter  $a$ . It is shown that if the hypotheses ensuring the uniqueness of a stationary measure are fulfilled, then the stationary measure continuously depends on  $a$ . This implies, in particular, that a solution of the Cauchy problem converges in distribution to a limiting solution *uniformly* in time. We next turn to the case of a high-frequency random kicks that converge to a spatially regular white noise and establish a similar result on convergence of stationary measures.

#### 4.3.1 Regular dependence on parameters

Let us consider the Navier–Stokes equations (3.84), in which  $\nu > 0$  and  $h \in H$  are fixed, and  $\eta$  is a spatially regular white noise depending on a parameter  $a$  that belongs to a metric space  $X$ :

$$\eta(t, x) = \eta_a(t, x) = \frac{\partial}{\partial t} \zeta_a(t, x), \quad \zeta_a(t, x) = \sum_{j=1}^{\infty} b_j(a) \beta_j(t) e_j(x). \quad (4.92)$$

Here  $\beta_j$  and  $e_j$  are the same as in (2.98), and  $b_j : X \rightarrow \mathbb{R}_+$  are some continuous functions. We shall always assume that

$$\mathfrak{B} := \sup_{a \in X} \sum_{j=1}^{\infty} b_j^2(a) < \infty. \quad (4.93)$$

We denote by  $P_t^a(u, \Gamma)$  the transition function associated with  $\eta_a$ , and by  $\mathfrak{P}_t(a)$  and  $\mathfrak{P}_t^*(a)$  the corresponding Markov semigroups. The following theorem shows that if the stationary measure is unique for some  $a = \hat{a}$ , then it is the only accumulation point of stationary distributions in the sense of weak convergence as  $a \rightarrow \hat{a}$ .

**Theorem 4.3.1.** *Under the above hypotheses, let  $b_j(\hat{a}) > 0$  for all  $j \geq 1$ , where  $\hat{a} \in X$  is a given point, and let  $\{\mu_a, a \in X\}$  be any family of stationary measures for Eq. (3.84) with  $\eta = \eta_a$ . Then  $\mu_a \rightarrow \mu_{\hat{a}}$  in the weak topology as  $a \rightarrow \hat{a}$ .*

*Proof.* We first outline the scheme of the proof. We wish to prove that

$$\|\mu_a - \mu_{\hat{a}}\|_L^* \rightarrow 0 \quad \text{as } \text{dist}_X(a, \hat{a}) \rightarrow 0. \quad (4.94)$$

To this end, we note that  $\mathfrak{P}_t^*(a)\mu_a = \mu_a$  for all  $t \geq 0$ , whence it follows that

$$\mu_a - \mu_{\hat{a}} = (\mathfrak{P}_t^*(a)\mu_a - \mathfrak{P}_t^*(\hat{a})\mu_a) + (\mathfrak{P}_t^*(\hat{a})\mu_a - \mu_{\hat{a}}). \quad (4.95)$$

It suffices to show that both terms on the right-hand side of (4.95) go to zero in the dual-Lipschitz metric as  $\text{dist}_X(a, \hat{a}) \rightarrow 0$ . Suppose for any  $T > 0$  and  $R > 0$  we have proved that

$$\sup_{0 \leq t \leq T} \sup_{|v|_2 \leq R} \|P_t^a(v, \cdot) - P_t^{\hat{a}}(v, \cdot)\|_L^* \rightarrow 0 \quad \text{as} \quad \text{dist}_X(a, \hat{a}) \rightarrow 0. \quad (4.96)$$

In this case, using Prokhorov's theorem, one can show that, for any  $T > 0$  and any tight family  $\Lambda \subset \mathcal{P}(H)$ , we have

$$\sup_{0 \leq t \leq T} \sup_{\lambda \in \Lambda} \|\mathfrak{P}_t^*(a)\lambda - \mathfrak{P}_t^*(\hat{a})\lambda\|_L^* \rightarrow 0 \quad \text{as} \quad \text{dist}_X(a, \hat{a}) \rightarrow 0. \quad (4.97)$$

Furthermore, by Theorem 3.3.1,

$$\sup_{\lambda \in \Lambda} \|\mathfrak{P}_t^*(\hat{a})\lambda - \mu_{\hat{a}}\|_L^* \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty. \quad (4.98)$$

Combining (4.95), (4.97), and (4.98) and noting that the family  $\{\mu_a\}$  is tight in  $H$ , we arrive at the required result. The accurate proof is divided into four steps.

*Step 1.* Let us prove (4.96), where  $T > 0$  is any fixed number. To this end, we first establish an estimate for the difference between two solutions for Eq. (3.84) with different right-hand sides. Let  $u^a(t, x)$  be a solution of problem (3.84), (3.85) in which  $\eta = \eta_a$  and  $u_0 = v$ . Then the difference  $u(t) = u^a(t) - u^{\hat{a}}(t)$  vanishes at  $t = 0$  and satisfies the equation

$$\dot{u} + \nu Lu + B(u^a(t), u) + B(u, u^{\hat{a}}(t)) = \xi(t, x),$$

where we set

$$\xi(t, x) = \frac{\partial}{\partial t} \sum_{j=1}^{\infty} (b_j(a) - b_j(\hat{a})) \beta_j(t) e_j(x).$$

Let us denote by  $z(t) = z(t; a, \hat{a})$  a solution of the linear problem

$$\dot{z} + \nu Lz = \xi(t, x), \quad z(0) = 0.$$

Then we can write  $u = z + w$ , where  $w$  vanishes at  $t = 0$  and satisfies the equation

$$\dot{w} + \nu Lw + B(u^a, z + w) + B(z + w, u^{\hat{a}}) = 0. \quad (4.99)$$

For any curve  $\xi \in C(0, T; H) \cap L^2(0, T; H^1)$ , let

$$\mathcal{E}_\xi(t) = |\xi(t)|_2^2 + \nu \int_0^t |\nabla \xi(s)|_2^2 ds.$$

The argument used in the proof of Proposition 2.4.2 (see the derivation of (2.109)) enables one to show that

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} \mathcal{E}_z(t) \right) \leq C_1 \sum_{j=1}^{\infty} |b_j(a) - b_j(\hat{a})|^2.$$

Recalling that  $b_j$  are continuous functions that satisfy (4.93), we conclude that

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} \mathcal{E}_z(t)\right) \rightarrow 0 \quad \text{as} \quad \text{dist}_X(a, \hat{a}) \rightarrow 0. \quad (4.100)$$

Furthermore, Corollary 2.4.11 implies that if  $|v|_2 \leq R$ , then

$$\sup_{a \in X} \mathbb{E}\left(\sup_{0 \leq t \leq T} \mathcal{E}_{u^a}(t)\right) \leq C_2 = C_2(R, T). \quad (4.101)$$

Taking the scalar product of (4.99) with  $2w$  and carrying out some standard transformations (cf. proof of Proposition 2.1.25), we obtain

$$\begin{aligned} \partial_t |w|_2^2 + 2\nu \|w\|_1^2 &\leq C_3 \|w\|_1 |w|_2 \|u^{\hat{a}}\|_1 \\ &\quad + C_3 (\|u^a\|_1 |u^a|_2 + \|u^{\hat{a}}\|_1 |u^{\hat{a}}|_2)^{1/2} (\|z\|_1 |z|_2)^{1/2} \|w\|_1, \end{aligned}$$

whence it follows that

$$\partial_t |w|_2^2 + \nu \|w\|_1^2 \leq C_4 \|u^{\hat{a}}\|_1^2 |w|_2^2 + C_4 (\|u^a\|_1 |u^a|_2 + \|u^{\hat{a}}\|_1 |u^{\hat{a}}|_2) \|z\|_1 |z|_2,$$

where  $C_4 > 0$  depends on  $\nu$ . Application of Gronwall's inequality results in

$$\sup_{0 \leq t \leq T} \mathcal{E}_w(t) \leq C_5(\nu) \exp\left(C_5(\nu) \int_0^T \|u^{\hat{a}}(t)\|_1^2 dt\right) K(a, \hat{a}), \quad (4.102)$$

where we set

$$K(a, \hat{a}) = \int_0^T (\|u^a\|_1 |u^a|_2 + \|u^{\hat{a}}\|_1 |u^{\hat{a}}|_2) \|z\|_1 |z|_2 dt.$$

Suppose we have shown that, for any positive constants  $\delta$  and  $R$ ,

$$\sup_{|v|_2 \leq R} \mathbb{P}\left\{\sup_{0 \leq t \leq T} \mathcal{E}_u(t) > \delta\right\} \rightarrow 0 \quad \text{as} \quad a \rightarrow \hat{a}. \quad (4.103)$$

Let us take any function  $f : H \rightarrow \mathbb{R}$  such that  $\|f\|_L \leq 1$  and fix an initial datum  $v \in B_H(R)$ . Denoting by  $\Omega_{\delta, a}$  the event on the left-hand side of (4.103), for  $0 \leq t \leq T$  we write

$$\begin{aligned} |\mathbb{E}(f(u^a(t)) - f(u^{\hat{a}}(t)))| &\leq \mathbb{E} \left| \mathbb{1}_{\Omega_{\delta, a}^c} (f(u^a(t)) - f(u^{\hat{a}}(t))) \right| + 2\mathbb{P}(\Omega_{\delta, a}) \\ &\leq \sqrt{\delta} + 2\mathbb{P}(\Omega_{\delta, a}). \end{aligned}$$

Since  $f$  and  $v$  were arbitrary, we arrive at the required convergence (4.96).

*Step 2.* We now prove (4.103). Since  $\mathcal{E}_u(t) \leq 2(\mathcal{E}_z(t) + \mathcal{E}_w(t))$  for all  $t \geq 0$ , in view of (4.100), it suffices to show that

$$\sup_{|v|_2 \leq R} \mathbb{P}\{\mathcal{E}(w) > \delta\} \rightarrow 0 \quad \text{as} \quad a \rightarrow \hat{a}, \quad (4.104)$$



where for a function  $g(t, x)$  we write

$$\mathcal{E}(g) = \sup_{0 \leq t \leq T} \mathcal{E}_g(t).$$

For any  $M > 0$ , we have

$$\begin{aligned} \mathbb{P}\{\mathcal{E}(w) > \delta\} &\leq \mathbb{P}\{\mathcal{E}(u^a) > M\} + \mathbb{P}\{\mathcal{E}(u^{\hat{a}}) > M\} \\ &\quad + \mathbb{P}\{\mathcal{E}(w) > \delta, \mathcal{E}(u^a) \leq M, \mathcal{E}(u^{\hat{a}}) \leq M\}. \end{aligned}$$

In view of (4.101), the first two terms in the right-hand side of this inequality can be made arbitrarily small, uniformly in  $v \in B_H(R)$ , by choosing a large  $M$ . Therefore, convergence (4.104) will be established if we show that

$$\sup_{|v|_2 \leq R} \mathbb{P}\{\mathcal{E}(w) > \delta, \mathcal{E}(u^a) \leq M, \mathcal{E}(u^{\hat{a}}) \leq M\} \rightarrow 0 \quad \text{as } a \rightarrow \hat{a}, \quad (4.105)$$

for any positive  $\delta$  and  $M$ . To this end, note that, in view of (4.102), on the set  $\{\mathcal{E}(u^a) \leq M, \mathcal{E}(u^{\hat{a}}) \leq M\}$ , we have

$$\mathcal{E}(w) \leq C_6 K(a, \hat{a}) \leq C_7 \mathcal{E}(z).$$

Combining this with (4.100) and Chebyshev's inequality, we easily prove (4.105).

*Step 3.* Let us prove (4.97). Let  $\Lambda \subset \mathcal{P}(H)$  be a tight family in  $H$ . Then, for any  $\varepsilon > 0$  and a sufficiently large  $R_\varepsilon > 0$ , we have

$$\sup_{\lambda \in \Lambda} \lambda(B_H^c(R_\varepsilon)) \leq \varepsilon. \quad (4.106)$$

On the other hand, the obvious relation

$$\mathfrak{P}_t^*(a)\lambda - \mathfrak{P}_t^*(\hat{a})\lambda = \int_H (P_t^a(z, \cdot) - P_t^{\hat{a}}(z, \cdot))\lambda(dz)$$

implies that, for any  $R > 0$ ,

$$\|\mathfrak{P}_t^*(a)\lambda - \mathfrak{P}_t^*(\hat{a})\lambda\|_L^* \leq \sup_{|v|_2 \leq R} \|P_t^a(v, \cdot) - P_t^{\hat{a}}(v, \cdot)\|_L^* + \lambda(B_H^c(R)).$$

Combining this inequality with (4.106) and (4.96) (where  $R = R_\varepsilon$ ), we get (4.97).

*Step 4.* To complete the proof of (4.94), it remains to show that the family  $\{\mu_a\} \subset \mathcal{P}(H)$  is tight in  $H$ . To see this, let us note that, in view of (2.157), we have

$$\int_H \|u\|_1^2 \mu_a(du) \leq C_8(v) \quad \text{for all } a \in X.$$

Chebyshev's inequality and the compactness of the embedding  $V \subset H$  prove that the family  $\{\mu_a\}$  is tight in  $H$ . The proof of Theorem 4.3.1 is complete.  $\square$

As a consequence of Theorem 4.3.1, we have the following result on uniform convergence of distributions of solutions for Cauchy's problem.

**Corollary 4.3.2.** *Under the hypotheses of Theorem 4.3.1, for any compact subset  $\Lambda \subset \mathcal{P}(H)$  there is a continuous function  $A_\Lambda(\rho) > 0$  going to zero with  $\rho$  such that*

$$\sup_{t \geq 0} \|\mathfrak{P}_t^*(a)\lambda_1 - \mathfrak{P}_t^*(\hat{a})\lambda_2\|_L^* \leq A_\Lambda(\|\lambda_1 - \lambda_2\|_L^* + \text{dist}_X(a, \hat{a})), \quad (4.107)$$

where  $\lambda_1, \lambda_2 \in \Lambda$  and  $a \in X$ .

In particular, if  $u^a(t, x)$  is a solution of (3.84), (3.85) with  $\eta = \eta_a$ , then

$$\sup_{t \geq 0} \|\mathcal{D}(u^a(t)) - \mathcal{D}(u^{\hat{a}}(t))\|_L^* \rightarrow 0 \quad \text{as } a \rightarrow \hat{a}.$$

*Proof of Corollary 4.3.2.* Let us fix a compact subset  $\Lambda \subset \mathcal{P}(H)$  and define the function

$$\Delta(t, a, \lambda_1, \lambda_2) = \|\mathfrak{P}_t^*(a)\lambda_1 - \mathfrak{P}_t^*(\hat{a})\lambda_2\|_L^*.$$

To establish (4.107), it suffices to prove that for any  $\varepsilon > 0$  there are positive constants  $T$  and  $\delta$  such that

$$\sup_{t \geq T} \Delta(t, a, \lambda_1, \lambda_2) \leq \varepsilon \quad \text{for } \lambda_1, \lambda_2 \in \Lambda, \text{dist}_X(a, \hat{a}) \leq \delta. \quad (4.108)$$

Indeed, an argument similar to that used in Step 1 of the proof of Theorem 4.3.1 enables one to show that

$$\sup_{0 \leq t \leq T} \Delta(t, \hat{a}, \lambda_1, \lambda_2) \rightarrow 0 \quad \text{as } \lambda_1, \lambda_2 \in \Lambda, \|\lambda_1 - \lambda_2\|_L^* \rightarrow 0.$$

Combining this with (4.97), we conclude that

$$\sup_{0 \leq t \leq T} \Delta(t, a, \lambda_1, \lambda_2) \rightarrow 0 \quad \text{as } \lambda_1, \lambda_2 \in \Lambda, \|\lambda_1 - \lambda_2\|_L^* \rightarrow 0, \text{dist}_X(a, \hat{a}) \rightarrow 0.$$

This convergence and inequality (4.108) imply the required result.

We now prove (4.108). Let  $\mu_a$  be a stationary measure for (3.84) with  $\eta = \eta_a$ . By Exercise 3.3.8, such a measure is unique for  $\text{dist}_X(a, \hat{a}) \leq \delta \ll 1$  and satisfies the inequality

$$\sup_{a \in B_X(\hat{a}, \delta)} \sup_{\lambda \in \Lambda} \|\mathfrak{P}_t^*(a)\lambda - \mu_a\|_L^* \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (4.109)$$

We write

$$\Delta(t, a, \lambda_1, \lambda_2) \leq \|\mathfrak{P}_t^*(a)\lambda_1 - \mu_a\|_L^* + \|\mathfrak{P}_t^*(\hat{a})\lambda_2 - \mu_{\hat{a}}\|_L^* + \|\mu_a - \mu_{\hat{a}}\|_L^*.$$

In view of (4.109), the first two terms on the right-hand side of this inequality go to zero uniformly in  $\lambda_1, \lambda_2 \in \Lambda$  as  $t \rightarrow \infty$ . Furthermore, by Theorem 4.3.1, the third term vanishes as  $a \rightarrow \hat{a}$ . These two observations immediately imply (4.108).  $\square$

### 4.3.2 Universality of white noise perturbations

In this section, we consider the Navier–Stokes system

$$\dot{u} + \nu Lu + B(u) = \eta, \tag{4.110}$$

where the space variable belongs to a bounded domain and the right-hand side  $\eta = \eta_\varepsilon$  is a high-frequency kick force of the form (2.69), i.e.,

$$\eta_\varepsilon(t, x) = \sqrt{\varepsilon} \sum_{k=1}^{\infty} \eta_k(x) \delta(t - k\varepsilon).$$

We assume that the random variables  $\eta_k$  satisfy the hypotheses of Theorem 3.2.9. In this case, if  $b_j \neq 0$  for all  $j \geq 1$ , then for any  $\varepsilon > 0$  the RDS associated with the equation

$$u_k = S_\varepsilon(u_{k-1}) + \sqrt{\varepsilon} \eta_k, \quad k \geq 1, \tag{4.111}$$

has a unique stationary distribution  $\mu_\varepsilon$ . On the other hand, if  $\eta$  is a spatially regular white noise of the form (2.66), then the Navier–Stokes system (4.110) also has a unique stationary distribution  $\mu$ . The following theorem shows that, under a mild regularity assumption, the family  $\{\mu_\varepsilon\}$  converges to  $\mu$  in the weak topology.

**Theorem 4.3.3.** *Let us assume that the above hypotheses are fulfilled and the coefficients  $b_j$  are such that  $\mathfrak{B}_1 = \sum_j \alpha_j b_j^2 < \infty$ . Then  $\mu_\varepsilon \rightarrow \mu$  as  $\varepsilon \rightarrow 0^+$  in the weak topology of  $\mathcal{P}(H)$ .*

*Proof.* We essentially repeat the scheme used in the proof of Theorem 4.3.1. Let  $P_t^\varepsilon(u, \Gamma)$  (where  $t \in \varepsilon\mathbb{Z}_+$ ,  $u \in H$ , and  $\Gamma \in \mathcal{B}(H)$ ) be the transition function for the Markov family defined by (4.111) and let  $\mathfrak{P}_t(\varepsilon)$  and  $\mathfrak{P}_t^*(\varepsilon)$  be the corresponding Markov operators; cf. Section 2.4.3. In this case, we can write (cf. (4.95))

$$\mu_\varepsilon - \mu = (\mathfrak{P}_t^*(\varepsilon)\mu_\varepsilon - \mathfrak{P}_t^*\mu_\varepsilon) + (\mathfrak{P}_t^*\mu_\varepsilon - \mu), \quad t \in \varepsilon\mathbb{Z}_+, \tag{4.112}$$

where  $\mathfrak{P}_t$  and  $\mathfrak{P}_t^*$  stand for the Markov semigroups associated with Eq. (4.110) in which  $\eta$  is given by (2.66). We wish to show that the right-hand side of (4.112) goes to zero in the dual-Lipschitz distance as  $t \rightarrow +\infty$  and  $\varepsilon \rightarrow 0^+$ . By Exercise 2.4.16, for any  $T > 0$  and any subset  $\Lambda \subset \mathcal{P}(H)$  satisfying the condition

$$\sup_{\lambda \in \Lambda} \int_H |u|_2^2 \lambda(du) < \infty, \tag{4.113}$$

we have (cf. (4.97))

$$\sup_{0 \leq t \leq T} \sup_{\lambda \in \Lambda} \|\mathfrak{P}_t^*(\varepsilon)\lambda - \mathfrak{P}_t^*\lambda\|_L^* \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0^+, \tag{4.114}$$

where the Markov semigroup  $\mathfrak{P}_t^*(\varepsilon)$  is extended to  $\mathbb{R}_+$  in a natural way; cf. Section 2.4.3. Furthermore, by Theorem 3.5.2, for any family  $\Lambda \subset \mathcal{P}(H)$  satisfying (4.113), we have

$$\sup_{\lambda \in \Lambda} \|\mathfrak{P}_t^*\lambda - \mu\|_L^* \rightarrow 0 \quad \text{as } t \rightarrow \infty. \tag{4.115}$$

If we prove that the family of measures  $\Lambda = \{\mu_\varepsilon\}$  satisfies (4.113), then combining (4.112) with relations (4.114) and (4.115), we see that  $\|\mu - \mu_\varepsilon\|_L^* \rightarrow 0$  as  $\varepsilon \rightarrow 0^+$ . Thus, it remains to show that (4.113) holds for the family  $\Lambda = \{\mu_\varepsilon\}$ .

In view of (2.51), we have  $|S_\varepsilon(u)|_2 \leq e^{-\varepsilon}|u|_2$  for  $u \in H$ . Since  $\eta_k$  and  $u_k$  are independent, it follows that

$$\mathbb{E}|u_k|_2^2 = \mathbb{E}|S_\varepsilon(u_{k-1})|_2^2 + \varepsilon \mathbb{E}|\eta_k|_2^2 \leq e^{-2\varepsilon} \mathbb{E}|u_{k-1}|_2^2 + \varepsilon \mathfrak{B},$$

where we used the relations

$$\mathbb{E}\eta_k = 0, \quad \mathbb{E}|\eta_k|_2^2 = \sum_{j=1}^{\infty} b_j^2 \mathbb{E}\xi_{jk}^2 = \sum_{j=1}^{\infty} b_j^2 = \mathfrak{B}.$$

Iterating the above inequality, we derive

$$\mathbb{E}|u_k|_2^2 \leq e^{-2\varepsilon k}|u_0|_2^2 + \frac{\varepsilon \mathfrak{B}}{1 - e^{-2\varepsilon}} \leq e^{-2\varepsilon k}|u_0|_2^2 + \frac{3\mathfrak{B}}{2} \quad \text{for all } k \geq 0.$$

The argument used in the proof of Theorem 2.5.3 implies now the required inequality (4.113).  $\square$

In contrast to Section 4.3.1, we cannot derive from Theorem 2.5.3 a result similar to Corollary 4.3.2 since we miss an analogue of (4.109). In other words, we cannot prove that the convergence  $\mathcal{D}(u_\varepsilon(t; v)) \rightarrow \mu_\varepsilon$  is uniform with respect to the parameter  $\varepsilon > 0$ .

## 4.4 Relevance of the results for physics

The Strong Law of Large Numbers proved in this chapter shows that, for the 2D turbulence driven by a non-degenerate random force, *the time average of an observable equals its ensemble average*. This equality is postulated in the theory of turbulence; e.g., see [Bat82], p. 17, or [Fri95], p. 58. It is important for numerical and experimental simulations, since it allows to calculate various averaged characteristics of a turbulent flow by running one experiment for a long time, rather than by making a large number of costly independent experiments in the stationary regime. The Law of Iterated Logarithm shows that the rate of convergence to the time-average is essentially the same as for the sum of independent random variables.

The Central Limit Theorem, in particular, justifies the widely accepted belief that, in the turbulent regime, *in large time-scales the velocity of the fluid has an approximately normal distribution*; see [Bat82], p. 174. For instance, an experiment carried out by A. A. Townsend in the middle of 20<sup>th</sup> century showed that the probability density function of the velocity of a turbulent fluid at an arbitrary point is very close to the Gaussian density; see [Bat82], p. 169. We believe that this observation is explained by the CLT, since any mechanical device measures not the instantaneous velocity, but its average over some time interval.

The random attractors and their deterministic counterparts are well-established tools for studying 2D Navier–Stokes equations. The fact that they have finite dimension for any positive value of the viscosity implies that large-time asymptotics of fluid motion, even in the turbulent regime, can be described by finite-dimensional dynamical systems. Theorem 4.2.19 shows that the random attractor has an additional natural structure—the probability measures  $\mu_\omega$ , which form a disintegration of the unique invariant measure.

Results of Section 4.3.1 indicate that, for statistical hydrodynamics, in striking difference with its deterministic analogue, the distribution of a velocity field  $u(t, \cdot)$  satisfying (0.1) continuously depends on parameters of the random force  $f$  uniformly in time  $t$ .

Finally, our results in Section 4.3.2 establish that, in certain sense, *white forces are universal* for the randomly forced 2D Navier–Stokes system (0.1). Namely, there we prove that if the system is perturbed by the force

$$f_\varepsilon(t, x) = \sqrt{\varepsilon} \eta(\varepsilon^{-1}t, x), \quad 0 < \varepsilon \ll 1, \quad (4.116)$$

where  $\eta$  is a non-degenerate kick-process with zero mean value, then its unique stationary measure converges, as  $\varepsilon \rightarrow 0$ , to the unique stationary measure of a white-forced system. Certainly this is a general phenomenon which holds true for many other fast oscillating random forces of the form (4.116), provided that the random field  $\eta(t, x)$  satisfies assumptions (a) and (b) from the Introduction.

## Notes and comments

The LLN, LIL, and CLT are well understood for independent random variables and for processes with strong mixing properties. Roughly speaking, these results hold as soon as the strong mixing coefficient decays to zero sufficiently fast; see the books [Has80, JS87, MT93, Rio00]. In the context of randomly forced PDE's, this condition can be satisfied only if the noise is rough with respect to the space variables; cf. Example 1.3 in [Shi06c]. In the context of randomly forced Navier–Stokes system, the SLNN and CLT were established by Kuksin [Kuk02a], using a coupling argument and some general limit theorems for stationary processes. The martingale approximation introduced by Gordin [Gor69] is a powerful tool for studying limit theorem for stochastic processes. It was used by Shirikyan [Shi06c] to derive the rates of convergence in the SLLN and CLT. The LIL for Navier–Stokes equations was announced by Denisov [Den04], however, a proof has never appeared. The SLLN and CLT discussed in this monograph can be derived from more general results of the book [DDL<sup>+</sup>07], while the LIL can be obtained as a consequence of the invariance principles established in the recent papers [DM10] and [BMS11].

A systematic study of random dynamical systems can be found in the book of L. Arnold [Arn98]. Random attractors are a natural extension of attractors for non-autonomous equations (cf. [CV02]). They were constructed in the papers of Crauel, Debussche, and Flandoli [CF94, CDF97], and their finite-dimensionality for various stochastic PDE's was proved by Debussche [Deb97,

[Deb98](#)], Langa, Robinson [[LR06](#)], and many others. Theorem [4.2.9](#) was established at various levels of generality by Ledrappier [[Led86](#)], Le Jan [[Le 87](#)], and Crauel [[Cra91](#)]. Theorems [4.2.16](#), [4.2.17](#), and [4.2.19](#) were proved by Kuksin and Shirikyan [[KS04a](#)]

The dependence of stationary measures on various parameters in the context of randomly forced PDE's was studied by Kuksin, Shirikyan [[KS03](#)], Chueshov, Kuksin [[CK08a](#)], and Hairer, Mattingly [[HM08](#)]. The presentation here follows the paper [[KS03](#)].

# Chapter 5

## Inviscid limit

In this chapter, we consider the Navier–Stokes system on the 2D torus  $\mathbb{T}^2$ , perturbed by a spatially regular white noise (with trivial deterministic part) whose amplitude is proportional to the square root of the viscosity:

$$\partial_t u + \nu Lu + B(u) = \sqrt{\nu} \eta(t, x), \quad (5.1)$$

$$\eta(t, x) = \frac{\partial}{\partial t} \zeta(t, x), \quad \zeta(t, x) = \sum_{s \in \mathbb{Z}_0^2} b_s \beta_s(t) e_s(x), \quad (5.2)$$

Here  $\{e_s\}$  is the basis defined in (2.29),  $\{\beta_s\}$  is a sequence of independent standard Brownian motions, and  $b_s \geq 0$  are some constants decaying to zero sufficiently fast. According to Theorem 3.3.1, if all the constants  $b_s$  are non-zero, then Eq. (5.1) has a unique stationary distribution  $\mu_\nu$ . In this chapter, we show that the family  $\{\mu_\nu\}$  is tight and investigate properties of the limiting measures. In particular, it will be proved that any such measure is the law of a stationary process concentrated on solutions of the homogeneous Euler equation. Some a priori estimates and a non-degeneracy property will also be obtained. The results of this chapter (with properly modified constants) remain true for the Navier–Stokes system on a non-standard torus  $\mathbb{R}^2/(a\mathbb{Z} \oplus b\mathbb{Z})$ ,  $a, b > 0$ .

### 5.1 Balance relations

#### 5.1.1 Energy and enstrophy

As before, for  $k \in \mathbb{Z}_+$  we denote

$$\mathfrak{B}_k = \sum_{s \in \mathbb{Z}_0} \alpha_s^k b_s^2 \leq \infty,$$

where  $\alpha_s = |s|^2$  stands for the eigenvalue associated with  $e_s$ . Note that  $\mathfrak{B}_0 = \mathfrak{B}$ .

Let us recall some results established in Section 2.5.2. Theorem 2.5.5 implies that, if  $\mathfrak{B}_0 < \infty$  and  $u_\nu(t, x)$  a stationary solution of problem (5.1), (5.2), then

$$\mathbb{E} \langle Lu_\nu, u_\nu \rangle = \frac{1}{2} \mathfrak{B}_0, \quad (5.3)$$

$$\mathbb{E} \exp(\varkappa_0 |u_\nu|_2^2) \leq C_0, \quad (5.4)$$

where  $\varkappa_0$  and  $C_0$  are positive constants not depending on  $\nu > 0$ . Furthermore, by Exercise 2.5.7, if  $\mathfrak{B}_1 < \infty$ , then

$$\mathbb{E} |Lu_\nu|_2^2 = \frac{1}{2} \mathfrak{B}_1, \quad (5.5)$$

$$\mathbb{E} \exp(\varkappa_1 \|u_\nu\|_1^2) \leq C_1, \quad (5.6)$$

where the positive constants  $\varkappa_1$  and  $C_1$  are also independent of  $\nu > 0$ . Note that relations (5.3) and (5.4) are true for any bounded domain with smooth boundary, whereas (5.5) and (5.6) are valid only for periodic boundary conditions, because their proof uses the crucial relation  $\langle B(u), u \rangle = 0$ .

A simple corollary of the above estimates is the inequality

$$\frac{\mathfrak{B}_0^2}{2\mathfrak{B}_1} \leq \mathbb{E} \int_{\mathbb{T}^2} |u_\nu(t, x)|^2 dx \leq \frac{\mathfrak{B}_0}{2} \quad \text{for all } \nu > 0. \quad (5.7)$$

Indeed, the right-hand estimate is a straightforward consequence of (5.3) and the obvious inequality  $|v|_2^2 \leq \langle Lv, v \rangle$ . To prove the left-hand estimate, note that, by the interpolation inequality for Sobolev spaces (see Property 1.1.4), we have

$$\langle Lv, v \rangle \leq |v|_2 |Lv|_2 \quad \text{for any } v \in V^2.$$

It follows that

$$\mathbb{E} \langle Lu_\nu, u_\nu \rangle \leq (\mathbb{E} |u_\nu|_2^2)^{1/2} (\mathbb{E} |Lu_\nu|_2^2)^{1/2}.$$

Recalling relations (5.3) and (5.5), we arrive at the left-hand inequality in (5.7).

### 5.1.2 Balance relations

Equalities (5.3) and (5.5) are related to the fact that the energy and enstrophy are integrals of motion for the 2D Euler equations on the torus. It is well known that the integral over  $\mathbb{T}^2$  of any function of vorticity also is preserved under the dynamics of the Euler system. It turns out that if the right-hand side of (5.1) is sufficiently smooth in  $x$ , then to each function of vorticity there corresponds a *balance relation* similar to (5.3) and (5.5). To formulate the corresponding result, we first introduce some notation.

Let us assume that  $\mathfrak{B}_2 < \infty$ . Then, by Exercise 2.5.8 with  $k = 2$ , almost every trajectory of a stationary solution  $u_\nu(t, x)$  for (5.1) belongs to the space  $L_{\text{loc}}^2(\mathbb{R}_+; V^3) \cap C(\mathbb{R}_+; V^2)$ . Applying the operator curl to (5.1) and using



the relation  $\text{curl}(\nabla q) = 0$ , we obtain the following equation<sup>1</sup> for the stationary process  $v_\nu = \text{curl } u_\nu$  whose almost every trajectory belongs to the space  $L^2_{\text{loc}}(\mathbb{R}_+; H^2) \cap C(\mathbb{R}_+; H^1)$ :

$$\partial_t v - \nu \Delta v + \langle u, \nabla \rangle v = \sqrt{\nu} \xi. \tag{5.8}$$

Here  $\xi$  is an  $H^1$ -smooth white noise given by the relation

$$\xi(t, x) = \text{curl} \eta(t, x) = \frac{\partial}{\partial t} \sum_{s \in \mathbb{Z}_0^2} b_s \beta_s(t) \varphi_s(x), \tag{5.9}$$

where we set

$$\varphi_s(x) = \frac{|s|}{\sqrt{2\pi}} \begin{cases} \cos\langle s, x \rangle, & s \in \mathbb{Z}_+^2, \\ -\sin\langle s, x \rangle, & s \in \mathbb{Z}_-^2. \end{cases} \tag{5.10}$$

**Theorem 5.1.1.** *Let  $\mathfrak{B}_2 < \infty$  and let  $g(r)$  be a continuous function satisfying the inequality*

$$|g(r)| \leq C(1 + |r|)^l \quad \text{for } r \in \mathbb{R}, \tag{5.11}$$

where  $C$  and  $l$  are some constants. Then for all  $\nu > 0$  and  $t \geq 0$  we have

$$\mathbb{E} \int_{\mathbb{T}^2} g(v_\nu) |\nabla v_\nu|^2 dx = \frac{1}{2} \sum_{s \in \mathbb{Z}_0^2} b_s^2 \mathbb{E} \int_{\mathbb{T}^2} g(v_\nu) \varphi_s^2 dx. \tag{5.12}$$

The meaning of the left-hand side of inequality (5.12) needs to be clarified, because the trajectories of  $v_\nu$  are continuous only as functions of time with range in  $H^1$ , and the integral in  $x$  does not necessarily converges. However, it follows from Exercise 2.5.8 that  $\mathbb{E} \|v_\nu(t)\|_2^2 < \infty$  for almost every  $t$ , and the mean value of the integral makes sense for almost every  $t$ . Since  $v_\nu$  is stationary, the left-hand side of (5.12) is in fact independent of time. Readers willing to avoid this type of complications may assume that  $\mathfrak{B}_3 < \infty$ , in which case all the terms are well defined without taking the mean value.

*Proof of Theorem 5.1.1.* Let  $G$  be the second primitive of  $g$  vanishing at zero together with its first derivative, that is,  $G''(r) = g(r)$  and  $G(0) = G'(0) = 0$ . Consider a functional  $F : H^1 \rightarrow \mathbb{R}$  defined by

$$F(w) = \int_{\mathbb{T}^2} G(w(x)) dx. \tag{5.13}$$

It is straightforward to see that  $F$  satisfies the hypotheses of Theorem 7.7.5 for the triple  $(V, H, V^*) = (H^2, H^1, L^2)$ , and therefore we can apply Itô's formula (7.24) to the process  $F(v_\nu(t))$ . Omitting the subscript  $\nu$  for the simplicity of notations, we get

$$F(v(t \wedge \tau_n)) = F(v(0)) + \int_0^{t \wedge \tau_n} \left( A(\theta) d\theta + \sum_{s \in \mathbb{Z}_0^2} B_s(\theta) d\beta_s \right), \tag{5.14}$$

---

<sup>1</sup>Note that Eq. (5.8) is valid for any solution whose trajectories possess the above-mentioned regularity.

where  $\tau_n = \inf\{t \geq 0 : \|v(t)\|_1 > n\}$ ,

$$A(t) = \int_{\mathbb{T}^2} \left( G'(v)(\nu \Delta v - \langle u, \nabla \rangle v) + \frac{\nu}{2} \sum_{s \in \mathbb{Z}_0^2} b_s^2 g(v) \varphi_s^2 \right) dx,$$

$$B_s(t) = \sqrt{\nu} b_s \int_{\mathbb{T}^2} G'(v) \varphi_s dx.$$

Suppose we have shown that condition (7.36) is satisfied for the process  $F(v)$ . Then, by Corollary 7.7.6, relation (5.14) implies that

$$F(v(t)) = F(v(0)) + \int_0^t A(\theta) d\theta + M_t, \quad (5.15)$$

where  $M_t$  stands for the corresponding stochastic integral. Taking the mean value of both sides of (5.15), using the stationarity of  $v$ , and recalling the definition of  $A$ , we get  $\mathbb{E}A(t) = 0$  for any  $t \geq 0$ . Noting that

$$\int_{\mathbb{T}^2} G'(v)(\nu \Delta v - \langle u, \nabla \rangle v) dx = -\nu \int_{\mathbb{T}^2} g(v) |\nabla v|^2 dx$$

for any functions  $v \in H^2$  and  $u \in V^2$ , we arrive at the required relation (5.12).

To complete the proof of the theorem, it remains to prove that we can apply Corollary 7.7.6. That is, we need to check that

$$\sum_{s \in \mathbb{Z}_0^2} \mathbb{E} \int_0^t |B_s(\theta)|^2 d\theta < \infty \quad \text{for any } t > 0. \quad (5.16)$$

In view of (5.10), (5.11), and (5.6), for any  $t > 0$  we have

$$\begin{aligned} \sum_{s \in \mathbb{Z}_0^2} \mathbb{E} \int_0^t |B_s(\theta)|^2 d\theta &= \nu \sum_{s \in \mathbb{Z}_0^2} b_s^2 \mathbb{E} \int_0^t \left| \int_{\mathbb{T}^2} G'(v) \varphi_s \right|^2 d\theta \\ &\leq C_1 \nu \sum_{s \in \mathbb{Z}_0^2} b_s^2 |s|^2 \int_0^t \mathbb{E} |v(\theta)|_{2(l+1)}^{2(l+1)} d\theta \leq C_2 \mathfrak{B}_1 t \nu, \end{aligned}$$

where  $C_1$  and  $C_2$  are some constants not depending on  $\nu$  and  $\{b_s\}$ . The proof is complete.  $\square$

*Exercise 5.1.2.* Calculate the first and second derivatives of the functional  $F$  defined by (5.13) and verify that the hypotheses of Theorem 7.7.5 are satisfied for  $F$ . *Hint:* Use the fact that the Sobolev space  $H^1$  is continuously embedded in  $L^p$  for any finite  $p \geq 1$ .

We now consider a particular case in which the stationary measure is unique, and the law of the random perturbation is invariant under translations and reflections. Namely, as was proved in Corollary 3.3.4, if the coefficients  $b_s$  are all non-zero and  $b_s \equiv b_{-s}$ , then the unique stationary measure  $\mu_\nu$  is invariant under translations and reflections. This property enables one to draw some further conclusions on the stationary distributions. Namely, we have the following result.

**Proposition 5.1.3.** *In addition to the hypotheses of Theorem 5.1.1, suppose that  $\mathfrak{B}_4 < \infty$  and*

$$b_s = b_{-s} \neq 0 \quad \text{for all } s \in \mathbb{Z}_0^2.$$

*Let  $\mu_\nu$  be the unique stationary measure of (5.1), let  $u_\nu$  be a corresponding stationary solution, and let  $v_\nu = \text{curl } u_\nu$ . Then for all  $\nu > 0$ ,  $t \geq 0$ , and  $x \in \mathbb{T}^2$  we have<sup>2</sup>*

$$\mathbb{E}(g(v_\nu(t, x))|\nabla v_\nu(t, x)|^2) = \frac{1}{2}(2\pi)^{-2}\mathfrak{B}_1 \mathbb{E}(g(v_\nu(t, x))). \quad (5.17)$$

Equalities (5.17) (with various functions  $g$ ) are called *balance relations* for a stationary solution  $v_\nu$ .

*Proof of Proposition 5.1.3.* Since  $b_s = b_{-s}$ , the right-hand side of (5.12) is equal to

$$\frac{1}{4}\mathbb{E} \sum_{s \in \mathbb{Z}_0^2} b_s^2 \int_{\mathbb{T}^2} g(v_\nu)(\varphi_s^2 + \varphi_{-s}^2) dx = \frac{1}{2}(2\pi)^{-2}\mathfrak{B}_1 \mathbb{E} \int_{\mathbb{T}^2} g(v_\nu) dx,$$

where we used the relation  $\varphi_s^2 + \varphi_{-s}^2 = \frac{|s|^2}{2\pi^2}$ . Substituting this into (5.12), we derive

$$\int_{\mathbb{T}^2} \mathbb{E} \left( g(v_\nu)(|\nabla v_\nu|^2 - \frac{1}{2}(2\pi)^{-2}\mathfrak{B}_1) \right) dx = 0.$$

Since the law of  $v_\nu$  is invariant under translations in  $x$ , the integrand is independent of  $x$  and vanishes. The proof is complete.  $\square$

Relation (5.3) and the space homogeneity of the process  $v_\nu(t, x)$  imply that  $\mathbb{E}|\nabla v_\nu(t, x)|^2 = \frac{1}{2}\mathfrak{B}_0(2\pi)^{-2}$ . So (5.17) means that the random variables  $g(v_\nu(t, x))$  and  $|\nabla v_\nu(t, x)|^2$  are uncorrelated: the expectation of their product equals the product of their expectations. Jointly the balance relations (5.17) (i.e., the assertion of Proposition 5.1.3) may be reformulated in terms of the conditional expectation with respect to the  $\sigma$ -algebra  $\mathcal{F}_{v_\nu(t, x)}$  generated by  $v_\nu(t, x)$ .

**Corollary 5.1.4.** *Under the hypotheses of Proposition 5.1.3, for all  $\nu > 0$ ,  $t \geq 0$ , and  $x \in \mathbb{T}^2$  we have*

$$\mathbb{E}(|\nabla v_\nu(t, x)|^2 | \mathcal{F}_{v_\nu(t, x)}) = \frac{1}{2}(2\pi)^{-2}\mathfrak{B}_1.$$

*Proof.* We first note that, in view of Exercise 2.5.8, almost all trajectories of  $v_\nu$  are continuous in time with range in  $H^3$ , whence it follows that, with probability 1, the expression  $|\nabla v_\nu(t, x)|$  is well defined for any  $t \geq 0$  and  $x \in \mathbb{T}^2$ . The required assertion is now a straightforward consequence of (5.17) and the definition of the conditional expectation.  $\square$

<sup>2</sup>Note that the left- and right-hand side of (5.17) are well defined, because almost every trajectory of  $v_\nu$  is a continuous function of time with range in  $H^3$ , and the corresponding norm has finite moments; see Exercise 2.5.8.

Other way round, the assertion of the corollary implies that

$$\begin{aligned}\mathbb{E}(g(v_\nu(t, x))|\nabla v_\nu(t, x)|^2) &= \mathbb{E}(g(v_\nu(t, x))|\nabla v_\nu(t, x)|^2 | \mathcal{F}_{v_\nu(t, x)}) \\ &= \frac{1}{2}(2\pi)^{-2}\mathfrak{B}_1 \mathbb{E}(g(v_\nu(t, x))).\end{aligned}$$

Thus, Corollary 5.1.4 is equivalent to the balance relations.

The following exercise gives yet another form of the balance relations (5.17).

*Exercise 5.1.5.* In addition to the hypotheses of Proposition 5.1.3, assume that  $\mathfrak{B}_6 < \infty$ . Prove that the following properties hold for any fixed  $t \geq 0$ .

- (i) For  $\tau \in \mathbb{R}$  and  $\omega \in \Omega$ , define  $\Gamma_t(\tau, \omega) = \{x \in \mathbb{T}^2 : v_\nu(t, x) = \tau\}$ . Then for almost all  $\omega \in \Omega$  and  $\tau \in \mathbb{R}$ , the set  $\Gamma_t(\tau, \omega)$  is a  $C^3$ -smooth curve in  $\mathbb{T}^2$ .
- (ii) Let  $d\ell$  be the length element on  $\Gamma_t(\tau, \omega)$  (whenever the latter is a smooth curve). Then for almost every  $\tau \in \mathbb{R}$ , we have the following *co-area form of the balance relations*:

$$\mathbb{E} \int_{\Gamma_t(\tau, \omega)} |\nabla v_\nu| d\ell = \frac{1}{2}(2\pi)^{-1}\mathfrak{B}_1 \mathbb{E} \int_{\Gamma_t(\tau, \omega)} |\nabla v_\nu|^{-1} d\ell.$$

*Hint:* Integrate (5.17) over  $dx$  and perform the co-area change of variables  $dx = |\nabla v_\nu|^{-1} d\tau d\ell$  in the integral. (We refer the reader to [Kuk06b] for a full proof of this result.)

### 5.1.3 Pointwise exponential estimates

Inequality (5.6), which is true for any stationary solution  $u_\nu$  of (5.1), implies that

$$\mathbb{P}\{\|u_\nu(t)\|_1 \geq R\} \leq C_1 e^{-\kappa_1 R^2} \quad \text{for all } R > 0.$$

Thus, large values of the enstrophy are very unlikely in a stationary regime. The aim of this section is to show that if the law of the random perturbation  $\eta$  is invariant under translations, then some similar estimates hold pointwise for  $u_\nu$  and  $\nabla u_\nu$ . Namely, we have the following result.

**Theorem 5.1.6.** *Under the hypotheses of Proposition 5.1.3, there are positive constants  $\sigma$  and  $K$ , depending only on  $\mathfrak{B}_1$ , such that, for all  $t \geq 0$ ,  $x \in \mathbb{T}^2$ ,  $\nu > 0$ , we have*

$$\mathbb{E} e^{\sigma|v_\nu(t, x)|} + \mathbb{E} e^{\sigma|u_\nu(t, x)|} + \mathbb{E} e^{\sigma|\nabla u_\nu(t, x)|^{1/2}} \leq K. \quad (5.18)$$

*Proof.* It suffices to estimate each term on the left-hand side of (5.18) by a constant depending only on  $\mathfrak{B}_1$ . We begin with the first term.

Integrating over  $x \in \mathbb{T}^2$  relation (5.17) in which  $g(r) = |r|^{2p}$  with  $p > 0$  and applying Hölder's inequality, we derive

$$\begin{aligned}\mathbb{E} \int_{\mathbb{T}^2} |v_\nu|^{2p} |\nabla v_\nu|^2 dx &= C_1 \mathfrak{B}_1 \mathbb{E} \int_{\mathbb{T}^2} |v_\nu|^{2p} dx \\ &\leq C_1 (2\pi)^{2/(p+1)} \left( \mathbb{E} \int_{\mathbb{T}^2} |v_\nu|^{2(p+1)} dx \right)^{\frac{p}{p+1}}.\end{aligned} \quad (5.19)$$

On the other hand, in view of generalised Poincaré's inequality (7.75), we have

$$\int_{\mathbb{T}^2} |v_\nu|^{2p} |\nabla v_\nu|^2 dx \geq C_2(p+1)^{-2} \int_{\mathbb{T}^2} |v_\nu|^{2(p+1)} dx.$$

Taking the mean value and substituting the resulting inequality into (5.19), we obtain

$$\mathbb{E} \int_{\mathbb{T}^2} |v_\nu|^{2(p+1)} dx \leq (C_3 \mathfrak{B}_1)^{p+1} (p+1)^{2(p+1)} \quad \text{for any } p > 0. \quad (5.20)$$

Since the law of  $v_\nu(t, x)$  is invariant under the translations, we conclude that

$$\mathbb{E} |v_\nu(t, x)|^m \leq (C_4 \mathfrak{B}_1)^{m/2} m^m \quad \text{for all integers } m \geq 0,$$

where we set  $0^0 = 1$ . It follows that if  $0 < \sigma \leq (e\sqrt{C_4 \mathfrak{B}_1})^{-1}$ , then

$$\begin{aligned} \mathbb{E} e^{\sigma |v_\nu(t, x)|} &= \sum_{m=0}^{\infty} \frac{\sigma^m}{m!} \mathbb{E} |v_\nu(t, x)|^m \leq \sum_{m=0}^{\infty} \frac{\sigma^m}{m!} (C_4 \mathfrak{B}_1)^{m/2} m^m \\ &\leq (1 - \sigma e\sqrt{C_4 \mathfrak{B}_1})^{-1}, \end{aligned} \quad (5.21)$$

where we used the well-known inequality  $m! \geq (m/e)^m$ . We have thus proved the required estimate for the first term on the left-hand side of (5.18).

The derivation of an upper bound for the second and third terms is based on similar ideas, and we shall confine ourselves to the proof for the more difficult third term. Let us recall that the velocity field  $u_\nu$  can be recovered from the vorticity  $v_\nu$  by the relation  $u_\nu = \text{curl}(\Delta^{-1}v_\nu)$ , where  $\Delta^{-1}$  stands for the inverse of the Laplace operator on the space of the  $L^2$  function on  $\mathbb{T}^2$  with zero mean value. It follows that the components of the matrix  $\nabla u_\nu$  can be written in the form  $Rv_\nu = \gamma(\partial_i \partial_j \Delta^{-1})v_\nu$ , where  $\gamma = \pm 1$ . Now note that  $R$  is a singular integral operator on  $\mathbb{T}^2$ , or more precisely, a composition of two Riesz projections; see Chapters II and III in [Ste70] for definitions. Therefore, by the Calderón–Zygmund theorem (see Sections 2.2 and 6.2 in Chapter II of [Ste70]), we have

$$|Rv_\nu|_q \leq Cq |v_\nu|_q \quad \text{for any } q \in [2, \infty).$$

Combining this estimate with  $q = 2(p+1)$  and inequality (5.20), we obtain

$$\mathbb{E} \int_{\mathbb{T}^2} |\nabla u_\nu|^{2(p+1)} dx \leq (C_5 \mathfrak{B}_1)^{p+1} (p+1)^{4(p+1)} \quad \text{for any } p > 0.$$

Since the law of  $\nabla u_\nu(t, x)$  is invariant under translations, it follows that

$$\mathbb{E} |v_\nu(t, x)|^{m/2} \leq (C_6 \mathfrak{B}_1)^{m/4} m^m \quad \text{for all } m \geq 0.$$

The required estimate can now be obtained by exactly the same argument as above; see (5.21).  $\square$

*Exercise 5.1.7.* Prove that the middle term in (5.18) can be bound by the right-hand side. *Hint:* Repeat the scheme of the proof of Theorem 5.1.6 using the fact that  $|u_\nu|_p \leq C |v_\nu|_p$  for any  $p \geq 2$ , where  $C > 0$  does not depend on  $p$ ; the latter follows from the elliptic regularity for the Laplacian and an interpolation theorem for Sobolev spaces.

## 5.2 Limiting measures

### 5.2.1 Existence of accumulation points

Throughout this subsection, we shall assume that

$$\mathfrak{B}_1 < \infty.$$

Let  $\{u_\nu, \nu > 0\}$  be a family of stationary solutions for (5.1) and let  $\mu_\nu$  be the law of  $u_\nu(t, \cdot)$ . In view of the results described in Section 5.1.1, we have

$$\mathbb{E} \|u_\nu(t, \cdot)\|_2^2 = \int_H \|v\|_2^2 \mu_\nu(dv) \leq C \quad \text{for all } \nu > 0. \quad (5.22)$$

Since the embedding  $H^2 \subset H^{2-\varepsilon}$  is compact for any  $\varepsilon > 0$ , Prokhorov's theorem implies that the family  $\{\mu_\nu\}$  is tight in  $H^{2-\varepsilon}$ , and therefore any countable subsequence  $\{\mu_{\nu_j}, \nu_j \rightarrow 0\}$  has a limiting point in the sense of weak convergence of measures on  $H^{2-\varepsilon}$ . Furthermore, if  $\mu$  is an accumulation point of the sequence  $\{\mu_{\nu_j}\}$ , then  $\mu$  is concentrated on  $H^2$ . Indeed, for any  $R > 0$  the ball  $B_{H^2}(R)$  is closed in  $H^{2-\varepsilon}$ , and  $\mu_{\nu_j}(B_{H^2}(R)) \geq 1 - CR^{-2}$  by (5.22) and Chebyshev's inequality. Therefore, by the portmanteau theorem, we have  $\mu(B_{H^2}(R)) \geq 1 - CR^{-2}$  for any  $R > 0$ , so  $\mu(H^2) = 1$ . Theorem 5.2.2 established below shows that the laws of the solutions  $u_{\nu_j}$  (regarded as elements of an appropriate space of trajectories) also converge to a limiting measure. We shall need some notations.

Let us denote by  $\mu_\nu$  the law of a stationary solution  $u_\nu$ . Since the trajectories of solutions of (5.1) almost surely belong to the space  $L_{\text{loc}}^2(\mathbb{R}_+; V)$ , we see that  $\mu_\nu$  is a probability measure on it.

*Exercise 5.2.1.* Show that if  $\mathfrak{B}_1 < \infty$ , then the measure  $\mu_\nu$  is concentrated on the space  $C(\mathbb{R}_+; V) \cap L_{\text{loc}}^2(\mathbb{R}_+; V^2)$ .

Given a Banach space  $X$ , a finite interval  $J \subset \mathbb{R}$ , and constants  $\alpha \in (0, 1)$  and  $p \geq 1$ , we define  $W^{\alpha,p}(J; X)$  as the space of curves  $f \in L^p(J; X)$  such that

$$\|f\|_{W^{\alpha,p}(J; X)} = \|f\|_{L^p(J; X)}^p + \int_J \int_J \frac{\|f(t) - f(s)\|_X^p}{|t - s|^{1+\alpha p}} ds dt < \infty.$$

The following theorem proves the tightness of the family  $\{\mu_\nu\}$  in an appropriate functional space and describes some properties of the limiting measures.

**Theorem 5.2.2.** *For any  $\varepsilon > 0$  the family of measures  $\{\mu_\nu\}$  is tight<sup>3</sup> in the space  $C(\mathbb{R}_+; H^{1-\varepsilon}) \cap L_{\text{loc}}^2(\mathbb{R}_+; H^{2-\varepsilon})$ . Moreover, if  $\mu$  is a limiting point for  $\{\mu_\nu\}$  and  $u(t, x)$  is a random process whose law coincides with  $\mu$ , then the following properties hold.*

<sup>3</sup>Recall that the space  $C(\mathbb{R}_+; V)$  is endowed with the topology of uniform convergence on bounded intervals of  $\mathbb{R}_+$ ; see Section 1.1.2. Similarly, the space  $L_{\text{loc}}^2(\mathbb{R}_+; H^s)$  is endowed with the topology generated by the norms of  $L^2(0, n; H^s)$ , where  $n \geq 1$  is any integer.

(i) Almost every realisation of  $u(t, x)$  belongs to the space

$$L^2_{\text{loc}}(\mathbb{R}_+; V^2) \cap W^{1,1}_{\text{loc}}(\mathbb{R}_+; V) \cap W^{1,\infty}_{\text{loc}}(\mathbb{R}_+; L^p) \quad (5.23)$$

for any  $p \in [1, 2)$  and satisfies the Euler equation

$$\dot{u} + B(u) = 0. \quad (5.24)$$

(ii) The process  $u(t, x)$  is stationary in time, and the functions  $|u(t)|_2$  and  $|\nabla u(t)|_2$  are time-independent random constants.

Before proving this result, we establish the following corollary.

**Corollary 5.2.3.** *Let  $\mu$  be a limiting point of the family  $\{\mu_\nu\}$  in the space  $C(\mathbb{R}_+; H^{1-\varepsilon})$  as  $\nu \rightarrow 0^+$ , where  $\varepsilon \in (0, 1)$ , and let  $\mu$  be the restriction of  $\mu$  at time  $t = 0$ . Then*

$$\int_H |u|_2^2 \mu(du) \geq \frac{\mathfrak{B}_0^2}{2\mathfrak{B}_1}, \quad (5.25)$$

$$\int_H \|u\|_1^2 \mu(du) = \frac{\mathfrak{B}_0}{2}, \quad (5.26)$$

$$\int_H |Lu|_2^2 \mu(du) \leq \frac{\mathfrak{B}_1}{2}, \quad (5.27)$$

$$\int_H \exp(\varkappa \|u\|_1^2) \mu(du) \leq C, \quad (5.28)$$

where  $\varkappa$  and  $C$  are some positive constants. If, in addition, the hypotheses of Proposition 5.1.3 are fulfilled, then for all  $x \in \mathbb{T}^2$  we have

$$\int_H (e^{\sigma|u(x)|} + e^{\sigma|\nabla u(x)|^{1/2}}) \mu(du) \leq K \quad (5.29)$$

with suitable positive constants  $\sigma$  and  $K$ . Moreover, in this case the measure  $\mu$  is invariant under translations and reflections of  $\mathbb{R}^2$ .

*Proof.* Inequality (5.25) is a straightforward consequence of (5.26), (5.27) and the interpolation inequality for Sobolev spaces; cf. proof of (5.7). The derivation of (5.26) – (5.29) is based on passing to the limit in similar estimates for stationary solutions of the Navier–Stokes system. The last assertion of the theorem follows from similar properties of the measures  $\mu_\nu$ . Therefore, we shall confine ourselves to the proof of (5.27), (5.28) and (5.26).

Since  $\mu$  is a limiting point of the family  $\{\mu_\nu\}$ , there is a sequence  $\nu_j \rightarrow 0^+$  such that  $\mu_{\nu_j} \rightarrow \mu$  in the sense of weak convergence of measures on  $C(\mathbb{R}_+; H^{1-\varepsilon})$ . It follows that  $\mu_{\nu_j} \rightarrow \mu$  as  $j \rightarrow \infty$ , where  $\mu_{\nu_j}$  stands for the restriction of  $\mu_{\nu_j}$  at time  $t = 0$ , and the convergence holds weakly in the space of Borel measures on  $H^{1-\varepsilon}$ . Now (5.27) easily follows from Lemma 1.2.17, where  $\pi_n$  is the orthogonal projection  $P_n : H \rightarrow H$  to the vector span of the first  $n$  eigenfunctions of the Stokes operator  $L$ ; see Section 2.1.5.

Let us prove (5.28). It follows from (5.27) that the family  $\{\mu_{\nu_j}\}$  is tight in  $H^{2-\delta}$  for any  $\delta > 0$ , and by Prokhorov's theorem, there is no loss of generality in assuming that  $\mu_{\nu_j} \rightarrow \mu$  in the weak topology of  $H^{2-\delta}$ . Using (5.6) and applying again Lemma 1.2.17, we derive inequality (5.28).

We now prove (5.26). It follows from (5.3) that, for any  $R > 0$ , we have

$$\int_{B_V(R)} \|u\|_1^2 \mu_{\nu_j}(du) + \int_{B_V^c(R)} \|u\|_1^2 \mu_{\nu_j}(du) = \frac{\mathfrak{B}_0}{2}.$$

Using (2.157), for any  $\delta > 0$  we can choose  $R = R_\delta > 0$  so large that the second term on the left-hand side is smaller than  $\delta$  uniformly in  $j$ . On the other hand, for any fixed  $R > 0$ , the first term converges to a similar integral for  $\mu$ . These two observations combined with Fatou's lemma immediately imply that

$$\frac{\mathfrak{B}_0}{2} - \varepsilon \leq \int_{B_V(R_\varepsilon)} \|u\|_1^2 \mu_{\nu_j}(du) \leq \int_H \|u\|_1^2 \mu_{\nu_j}(du) \leq \frac{\mathfrak{B}_0}{2}.$$

Since  $\varepsilon > 0$  is arbitrary, we obtain the required inequality.  $\square$

*Exercise 5.2.4.* Prove inequality (5.29). *Hint:* Use the fact that the functionals taking  $u \in V$  to  $|u(x)|$  and  $|\nabla u(x)|$  are measurable on  $V$ , and apply Fatou's lemma.

*Proof of Theorem 5.2.2.* We first prove the tightness of the family  $\{\mu_\nu\}$ . To this end, it suffices to show that its restriction to any finite interval  $J = [0, n]$  is tight in  $C(J; H^{1-\varepsilon}) \cap L^2(J; H^{2-\varepsilon})$ . To simplify the formulas, we assume that  $J = [0, 1]$ , and with a slight abuse of notation, denote the restriction of measures to  $J$  by the same symbol.

We fix some constants  $\varepsilon > 0$  and  $\alpha \in (1/4, 1/2)$  and introduce the functional spaces<sup>4</sup>

$$\begin{aligned} \mathcal{X} &= L^2(J; V^2) \cap (H^1(J; H) + W^{\alpha,4}(J; V)) = \mathcal{X}_1 + \mathcal{X}_2, \\ \mathcal{Y} &= L^2(J; H^{2-\varepsilon}) \cap C(J; H^{1-\varepsilon}) \end{aligned}$$

where we set

$$\mathcal{X}_1 = L^2(J; V^2) \cap H^1(J; H), \quad \mathcal{X}_2 = L^2(J; V^2) \cap W^{\alpha,4}(J; V).$$

Let us note that we have a compact embedding  $\mathcal{X} \subset \mathcal{Y}$ . Indeed, it follows from Theorem 5.1 and 5.2 in [Lio69, Chapter I] that  $\mathcal{X}_1$  and  $\mathcal{X}_2$  are compactly embedded in  $L^2(J; H^{2-\varepsilon})$ . Furthermore, in view of Theorem 3.1 in [LM72], we have a continuous embedding  $\mathcal{X}_1 \subset C(J; V)$ , whence it follows that the embedding  $\mathcal{X}_1 \subset C(J, H^{1-\varepsilon})$  is compact. Finally, by Lemma II.2.4 of [Kry02],

<sup>4</sup>Recall that if  $B_1$  and  $B_2$  are Banach spaces embedded into a topological vector space  $L$ , then their sum  $B_1 + B_2$  is defined as the vector space of those elements  $u \in L$  that are representable in the form  $u = u_1 + u_2$ , where  $u_i \in B_i$ . It is a Banach space with respect to the norm  $\|u\| = \inf(\|u_1\|_{B_1} + \|u_2\|_{B_2})$ , where the infimum is taken over all pairs  $(u_1, u_2)$  whose sum is equal to  $u$ .



we have a continuous embedding  $W^{\alpha,4}(J;V) \subset C^{\alpha-\frac{1}{4}}(J;V)$ , whence it follows that  $\mathcal{X}_2$  is also compactly embedded in  $C(J;V)$ .

Let us denote by  $u_\nu(t,x)$  a stationary solution for (5.1) on  $J$  whose law coincides with  $\mu_\nu$ . The tightness of  $\{\mu_\nu\}$  in  $\mathcal{Y}$  will be established if we prove that

$$\mathbb{E} \|u_\nu\|_{\mathcal{X}}^2 \leq C_1, \quad (5.30)$$

where  $C_1 > 0$  is a constant not depending on  $\nu$ . To this end, we first note that, in view of (5.5), we have

$$\mathbb{E} \int_0^1 |Lu_\nu(t)|_2^2 dt = \frac{\mathfrak{B}_1}{2}. \quad (5.31)$$

Equation (5.1) implies that

$$u_\nu(t) = u_\nu^1(t) + u_\nu^2(t) + \zeta_\nu(t),$$

where we set  $\zeta_\nu = \sqrt{\nu} \zeta$ ,

$$u_\nu^1(t) = u_\nu(0) - \int_0^t Lu_\nu(s) ds, \quad u_\nu^2(t) = - \int_0^t B(u_\nu(s)) ds.$$

It follows from (5.31) that

$$\mathbb{E} \|u_\nu^1\|_{H^1(J;H)}^2 \leq C_2. \quad (5.32)$$

We now need the following properties of the nonlinear term  $B(u)$ .

*Exercise 5.2.5.* For any  $p \in [1, 2)$ , the bilinear map  $(u, v) \mapsto B(u, v)$  satisfies the following inequalities:

$$|B(u, v)|_2 \leq C (|u|_2 \|u\|_2)^{1/2} \|v\|_1 \quad \text{for } u \in H^2, v \in H^1, \quad (5.33)$$

$$\|B(u, v)\|_1 \leq C \|u\|_2 \|v\|_2 \quad \text{for } u \in H^2, v \in H^2, \quad (5.34)$$

$$|B(u, v)|_p \leq C_p \|u\|_1 \|v\|_1 \quad \text{for } u \in H^1, v \in H^1. \quad (5.35)$$

Combining (5.33), (5.31) and (5.6), we see that

$$\mathbb{E} \|u_\nu^2\|_{H^1(J;H)}^2 \leq C_3 \mathbb{E} \int_0^1 \|u_\nu^2\|_1^3 \|u_\nu^2\|_2 dt \leq C_4 \quad (5.36)$$

Furthermore, since for a centred Gaussian random variable  $\xi$  with variance  $\sigma^2$  we have  $\mathbb{E} \xi^4 = 3\sigma^4$ , then

$$\begin{aligned} \mathbb{E} \left( \sum_{s \in \mathbb{Z}_0^2} c_s \beta_s^2(t) \right)^2 &= \sum_{s, s' \in \mathbb{Z}_0^2} c_s c_{s'} \mathbb{E} (\beta_s^2(t) \beta_{s'}^2(t)) \\ &\leq \sum_{s, s' \in \mathbb{Z}_0^2} c_s c_{s'} \mathbb{E} \beta_s^4(t) = 3t^2 \left( \sum_{s \in \mathbb{Z}_0^2} c_s \right)^2, \\ \mathbb{E} \left( \sum_{s \in \mathbb{Z}_0^2} c_s |\beta_s(t) - \beta_s(r)|^2 \right)^2 &\leq 3(t-r)^2 \left( \sum_{s \in \mathbb{Z}_0^2} c_s \right)^2, \end{aligned}$$

where  $c_s \in \mathbb{R}$  are arbitrary constants. It follows that

$$\mathbb{E} \int_0^1 \|\zeta\|_1^4 dt = \mathbb{E} \int_0^1 \left( \sum_{s \in \mathbb{Z}_0^2} \alpha_s b_s^2 \beta_s^2(t) \right)^2 dt \leq C_5 \mathfrak{B}_1^2,$$

$$E \int_0^1 \int_0^1 \frac{\|\zeta(t) - \zeta(r)\|_1^4}{|t - r|^{1+4\alpha}} dr dt \leq C_6 \mathfrak{B}_1^2.$$

Combining these inequalities with (5.32) – (5.36), we arrive at (5.30).

We now prove assertions (i) and (ii). As before, it suffices to consider the restrictions of the measures and of the corresponding processes to the interval  $J = [0, 1]$ . It is easy<sup>5</sup> to see that all needed properties of the process  $u(t)$  can be reformulated in terms of its law  $\boldsymbol{\mu}$ . For instance, the energy and enstrophy of  $u$  are random constants if and only if the image of  $\boldsymbol{\mu}$  under the mapping

$$w \mapsto (|w|_2^2, |\nabla w|_2^2), \quad \mathcal{Y} \rightarrow C(J; \mathbb{R}^2),$$

is concentrated on the space of constant vector functions. Thus, it suffices to establish assertions (i) and (ii) for any particular choice of the process  $u$ .

Now let  $\boldsymbol{\mu}$  be a limiting measure for the family  $\{\boldsymbol{\mu}_\nu\}$  as  $\nu \rightarrow 0^+$  in the sense of weak convergence in the space  $\mathcal{Y}$ . Relation (5.30) and Lemma 1.2.17 with  $f(u(\cdot)) = \int_J \|u\|_2^2 dt$  and  $\pi_n(u(\cdot)) = \mathbb{P}_n$  (where  $\mathbb{P}_n$  is the projection from Section 2.1.5) imply that  $\boldsymbol{\mu}(L^2(J; V^2)) = 1$ .

By Skorokhod's embedding theorem (see Theorem 11.7.2 in [Dud02]), there is a sequence of stationary processes  $\tilde{u}_{\nu_j}(t, x)$  with  $\nu_j \rightarrow 0^+$  and an  $\mathcal{Y}$ -valued random variable  $u(t, x)$  that are defined on the same probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and satisfy the properties

$$\mathcal{D}(\tilde{u}_{\nu_j}) = \boldsymbol{\mu}_{\nu_j}, \quad \mathcal{D}(u) = \boldsymbol{\mu}, \quad (5.37)$$

$$\mathbb{P}\{\tilde{u}_{\nu_j} \rightarrow u \text{ in } \mathcal{Y} \text{ as } j \rightarrow \infty\} = 1. \quad (5.38)$$

As  $\boldsymbol{\mu}(L^2(J; V^2)) = 1$ , then  $u \in L^2(J; V^2)$  almost surely. Since  $\tilde{u}_{\nu_j}$  is a stationary process, so is  $u$ .

Let us prove that almost every trajectory of  $u$  satisfies the Euler equation (5.24). Indeed, since  $\boldsymbol{\mu}_{\nu_j}$  is the law of a stationary solution for (5.1), the process  $u_{\nu_j}$  can be written as

$$u_{\nu_j}(t) = u_{\nu_j}(0) - \int_0^t (\nu_j L u_{\nu_j} + B(u_{\nu_j})) ds + \sqrt{\nu_j} \zeta_j(t), \quad (5.39)$$

where  $\zeta_j$  is a  $V$ -valued spatially regular white noise distributed as  $\zeta$ . Since the random sequence  $\{\sqrt{\nu_j} \zeta_j\} \subset C(J; H)$  goes to zero in probability, passing to a subsequence, we can assume that

$$\mathbb{P}\{\sqrt{\nu_j} \zeta_j \rightarrow 0 \text{ in } C(J; H) \text{ as } j \rightarrow \infty\} = 1. \quad (5.40)$$

<sup>5</sup>It is not straightforward that the space defined by (5.23) is a Borel subset of  $C(\mathbb{R}_+; H^{1-\varepsilon}) \cap L^2(\mathbb{R}_+; H^{2-\varepsilon})$ . The verification of this property is the content of Exercise 5.2.6.

Finally, since  $\mathbb{E}\|u_{\nu_j}\|_{L^2(J, H^2)}^2 \leq C$  (see (5.31)), passing again to a subsequence, we see that

$$\mathbb{P}\{\nu_j L u_{\nu_j} \rightarrow 0 \text{ in } L^2(J; H) \text{ as } j \rightarrow \infty\} = 1. \quad (5.41)$$

Let us denote by  $\Omega_*$  the intersection of the events in the left-hand sides of (5.38), (5.40) and (5.41). Then  $\mathbb{P}(\Omega_*) = 1$ , and for any  $\omega \in \Omega_*$ , we can pass to the limit in relation (5.39) as  $j \rightarrow \infty$ , regarding it as an equality in  $L^1(J \times \mathbb{T}^2)$ . This results in the relation

$$u(t) = u(0) - \int_0^t B(u) ds, \quad t \in J,$$

which is equivalent to (5.24).

We now prove some further regularity properties for  $u$ . It follows from (5.34) and (5.35) that

$$\int_0^1 \|B(u)\|_1 \leq C_8 \int_0^1 \|u\|_2^2 dt, \quad \operatorname{ess\,sup}_{t \in J} |B(u)|_p \leq C_9 \operatorname{ess\,sup}_{t \in J} \|u\|_1^2.$$

Since  $u \in L^2(J; V^2) \cap C(J; V)$  with probability 1, we conclude that the right-hand sides in the above inequalities are almost surely finite. Combining this with Eq. (5.24), we see that  $u \in W^{1,1}(J; V) \cap W^{1,\infty}(J; L^p)$ .

It remains to prove that the energy and enstrophy are random constants. Let us fix any  $\omega \in \Omega$  for which  $u \in L^2(J; V^2) \cap W^{1,1}(J, V)$ . Then the function  $t \mapsto |\nabla u(t)|^2$  is absolutely continuous, and in view of (5.24) and Lemma 2.1.16, for almost all  $t \in J$  we have

$$\frac{d}{dt} |\nabla u(t)|^2 = 2 \langle \nabla \dot{u}(t), \nabla u(t) \rangle = -2 \langle B(u(t)), \Delta u(t) \rangle = 0.$$

This implies that  $|\nabla u(t)|_2^2$  does not depend on  $t$ . A similar argument shows that  $|u(t)|^2$  is also independent of time. The proof of the theorem is complete.  $\square$

*Exercise 5.2.6.* Prove that the space defined by (5.23) is a Borel subset of the spaces  $C(\mathbb{R}_+; H^{s-1})$  and  $L_{\text{loc}}^2(\mathbb{R}_+; H^s)$  with  $s \leq 2$ . *Hint:* It suffices to consider the restrictions to finite intervals and to show that the closed balls are Borel subsets. This can be done by describing balls with the help of some countable system of inequalities for functionals.

In Theorem 5.2.2, we showed that the energy and enstrophy of a limiting process are random constants. The proof of this fact was based on some well-known properties of the flow for the homogeneous Euler equations. It turns out that the regularity we obtained for the limiting measure is sufficient to derive an entire family of integrals of motions.

*Exercise 5.2.7.* Under the hypotheses of Theorem 5.2.2, show that if  $g: \mathbb{R} \rightarrow \mathbb{R}$  is a continuous function with at most a polynomial growth at infinity, then the expression

$$\langle g(\operatorname{curl} u(t)) \rangle = (2\pi)^{-2} \int_{\mathbb{T}^2} g(\operatorname{curl} u(t, x)) dx \quad (5.42)$$

defines a random constant not depending on time. *Hint:* In the case of regular solutions of the Euler equations, quantity (5.42) is preserved. This fact follows immediately from the observation that the vorticity satisfies a transport equation defined by the vector field  $u(t, x)$ . Prove that a similar argument applies in our case, on condition that  $g$  is a  $C^1$ -function with a bounded derivative, and then use a standard regularisation procedure to treat the general situation.

We now construct a topological space  $X \subset V$  in which the Euler equation is well posed and any limit point of stationary distributions  $\{\mu_\nu\}$  is an invariant measure for the corresponding flow. Let us note that if  $\mu$  is a limiting measure obtained in Corollary 5.2.3, then for  $\mu$ -almost every initial function  $u_0 \in V$  the Euler equation (5.24) has a global solution belonging to space (5.23). The following exercise shows that the constructed solution is unique.

*Exercise 5.2.8.* Let  $u_1$  and  $u_2$  be two solutions of (5.24) that belong to the space  $L^2(J; V^2) \cap C(J; V)$ , where  $J = [0, T]$  is a finite interval. Assume that  $u_1(\tau) = u_2(\tau)$  for some  $\tau \in J$ . Show that  $u_1$  and  $u_2$  coincide. *Hint:* The proof can be carried out with the help of an argument due to Yudovich [Jud63] (see also [Che98]).

Let us introduce the Fréchet space  $\mathcal{K} = L^2_{\text{loc}}(\mathbb{R}; V^2) \cap W^{1,1}_{\text{loc}}(\mathbb{R}; V)$  and denote by  $\mathcal{K}_E$  the set of curves  $u \in \mathcal{K}$  that satisfy the Euler equation (5.24) on the real line  $\mathbb{R}$ . We note that  $\mathcal{K}_E$  is a closed subset of  $\mathcal{K}$  and provide it with the distance induced from  $\mathcal{K}$ . Then  $\mathcal{K}$  becomes a Polish space. Clearly, the restriction of a function  $u \in \mathcal{K}$  to any interval  $J \subset \mathbb{R}$  belongs to the space  $L^2(J; V^2) \cap C(J; V)$ . Therefore, by Exercise 5.2.8, if  $u_1, u_2 \in \mathcal{K}_E$  are such that  $u_1(0) = u_2(0)$ , then  $u_1 = u_2$ .

Let  $\pi : \mathcal{K} \rightarrow V$  be the linear continuous operator that takes  $u$  to  $u(0)$  and let  $\pi_E$  be its restriction to  $\mathcal{K}_E$ . What has been said above implies that  $\pi_E$  is an injective mapping. Let us denote by  $X$  its image and endow it with the induced distance.<sup>6</sup> Thus,  $X$  is a complete metric space embedded in  $V$  and homeomorphic to  $\mathcal{K}_E$ . The very definition of  $X$  implies that for any  $u_0 \in X$  the function  $\pi_E^{-1}(u_0) \in \mathcal{K}_E$  is the unique solution of the Euler equation (5.24) supplemented with the initial condition

$$u(0) = u_0, \tag{5.43}$$

Moreover, this solution continuously depends on  $u_0$ . Thus, we can define a group of homeomorphisms  $S_t : X \rightarrow X$  such that  $\{S_t u_0, t \in \mathbb{R}\}$  is a unique global solution of (5.24), (5.43). The mapping  $\pi_E$  conjugates any  $S_t$  with the time shift  $\theta_t : \mathcal{K}_E \ni u(\cdot) \mapsto u(\cdot + t)$ . The following theorem establishes some further properties of  $X$  and  $S_t$ .

**Theorem 5.2.9.** (i) *The space  $V^3$  is continuously embedded in  $X$ .*

(ii) *For any  $u_0 \in X$ , the curve  $\{S_t u_0, t \in \mathbb{R}\} \subset X$  is continuous.*

(iii) *For any  $u \in \mathcal{K}_E$  and any bounded continuous function  $g : \mathbb{R} \rightarrow \mathbb{R}$ , the integral (5.42) is independent of time, as is the  $L^2$ -norm  $|u(t)|_2$ .*

<sup>6</sup>In other words,  $\text{dist}_X(\pi_E(u), \pi_E(v)) = \text{dist}_{\mathcal{K}_E}(u, v)$  for all  $u, v \in \mathcal{K}_E$ .

(iv) Under the hypotheses of Theorem 5.2.2, any limit point of the family  $\{\mu_\nu\}$  as  $\nu \rightarrow 0^+$  is concentrated on  $X$  and is an invariant measure for  $S_t$ .

*Proof.* (i) It is well known that the 2D Euler equations are well posed in the space  $V^3$ , and the corresponding solutions belong to  $C(\mathbb{R}; V^3)$ ; e.g., see Chapter 17 in [Tay97]. Using the equation and properties of the nonlinear term, it is straightforward to see that any solution  $u$  with an initial condition from  $V^3$  satisfies the inclusion  $\dot{u} \in C(\mathbb{R}; V^2)$ , whence we conclude that  $u \in \mathcal{K}_E$ . This means that  $V^3 \subset X$ .

(ii) This follows from the fact that the shifts  $\theta_t$  are continuous homeomorphisms of  $\mathcal{K}_E$ .

(iii) We confine ourselves to the sketch of the proof, leaving the details to the reader as an exercise. It suffices to show that, for any time-dependent vector field  $u \in \mathcal{K}_E$ , the Cauchy problem

$$\dot{y} = u(t, y), \quad y(s) = y_0, \quad s \in \mathbb{R}, \quad y_0 \in \mathbb{T}^2, \quad (5.44)$$

is well posed, and the corresponding flow  $U(t, s) : \mathbb{T}^2 \rightarrow \mathbb{T}^2$  that takes  $y_0$  to  $y(t)$  preserves the Lebesgue measure on  $\mathbb{T}^2$ . Indeed, if these properties are established, then the classical argument of the theory of first-order linear PDE's show that  $(\operatorname{curl} u)(t, x) = (\operatorname{curl} u)(0, U(0, t)x)$ . Combining this with the measure-preserving property of  $U$ , for any function  $g \in C_b(\mathbb{R})$  we obtain

$$\int_{\mathbb{T}^2} g(\operatorname{curl} u(t, x)) dx = \int_{\mathbb{T}^2} g(\operatorname{curl} u(0, U(0, t)x)) dx = \int_{\mathbb{T}^2} g(\operatorname{curl} u(0, x)) dx.$$

We now prove the above-mentioned result on problem (5.44). It is well known that any function from the unit ball in  $H^2$  possesses the modulus of continuity  $\sigma(r) = r |\log r|$ . It follows that any time-dependent vector field  $u \in \mathcal{K}$  has a modulus of continuity in  $x$  that has the form  $\varphi(t)\sigma(r)$ , where  $\varphi \in L^2_{\text{loc}}(\mathbb{R})$ . Thus, by Osgood's criterion (see Corollary 6.2 in Chapter III of [Har64]), problem (5.44) is well posed. The fact that the corresponding flow preserves the Lebesgue measure would follow from the relation  $\operatorname{div}_x u \equiv 0$  if the vector field  $u$  was  $C^1$ -smooth. In our situation, it suffices to approximate it by regular divergence-free vector field and to pass to the limit.

(iv) Let  $\mu \in \mathcal{P}(H)$  be a limit point of the family  $\{\mu_\nu\}$ . Then there is a subsequence  $\nu_j \rightarrow 0^+$  such that  $\mu_{\nu_j} \rightarrow \mu$  as  $j \rightarrow \infty$ . Let  $\{u_j\}$  be a sequence of stationary solutions of the Navier–Stokes system (5.1) with  $\nu = \nu_j$  that are defined throughout the real line and distributed at any fixed time as  $\mu_{\nu_j}$  and let  $\mu_j$  be the law of the process  $u_j$ . Analogues of Theorem 5.2.2 and Corollary 5.2.3 for the real line imply that  $\mu_j$  converges weakly to the law of a stationary process  $u(t)$  whose almost every trajectory belongs to the set  $\mathcal{K}_E$ . It follows from what has been said that  $\mathcal{D}(u(t)) = \mu$  for any  $t \in \mathbb{R}$  and  $\mu(X) = 1$ . Furthermore, with probability 1, we have

$$S_t(u(0)) = u(t) \quad \text{for all } t \in \mathbb{R},$$

whence we conclude that  $\mu$  is an invariant measure for  $S_t$ . The proof of the theorem is complete.  $\square$

*Remark 5.2.10.* It is an interesting open question to decide whether  $X$  is a *vector* subspace of  $V$ . We believe that it is not the case, though we cannot prove this fact.

We proved that the group of homeomorphisms  $S_t : X \rightarrow X$  possesses an invariant measure  $\mu$ , which was obtained as an inviscid limit of stationary distributions for Navier–Stokes system. Let us define the space  $Y = \mathcal{P}(\mathbb{R}) \times \mathbb{R}_+$  with the natural metric on the product space (where  $\mathcal{P}(\mathbb{R})$  is endowed with the dual-Lipschitz distance) and consider the mapping

$$\Psi : X \rightarrow Y, \quad u \mapsto ((\operatorname{curl} u)_*(dx), |u|_2),$$

where  $dx = (2\pi)^{-2}dx$  is the normalised Lebesgue measure on the torus  $\mathbb{T}^2$ . Using the continuity of the embedding  $X \subset V$ , it is straightforward to see that  $\Psi$  is continuous. Furthermore, assertion (iii) of Theorem 5.2.9 implies that  $\Psi(S_t u_0)$  is constant for any  $u_0 \in X$ . In other words, we have  $\Psi \circ S_t = \Psi$  for any  $t \in \mathbb{R}$ . Thus, denoting by  $X_{\mathbf{b}}$  the pre-image<sup>7</sup> of  $\mathbf{b} \in Y$  under  $\Psi$  and endowing it with the topology induced from  $X$ , we see that  $S_t : X_{\mathbf{b}} \rightarrow X_{\mathbf{b}}$  is a homeomorphism for all  $t \in \mathbb{R}$ . Let us set  $\lambda = \Psi_*(\mu)$ . The following result is an immediate consequence of a disintegration theorem for probability measures; see Sections III.70–73 in [DM78].

**Theorem 5.2.11.** *There is a random probability measure  $\{\lambda_{\mathbf{b}}, \mathbf{b} \in Y\}$  on the space  $X$  such that the following properties hold.*

(i) *We have  $\lambda_{\mathbf{b}}(X_{\mathbf{b}}) = 1$  for all  $\mathbf{b} \in Y$  and*

$$\mu(A \cap \Psi^{-1}(B)) = \int_B \lambda_{\mathbf{b}}(A) \lambda(d\mathbf{b}) \quad \text{for any } A \in \mathcal{B}(X), B \in \mathcal{B}(Y).$$

(ii) *For  $\lambda$ -almost every  $\mathbf{b} \in Y$ , the measure  $\lambda_{\mathbf{b}}$  is invariant for the dynamical system  $S_t : X_{\mathbf{b}} \rightarrow X_{\mathbf{b}}$ .*

*Proof.* Assertion (i) is just a reformulation of the disintegration theorem, and we confine ourselves to the proof of (ii). The relation  $\Psi \circ S_t = \Psi$  implies that  $S_t(\Psi^{-1}(B)) = \Psi^{-1}(B)$  for any Borel subset  $B \subset Y$ . Since  $(S_t)_*\mu = \mu$ , we have

$$\int_B \lambda_{\mathbf{b}}(S_t(A)) \lambda(d\mathbf{b}) = \int_B \lambda_{\mathbf{b}}(A) \lambda(d\mathbf{b}) \quad \text{for any } A \in \mathcal{B}(X), B \in \mathcal{B}(Y).$$

It follows that  $(S_t)_*\lambda_{\mathbf{b}} = \lambda_{\mathbf{b}}$  for any  $t \in \mathbb{R}$  and  $\lambda$ -almost every  $\mathbf{b} \in Y$ . Combining this with the continuity of  $S_t$  in time, we arrive at the required result.  $\square$

<sup>7</sup>Note that  $X_{\mathbf{b}}$  is a closed subset of  $X$ , since it is the pre-image of a closed set under a continuous mapping.

### 5.2.2 Estimates for the densities of the energy and enstrophy

The balance relations established in Theorem 5.1.1 and Proposition 5.1.3 give infinitely many identities for stationary measures of the Navier–Stokes system (5.1). Unfortunately, the convergence established in Theorem 5.2.2 is not strong enough to pass to the limit in those relations and to obtain similar information for the accumulation points of the family of stationary distributions, even though both sides of relation (5.12) are well defined for a limit process  $u$  and any bounded continuous function  $g$ . Instead, in this section we prove that the concept of local time, combined with relation (5.5), enables one to get some qualitative properties of the energy and enstrophy functionals in the limiting stationary regimes. Namely, we establish the following result.

**Theorem 5.2.12.** *Let  $\mathfrak{B}_1 < \infty$  and let  $\{\mu_\nu, \nu > 0\}$  be a family of stationary measures for (5.1). Then the following properties hold for any limit point  $\mu_0$  of the family  $\{\mu_\nu\}$  in the sense of weak convergence of measures on  $H$ .*

- (i) *Let  $b_s \neq 0$  for at least two indices in  $\mathbb{Z}_0^2$ . Then  $\mu_0$  has no atom at zero. Moreover, there is a constant  $C > 0$  not depending on  $\{b_s\}$  such that*

$$\mu_0(B_H(\delta)) \leq C \gamma^{-1} \sqrt{\mathfrak{B}_1} \delta \quad \text{for all } \delta > 0, \quad (5.45)$$

where we set  $\gamma = \inf\{\mathfrak{B}_0 - b_s^2, s \in \mathbb{Z}_0^2\}$ .

- (ii) *Let  $b_s \neq 0$  for all  $s \in \mathbb{Z}_0^2$ . Then there is an increasing continuous function  $p(r)$ , vanishing at  $r = 0$  and depending only on the sequence  $\{b_s\}$ , such that*

$$\mu_0(\{u \in H : |u|_2 \in \Gamma\}) + \mu_0(\{u \in H : |\nabla u|_2 \in \Gamma\}) \leq p(\ell(\Gamma)) \quad (5.46)$$

for any Borel subset  $\Gamma \subset \mathbb{R}$ , where  $\ell$  stands for the Lebesgue measure on  $\mathbb{R}$ .

Before presenting a proof of this result, we outline its main idea. We claim that it suffices to establish inequalities (5.45) and (5.46) for any measure of the family  $\{\mu_\nu\}$  with some universal constant  $C$  and function  $p$ . Indeed, consider, for instance, the case of the second inequality. It follows from (1.29) that inequality (5.46) remains true for any limiting measure  $\mu_0$  and any open subset  $\Gamma \subset \mathbb{R}$ . Now recall that the Lebesgue measure is regular:

$$\ell(\Gamma) = \inf\{\ell(G) : G \supset \Gamma, G \text{ is open}\}.$$

Combining this relation with the continuity of  $p$ , we conclude that (5.46) is true for any Borel set  $\Gamma \subset \mathbb{R}$ .

The proof of inequalities (5.45) and (5.46) for the measures of the family  $\{\mu_\nu\}$  is based on an application of some properties of the local times for Itô processes. If we were just interested in absolute continuity of the law for the components  $\langle u, e_s \rangle$  under a measure  $\mu_\nu$  with a fixed  $\nu > 0$ , then it would be sufficient to remark that the corresponding stationary process  $\langle u_\nu(t), e_s \rangle$  is

an Itô process with a non-zero constant diffusion, and the required assertion would follow immediately from a well-known technique involving the concept of local time (see Step 2 in the proof below) or from Krylov's estimate (see Theorem 7.9.1). In our situation, the estimates must be uniform in  $\nu > 0$ , and diffusion terms in the equations satisfied by  $|u_\nu(t)|_2^2$  and  $|\nabla u_\nu(t)|_2^2$  are not separated from zero. This difficulty will be overcome with the help of a thorough study of the local times for the processes in question.

*Proof of Theorem 5.2.12.* As was mentioned above, it suffices to find a constant  $C > 0$  and a function  $p$  with the properties stated in the theorem such that (5.45) and (5.46) hold for any measure  $\mu'$  of the family  $\{\mu_\nu\}$ . This will be done in several steps. In what follows, we denote by  $u(t, x)$  a stationary solution for (5.1) whose law coincides with  $\mu'$  and write  $u_s(t) = \langle u(t), e_s \rangle$  for its components in the Hilbert basis  $\{e_s, s \in \mathbb{Z}_0^2\}$ .

*Step 1.* We first show that, for any real-valued function  $g \in C^2(\mathbb{R})$  whose second derivative has at most a polynomial growth at infinity and for any Borel subset  $\Gamma \subset \mathbb{R}$ , we have

$$\begin{aligned} \mathbb{E} \int_{\Gamma} \mathbb{I}_{(a, \infty)}(g(|u|_2^2)) \left( g'(|u|_2^2) \left( \frac{\mathfrak{B}_0}{2} - |\nabla u|_2^2 \right) + g''(|u|_2^2) \sum_{s \in \mathbb{Z}_0^2} b_s^2 u_s^2 \right) da \\ + \sum_{s \in \mathbb{Z}_0^2} b_s^2 \mathbb{E} \left( \mathbb{I}_{\Gamma}(g(|u|_2^2)) (g'(|u|_2^2) u_s)^2 \right) = 0. \end{aligned} \quad (5.47)$$

Indeed, let us fix any function  $g \in C^2(\mathbb{R})$  and consider the process  $f(t) = g(|u(t)|_2^2)$ . By Itô's formula, we have

$$f(t) = f(0) + \nu \int_0^t A(s) ds + 2\sqrt{\nu} \sum_{s \in \mathbb{Z}_0^2} b_s \int_0^t g'(|u|_2^2) u_s d\beta_s,$$

where we set

$$A(t) = 2 \left( g'(|u|_2^2) \left( \frac{\mathfrak{B}_0}{2} - |\nabla u|_2^2 \right) + g''(|u|_2^2) \sum_{s \in \mathbb{Z}_0^2} b_s^2 u_s^2 \right).$$

Let  $\Lambda_t(a)$  be the local time for  $f$ . Then, in view of relation (7.42) with  $h = \mathbb{I}_{\Gamma}$ , we have

$$2 \int_{\Gamma} \Lambda_t(a) da = 4\nu \sum_{s \in \mathbb{Z}_0^2} b_s^2 \int_0^t \mathbb{I}_{\Gamma}(f(r)) (g'(|u|_2^2) u_s)^2 dr.$$

Taking the mean value and using the stationarity of  $u$ , we derive

$$\int_{\Gamma} (\mathbb{E} \Lambda_t(a)) da = 2\nu t \sum_{s \in \mathbb{Z}_0^2} b_s^2 \mathbb{E} \left( \mathbb{I}_{\Gamma}(f) (g'(|u|_2^2) u_s)^2 \right).$$



On the other hand, taking the mean value in (7.45) and using again the stationarity of  $u$ , we obtain

$$\mathbb{E}\Lambda_t(a) = -\nu t \mathbb{E}(\mathbb{I}_{(a,\infty)}(f(0)) A(0)).$$

The last two equalities and the definition of the drift  $A$  imply the required relation (5.47).

*Step 2.* We now prove that (cf. (5.45))

$$\mu'(\{u \in H : 0 < |u|_2 \leq \delta\}) \leq C \gamma^{-1} \sqrt{\mathfrak{B}_1} \delta \quad \text{for any } \delta > 0. \quad (5.48)$$

Let us apply relation (5.47) in which  $\Gamma = [\alpha, \beta]$  with  $\alpha > 0$  and  $g \in C^2(\mathbb{R})$  is a function that coincides with  $\sqrt{x}$  for  $x \geq \alpha$  and vanishes for  $x \leq 0$ . This results in

$$\begin{aligned} \mathbb{E} \int_{\alpha}^{\beta} \mathbb{I}_{(a,\infty)}(|u|_2) \left( \frac{\mathfrak{B}_0 - 2|\nabla u|_2^2}{4|u|_2} - \frac{1}{4|u|_2^3} \sum_{s \in \mathbb{Z}_0^2} b_s^2 u_s^2 \right) da \\ + \frac{1}{4} \sum_{s \in \mathbb{Z}_0^2} b_s^2 \mathbb{E}(\mathbb{I}_{[\alpha,\beta]}(|u|_2) |u|_2^{-2} u_s^2) = 0. \end{aligned}$$

It follows that

$$\mathbb{E} \int_{\alpha}^{\beta} \frac{\mathbb{I}_{(a,\infty)}(|u|_2)}{|u|_2^3} \left( \mathfrak{B}_0 |u|_2^2 - \sum_{s \in \mathbb{Z}_0^2} b_s^2 u_s^2 \right) da \leq 2(\beta - \alpha) \mathbb{E} \left( \frac{|\nabla u|_2^2}{|u|_2} \right). \quad (5.49)$$

Now note that

$$\begin{aligned} \mathfrak{B}_0 |u|_2^2 - \sum_{s \in \mathbb{Z}_0^2} b_s^2 u_s^2 &= \sum_{s \in \mathbb{Z}_0^2} (\mathfrak{B}_0 - b_s^2) u_s^2 \geq \gamma |u|_2^2, \\ \mathbb{E} \left( \frac{|\nabla u|_2^2}{|u|_2} \right) &\leq C_1 \mathbb{E} |Lu|_2 \leq C_1 \sqrt{\mathfrak{B}_1}, \end{aligned}$$

where we used interpolation and inequality (5.27). Substituting these estimates into (5.49) and passing to the limit as  $\alpha \rightarrow 0^+$ , we derive

$$\mathbb{E} \int_0^{\beta} \mathbb{I}_{(a,\infty)}(|u|_2) |u|_2^{-1} da \leq C_1 \gamma^{-1} \sqrt{\mathfrak{B}_1} \beta. \quad (5.50)$$

We now fix a constant  $\delta > 0$  and note that the left-hand side of (5.50) can be minorised by

$$\begin{aligned} \mathbb{E} \int_0^{\beta} \mathbb{I}_{(a,\delta]}(|u|_2) |u|_2^{-1} da &\geq \delta^{-1} \mathbb{E} \int_0^{\beta} \mathbb{I}_{(a,\delta]}(|u|_2) da \\ &= \delta^{-1} \int_0^{\beta} \mathbb{P}(\{a < |u|_2 \leq \delta\}) da. \end{aligned}$$

Substituting this inequality into (5.50), we obtain

$$\frac{1}{\beta} \int_0^\beta \mathbb{P}(\{a < |u|_2 \leq \delta\}) da \leq C_1 \gamma^{-1} \sqrt{\mathfrak{B}_1} \delta.$$

Passing to the limit as  $\beta \rightarrow 0^+$ , we arrive at the required inequality (5.48).

*Step 3.* In view of (5.48), inequality (5.45) will be established if we prove that  $\mu'$  has no atom at zero. To this end, we could apply the local time technique used in Step 1 to a one-dimensional projection of  $u$ . However, the following proof based on Krylov's estimate for Itô processes is simpler.

Let us fix an index  $s \in \mathbb{Z}_0^2$  such that  $b_s \neq 0$ . The stationary process  $u_s(t)$  satisfies the equation

$$u_s(t) = u_s(0) + \int_0^t v(r) dr + \sqrt{\nu} b_s \beta_s(t),$$

where we set  $v(t) = \langle -\nu Lu - B(u), e_s \rangle$ . Since  $|u|_2 \geq |u_s|$ , it suffices to show that  $P = \mathbb{P}\{u_s(0) = 0\} = 0$ . By Theorem 7.9.1 with  $d = 1$  and  $f = \mathbb{I}_{(-\varepsilon, \varepsilon)}$ , we have

$$\sqrt{\nu} b_s \mathbb{P}\{|u_s(0)| < \varepsilon\} \leq C\varepsilon \quad \text{for any } \varepsilon > 0,$$

whence it follows that  $P = 0$ .

*Step 4.* It remains to establish (5.46). We shall confine ourselves to the proof of the fact that the first term on the left-hand side can be bounded by  $p(\ell(\Gamma))$ . A similar argument shows that the same is true for the second term.

Applying relation (5.47) with  $g(x) = x$ , we obtain

$$\mathbb{E} \left( \mathbb{I}_\Gamma(|u|_2^2) \sum_{s \in \mathbb{Z}_0^2} b_s^2 u_s^2 \right) \leq \int_\Gamma \mathbb{E} \left( \mathbb{I}_{(a, \infty)}(|u|_2^2) |\nabla u|_2^2 \right) da \leq \frac{1}{2} \mathfrak{B}_0 \ell(\Gamma), \quad (5.51)$$

where we used (5.3) to get the second inequality. We wish to estimate the left-hand side of this inequality from below. To this end, note that if  $|u|_2 \geq \delta$  and  $|Lu|_2 \leq \delta^{-1/2}$ , then for any integer  $N \geq 1$  we have

$$\begin{aligned} \sum_{s \in \mathbb{Z}_0^2} b_s^2 u_s^2 &\geq \underline{b}_N^2 \sum_{0 < |s| \leq N} u_s^2 = \underline{b}_N^2 \left( |u|_2^2 - \sum_{|s| > N} u_s^2 \right) \\ &\geq \underline{b}_N^2 \left( |u|_2^2 - N^{-4} |Lu|_2^2 \right) \geq \underline{b}_N^2 (\delta^2 - N^{-4} \delta^{-1}), \end{aligned}$$

where  $\underline{b}_N = \min\{b_s, |s| \leq N\}$ . Choosing  $N = N(\delta)$  sufficiently large, we find an increasing function  $\varepsilon(\delta) > 0$  going to zero with  $\delta$  such that

$$\sum_{s \in \mathbb{Z}_0^2} b_s^2 u_s^2 \geq \varepsilon(\delta) \quad \text{for } |u|_2 \geq \delta, \|u\|_2 \leq \delta^{-1/2}. \quad (5.52)$$

Define now the event  $G_\delta = \{|u(0)|_2 \leq \delta \text{ or } |Lu(0)|_2 \geq \sqrt{\delta}\}$  and note that, in view of (5.5), (5.45), and Chebyshev's inequality, we have

$$\mathbb{P}(G_\delta) \leq \mathbb{P}\{|u|_2 \leq \delta\} + \mathbb{P}\{|Lu|_2 \geq \sqrt{\delta}\} \leq C_5 \delta.$$

Combining this with (5.51) and (5.52), we write

$$\begin{aligned} \mathbb{P}\{|u|_2^2 \in \Gamma\} &= \mathbb{P}(\{|u|_2^2 \in \Gamma\} \cap G_\delta) + \mathbb{P}(\{|u|_2^2 \in \Gamma\} \cap G_\delta^c) \\ &\leq C_5\delta + \varepsilon(\delta)^{-1} \mathbb{E} \left( \mathbb{I}_\Gamma(|u|_2^2) \sum_{s \in \mathbb{Z}_0^2} b_s^2 u_s^2 \right) \\ &\leq C_5\delta + C_6\varepsilon(\delta)^{-1} \ell(\Gamma). \end{aligned}$$

This inequality immediately implies the required result.  $\square$

*Exercise 5.2.13.* Prove that the second term on the left-hand side of (5.46) can be bounded by  $p(\ell(\Gamma))$ . *Hint:* A relation similar to (5.47) holds for the process  $g(|\nabla u|_2^2)$ . To see this, it suffices to repeat the arguments used in Step 1 of the above proof.

### 5.2.3 Further properties of the limiting measures

In the previous subsection, we have shown that the limiting measures of the family of stationary distributions for (5.1) possess some non-degeneracy properties. Indeed, Theorem 5.2.12 implies, in particular, that the image of such a measure under the energy and enstrophy functionals has no atoms. In this section, we combine balance relations with Krylov's estimate for Itô diffusions to show that a similar assertion is true for a large class of multidimensional functionals.

Let  $f_1, \dots, f_d : \mathbb{R} \rightarrow \mathbb{R}$  be real-analytic functions whose second derivatives have at most a polynomial growth at infinity and are bounded from below:

$$f_k''(z) \geq -C \quad \text{for } z \in \mathbb{R}, k = 1, \dots, d. \quad (5.53)$$

These hypotheses are satisfied, for instance, if the functions  $f_k$  are trigonometric polynomials or polynomials of even degree with positive leading coefficients. Since  $H^1(\mathbb{T}^2)$  is continuously embedded in  $L^p(\mathbb{T}^2)$  for any  $p < \infty$ , we can consider the map  $F : V^2 \rightarrow \mathbb{R}^d$  defined by

$$F(u) = (F_1(u), \dots, F_d(u)), \quad F_k(u) = \int_{\mathbb{T}^2} f_k(\operatorname{curl} u(x)) dx.$$

Assume that the functions  $f_k'$  are independent modulo constants, that is, if  $c_1 f_1' + \dots + c_d f_d' \equiv \text{const}$ , then all  $c_k$  vanish. The following theorem gives some information on the distribution of  $F$  under any limiting point of the family of stationary measures for (5.1).

**Theorem 5.2.14.** *Let  $b_s = b_{-s} \neq 0$  for all  $s \in \mathbb{Z}_0^2$ , let  $\mathfrak{B}_2 < \infty$ , and let functions  $f_1, \dots, f_d$  be as above. Then there is an increasing continuous function  $p(r)$  vanishing at  $r = 0$  such that for any limiting point  $\mu$  of the family  $\{\mu_\nu\}$ , as  $\nu \rightarrow 0$ , in the sense of weak convergence on  $H$  and any Borel subset  $\Gamma \in \mathbb{R}^d$ , we have<sup>8</sup>*

$$F_*(\mu)(\Gamma) \leq p(\ell(\Gamma)). \quad (5.54)$$

<sup>8</sup>Recall that, by (5.27), any limiting point  $\mu$  is concentrated on the space  $V^2$ .

Before proving this result, we establish a corollary showing that the limiting measures are not concentrated on subsets of low dimension. Its formulation uses the concept of Hausdorff dimension (e.g., see Section X.1 of the book [BV92]).

**Corollary 5.2.15.** *Under the hypotheses of Proposition 5.1.3, let  $X \subset H$  be a closed subset whose intersection with any compact set  $\mathcal{K} \subset V^2$  has finite Hausdorff dimension with respect to the metric of  $V^2$  and let  $\mu$  be a limiting point of  $\{\mu_\nu\}$ , as  $\nu \rightarrow 0$ , in the sense of weak convergence on  $H$ . Then  $\mu(X) = 0$ .*

*Proof.* We shall only outline the proof, leaving it to the reader to fill in the details. We know that  $\mu$  is concentrated on  $V^2$ . By Ulam's theorem (see Section 1.2.1), any Borel measure on a Polish space is regular, and therefore we can find an increasing sequence of compact subsets  $\mathcal{K}_n \subset V^2$  such that  $\mu(\mathcal{K}_n) \rightarrow 1$  as  $n \rightarrow \infty$ . Hence, it suffices to show that  $\mu(\mathcal{K}_n \cap X) = 0$  for any  $n \geq 1$ . Thus, we can assume from the very beginning that  $X$  is a compact subset in  $V^2$  of finite Hausdorff dimension.

Let us take an integer  $d \geq 1$  larger than Hausdorff's dimension of  $X$  and consider any map  $F : V^2 \rightarrow \mathbb{R}^d$  satisfying the hypotheses of Theorem 5.2.14. It is easy to see that  $F$  is uniformly Lipschitz continuous on any compact subset of  $V^2$ . Combining this with the definition of the Hausdorff dimension, we see that the compact set  $F(X) \subset \mathbb{R}^d$  has zero Lebesgue measure. Inequality (5.54) with  $\Gamma = F(X)$  now implies that  $\mu(X) = 0$ .  $\square$

*Proof of Theorem 5.2.14.* It suffices to prove that inequality (5.54) holds for any measure  $\mu_\nu$  with a function  $p$  not depending on  $\nu$ . Let us fix a stationary solution  $u(t)$  distributed as  $\mu_\nu$  and define the stochastic process  $y(t) = F(u(t))$ . The first and second derivatives of the mapping  $F : V^2 \rightarrow \mathbb{R}^d$  have the form

$$\begin{aligned} F'(u; h) &= \left( \int_{\mathbb{T}^2} f'_k(\operatorname{curl} u(x)) \operatorname{curl} h(x) dx, k = 1, \dots, d \right), \\ F''(u; h, h) &= \left( \int_{\mathbb{T}^2} f''_k(\operatorname{curl} u(x)) (\operatorname{curl} h(x))^2 dx, k = 1, \dots, d \right). \end{aligned}$$

Combining these relations with estimates (2.162), Corollary 7.7.6, and Eq. (5.8) for the vorticity  $v = \operatorname{curl} u$ , we prove that  $y(t)$  is an Itô process whose components are representable in the form (cf. proof of Proposition 5.1.3)

$$y_k(t) = y_k(0) + \int_0^t A_k(r) dr + \sum_{s \in \mathbb{Z}_0^2} \int_0^t B_{ks}(r) d\beta_s, \quad k = 1, \dots, d, \quad (5.55)$$

where we set

$$\begin{aligned} A_k(t) &= \int_{\mathbb{T}^2} f'_k(v(t, x)) \Delta v(t, x) dx + \frac{1}{2} \sum_{s \in \mathbb{Z}_0^2} b_s^2 \int_{\mathbb{T}^2} f''_k(v(t, x)) \varphi_s^2(x) dx \\ &= \int_{\mathbb{T}^2} f''_k(v(t, x)) (-|\nabla v|^2 + \frac{\mathfrak{B}_1}{8\pi^2}) dx, \\ B_{ks}(t) &= b_s \int_{\mathbb{T}^2} f'_k(v(t, x)) \varphi_s(x) dx. \end{aligned}$$

We now apply Krylov’s estimate (7.46) with  $f = \mathbb{I}_\Gamma$  to the process  $y$ . In view of stationarity of  $y$ ,  $A_k$ , and  $B_{ks}$ , this results in

$$\mathbb{E}((\det \sigma_0)^{1/d} \mathbb{I}_\Gamma(y(0))) \leq C_d \ell(\Gamma)^{1/d} \mathbb{E}|A(0)|, \tag{5.56}$$

where  $A = (A_1, \dots, A_d)$ , and  $\sigma_t$  stands for the  $d \times d$  diffusion matrix  $BB^*$  (that is,  $\sigma_t^{kl} = \sum_s B_{ks}(t)B_{ls}(t)$ ). Next, we estimate  $\mathbb{E}|A(0)|$  and  $\det \sigma_0$  from above and below, respectively.

It follows from (5.53) that  $|f_k''(z)| \leq f_k''(z) + 2C$  for all  $z \in \mathbb{R}$ . Therefore,

$$\begin{aligned} \mathbb{E}|A_k(0)| &\leq \mathbb{E} \int_{\mathbb{T}^2} |f_k''(v(t, x))| \cdot |-\nabla v|^2 + \frac{\mathfrak{B}_1}{8\pi^2} dx \\ &\leq \mathbb{E} \int_{\mathbb{T}^2} (f_k''(v(t, x)) + 2C)(|\nabla v|^2 + \frac{\mathfrak{B}_1}{8\pi^2}) dx. \end{aligned}$$

The balance relation (5.17) with  $g = f_k''$  implies that

$$\mathbb{E} \int_{\mathbb{T}^2} f_k''(v(t, x)) |\nabla v(t, x)|^2 dx = C \mathbb{E} \int_{\mathbb{T}^2} f_k''(v(t, x)) dx.$$

Combining this with the above inequality, we derive

$$\mathbb{E}|A_k(0)| \leq C_1 \mathbb{E} \int_{\mathbb{T}^2} (|f_k''(v(t, x))| + |\nabla v(t, x)|^2 + 1) dx.$$

Recalling that  $f_k''$  has at most a polynomial growth at infinity and invoking estimates (5.5) and (5.6), we obtain

$$\mathbb{E}|A_k(0)| \leq C_2 \quad \text{for } k = 1, \dots, d, \tag{5.57}$$

where  $C_2 > 0$  is a constant depending only on the sequence  $\{b_s\}$ .

To estimate from below the quantity  $\det \sigma_0$ , we need an auxiliary lemma established at the end of this section. Recall that  $\dot{H}^m(\mathbb{T}^2)$  stands for the space of functions in  $H^m(\mathbb{T}^2)$  with zero mean value. We abbreviate  $f'(w) = (f_1'(w), \dots, f_d'(w))$  and define a function  $D : \dot{H}^1(\mathbb{T}^2) \rightarrow \mathbb{R}$  by the relation  $D(w(\cdot)) = \det \sigma(f' \circ w)$ . Here for a vector function  $g : \mathbb{T}^2 \rightarrow \mathbb{R}^d$  we denote by  $\sigma(g)$  a  $d \times d$  matrix with the entries

$$\sigma^{kl}(g) = \sum_{s \in \mathbb{Z}_0^2} b_s^2 \int_{\mathbb{T}^2} g_k(x) \varphi_s(x) dx \int_{\mathbb{T}^2} g_l(x) \varphi_s(x) dx;$$

cf. definition of the diffusion matrix  $\sigma_t$ .

**Lemma 5.2.16.** *For any functions  $f_1, \dots, f_d$  satisfying the hypotheses of Theorem 5.2.14, there is  $r < 1$  such that  $D(w)$  admits a continuous extension to the space  $\dot{H}^r = \dot{H}^r(\mathbb{T}^2)$  and there does not vanish outside the origin.*

Let us note that  $\det \sigma_0 = D(v(0))$ . For any  $\delta > 0$ , we introduce the event  $Q_\delta = \{v(0) \in \mathcal{O}_\delta\}$ , where  $\mathcal{O}_\delta = \{w \in \dot{H}^1(\mathbb{T}^2) : \|w\|_1 \leq \delta^{-1}, \|w\|_r \geq \delta\}$ , and  $r$  denotes the constant constructed in the above lemma. Since the embedding  $\dot{H}^1(\mathbb{T}^2) \subset \dot{H}^r(\mathbb{T}^2)$  is compact,  $\mathcal{O}_\delta$  is a compact subset of  $\dot{H}^r(\mathbb{T}^2)$  not containing zero. Lemma 5.2.16 implies that there  $D(w) \geq c_3(\delta) > 0$ . It follows that

$$\det \sigma_0 \geq c_3(\delta) \quad \text{for } \omega \in Q_\delta^c.$$

Combining this with (5.56) and (5.57), we obtain

$$\begin{aligned} \mathbb{P}\{y(0) \in \Gamma\} &\leq \mathbb{E}(\mathbb{I}_{Q_\delta} \mathbb{I}_\Gamma(y(0))) + \mathbb{P}(Q_\delta^c) \\ &\leq (c_3(\delta))^{-1/d} \mathbb{E}((\det \sigma_0)^{1/d} \mathbb{I}_\Gamma(y(0))) + \mathbb{P}(Q_\delta^c) \\ &\leq C_4(\delta) \ell(\Gamma)^{1/d} + \mathbb{P}(Q_\delta^c). \end{aligned} \tag{5.58}$$

Let us show that the second term in the right-hand side goes to zero with  $\delta$ . Indeed, we have

$$\begin{aligned} \mathbb{P}(Q_\delta^c) &\leq \mathbb{P}\{\|v(0)\|_1 \geq \delta^{-1}\} + \mathbb{P}\{\|v(0)\|_r \geq \delta\} \\ &\leq \mathbb{P}\{\|u(0)\|_2 \geq \delta^{-1}\} + \mathbb{P}\{|u(0)|_2 \geq \delta\}, \end{aligned}$$

where we used the inequalities

$$\|\operatorname{curl} u\|_1 \leq \|u\|_2, \quad \|\operatorname{curl} u\|_r \geq |\operatorname{curl} u|_2 \geq |u|_2.$$

It follows from (5.5) and (5.45) that  $\mathbb{P}(Q_\delta^c) \rightarrow 0$  as  $\delta \rightarrow 0$ . Combining this with (5.58), we arrive at the required inequality (5.54).  $\square$

*Proof of Lemma 5.2.16.* The conditions imposed on  $f_k$  implies that

$$|f'_k(z)| + |f''_k(z)| \leq C_1(1 + |z|)^q \quad \text{for all } z \in \mathbb{R},$$

where  $q \geq 1$  is a constant. Let  $r < 1$  be such that  $\dot{H}^r$  is continuously embedded into  $L^{2q}(\mathbb{T}^2)$ . By the Cauchy–Schwarz inequality,

$$|\sigma^{kl}(g)| \leq \left( \sum_{s \in \mathbb{Z}_0^2} b_s^2 |\varphi_s|_2^2 \right) |g_k|_2 |g_l|_2 \leq C_5 \mathfrak{B}_1 |g|_2^2.$$

Therefore  $D$  is continuous as the composition of the continuous maps

$$\begin{aligned} w &\mapsto g = f'(w(x)), & \dot{H}^r(\mathbb{T}^2) &\rightarrow L^2(\mathbb{T}^2; \mathbb{R}^d), \\ g &\mapsto \det \sigma(g), & L^2(\mathbb{T}^2; \mathbb{R}^d) &\rightarrow \mathbb{R}. \end{aligned}$$

We now prove that  $D(w) \neq 0$  for  $w \in \dot{H}^r \setminus \{0\}$ . It is an elementary exercise in linear algebra to show that if  $\mathcal{H}$  is a Hilbert space and  $B : \mathcal{H} \rightarrow \mathbb{R}^d$  is a surjective continuous operator, then the application  $BB^* : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is an isomorphism. Let us denote by  $\dot{L}^2(\mathbb{T}^2)$  the space of  $L^2$  functions with zero mean value and

define a family of continuous operators  $B(w) : \dot{L}^2(\mathbb{T}^2) \rightarrow \mathbb{R}^d$  by the relation (cf. the relation for the first derivative of  $F$ )

$$B(w)\xi = \left( \sum_{s \in \mathbb{Z}_0^2} |s|^{-1} b_s \langle \xi, \varphi_s \rangle \int_{\mathbb{T}^2} f'_k(w(x)) \varphi_s(x) dx, k = 1, \dots, d \right).$$

It is straightforward to check that

$$B_{ks}(w) = b_s \int_{\mathbb{T}^2} f'_k(w(x)) \varphi_s(x) dx, \quad 1 \leq k \leq d, \quad s \in \mathbb{Z}_0^2,$$

are the matrix entries of the operator  $B(w)$  with respect to the orthonormal basis  $\{|s|^{-1} \varphi_s, s \in \mathbb{Z}_0^2\}$ . Therefore  $\sigma(f'(w)) = B(w)B^*(w)$ . It follows that the application  $\sigma(w) : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is an isomorphism and its matrix is nonsingular for those  $w \in H^r$  for which  $B(w)$  is surjective. Hence, the required assertion will be established if we show that  $B(w)$  is surjective for any non-zero function  $w \in \dot{H}^r$ .

Suppose that  $B(w)$  is not surjective. Then there exists a non-zero vector  $c = (c_1, \dots, c_d) \in \mathbb{R}^d$  such that

$$\langle c, B(w)\xi \rangle = \sum_{k=1}^d c_k \sum_{s \in \mathbb{Z}_0^2} |s|^{-1} b_s \langle \xi, \varphi_s \rangle \int_{\mathbb{T}^2} f'_k(w(x)) \varphi_s(x) dx = 0$$

for any  $\xi \in \dot{L}^2(\mathbb{T}^2)$ . That is,

$$\sum_{s \in \mathbb{Z}_0^2} |s|^{-1} b_s \langle \xi, \varphi_s \rangle \int_{\mathbb{T}^2} h(w(x)) \varphi_s(x) dx = 0 \quad \text{for any } \xi \in \dot{L}^2(\mathbb{T}^2),$$

where  $h(w) = c_1 f'_1(w) + \dots + c_d f'_d(w)$ . Choosing  $\xi = \varphi_s$  with  $s$  varying in  $\mathbb{Z}_0^2$  and recalling that the numbers  $b_s$  are non-zero, we get

$$\int_{\mathbb{T}^2} h(w(x)) \xi(x) dx = 0 \quad \text{for any } \xi \in C^\infty(\mathbb{T}^2) \text{ with zero mean value.}$$

Therefore there exists a constant  $C = C_w$  such that  $h(w(x)) = C$  almost everywhere, whence we conclude that

$$w(x) \in N_C \quad \text{for almost every } x \in \mathbb{T}^2, \quad (5.59)$$

where  $N_C = \{z \in \mathbb{R} : h(z) = C\}$ . Recalling that  $f'_1, \dots, f'_d$  are analytic functions linearly independent modulo constants, we see that  $N_C$  is a discrete<sup>9</sup> subset of  $\mathbb{R}$ . Combining (5.59) with Lemma 7.15.1, we see that the function  $w$  must be a constant almost everywhere. Since the mean value of  $w$  is zero, we conclude that  $w \equiv 0$ .  $\square$

<sup>9</sup>Recall that a subset of  $\mathbb{R}$  is said to be *discrete* if it has no finite accumulation points.

### 5.2.4 Other scalings

We have studied so far the inviscid limit of stationary measures for the Navier–Stokes equations in the case when the random perturbation is white in time and has an amplitude proportional to  $\sqrt{\nu}$ . A natural question is to find out what happens for white noise perturbations of different size. In this section, we prove that the scaling we studied is the only one for which a nontrivial limit exists.

Let us consider the equation (cf. (5.1))

$$\partial_t u + \nu Lu + B(u) = \nu^a \eta(t, x), \quad (5.60)$$

where  $a \in \mathbb{R}$  is a constant, and  $\eta$  is a regular white noise.

**Theorem 5.2.17.** *Let  $\{\mu_\nu, \nu > 0\}$  be a family of stationary measures for (5.60). Then the following assertions hold.*

- (i) *Let  $a > 1/2$ . Then  $\{\mu_\nu\}$  converges weakly on  $V^2$  to the Dirac measure  $\delta_0$  as  $\nu \rightarrow 0^+$ .*
- (ii) *Let  $a < 1/2$ . Then  $\{\mu_\nu\}$  has no accumulation points in the sense of weak convergence on  $H$  as  $\nu \rightarrow 0^+$ .*

*Proof.* We first assume that  $a > 1/2$ . By (2.161), we have

$$\int_H |Lu|_2^2 \mu_\nu(du) = \frac{\nu^{2a-1} \mathfrak{B}_1}{2}. \quad (5.61)$$

Let  $f \in L_b(V^2)$  be an arbitrary function such that  $\|f\|_L \leq 1$ . Then, applying the Cauchy–Schwarz inequality and using (5.61), we derive

$$\left| \int_H f(u) \mu_\nu(du) - f(0) \right| \leq \int_H \|u\|_2 \mu_\nu(du) \leq C_1 \nu^{a-1/2}.$$

Recalling assertion (ii) of Theorem 1.2.15, we conclude that  $\mu_\nu \rightarrow \delta_0$  weakly on  $V$ .

We now assume that  $a < 1/2$ . Let  $u_\nu(t)$  be a stationary solution for (5.60) whose law coincides with  $\mu_\nu$ . We fix a constant  $b > 0$  and consider the random process  $v_\nu(t) = \nu^b u_\nu(\nu^b t)$ . It is a matter of direct verification to show that  $v_\nu$  is a stationary solution of the equation

$$\partial_t v + \nu^{1+b} Lv + B(v) = \nu^{a+\frac{3b}{2}} \partial_t \tilde{\zeta}, \quad (5.62)$$

where  $\tilde{\zeta}(t) = \nu^{-\frac{b}{2}} \zeta(\nu^b t)$ . Using the scaling properties of the Brownian motion, it is easy to see that  $\tilde{\zeta}$  is a process with the same distribution as  $\zeta$ . Choosing  $b = \frac{1}{2} - a$ , we conclude from (5.62) that  $v_\nu$  satisfies Eq. (5.1), in which  $\zeta$  and  $\nu$  are replaced by  $\tilde{\zeta}$  and  $\tilde{\nu} = \nu^{1+b}$ , respectively.

To prove that  $\{\mu_\nu\}$  has no accumulation points as  $\nu \rightarrow 0^+$ , we argue by contradiction: suppose that  $\{\mu_{\nu_j}\}$  converges weakly on  $H$  for some sequence  $\nu_j \rightarrow 0^+$ . Then, by Prokhorov’s theorem, for any  $\varepsilon > 0$  there is  $R > 0$  such that

$$\mu_{\nu_j}(B_H(R)^c) = \mathbb{P}\{|u_{\nu_j}(0)|_2 > R\} \leq \varepsilon \quad \text{for all } j \geq 1.$$



Recalling the definition of  $v_\nu$ , we see that

$$\mathbb{P}\{|v_{\nu_j}(0)|_2 > \nu_j^b R\} = \mathbb{P}\{|u_{\nu_j}(0)|_2 > R\} \leq \varepsilon \quad \text{for all } j \geq 1,$$

whence it follows that the law of  $v_{\nu_j}(0)$  converges weakly to the Dirac measure  $\delta_0$  as  $j \rightarrow \infty$ . On the other hand, by Theorem 5.2.2 and Corollary 5.2.3, the accumulation points of the family of stationary measures for (5.1) are non-degenerate in the sense that they are not concentrated at zero (see (5.25)). Since the law of  $v_{\nu_j}(0)$  is a stationary measure for (5.1) with  $\nu = \nu_j^{1+b}$ , we arrive at a contradiction. The proof of the theorem is complete.  $\square$

*Exercise 5.2.18.* Prove that the process  $v_\nu$  defined in the above proof satisfies (5.62). *Hint:* Rewrite Eq. (5.1) in the form

$$u(t) = u(0) - \int_0^t (\nu Lu + B(u)) ds + \sqrt{\nu} \zeta(t), \quad t \geq 0,$$

and make a suitable change of the time variable.

*Remark 5.2.19.* If in Eq. (5.1) we replace the viscous term  $\nu Lu$  with a hyperviscous one of the form  $\nu L^m u$ ,  $m > 1$ , then still the equation has a unique stationary measure  $\mu_\nu^m$ , and there exist limits along sequences  $\mu_0^m = \lim_{\nu_j \rightarrow 0} \mu_{\nu_j}^m$ . For the same reason as in Section 5.2.1, we have  $\mathbb{E}_{\mu_0^m} \|u\|_m^2 = \frac{1}{2} \mathfrak{B}_0$  and  $\mathbb{E}_{\mu_0^m} \|u\|_{m+1}^2 \leq \frac{1}{2} \mathfrak{B}_1$ . As  $\mathbb{E}_{\mu_0} \|u\|_1^2 = \frac{1}{2} \mathfrak{B}_0$ , then the non-degeneracy property of Corollary 5.2.15 implies that  $\mathbb{E}_{\mu_0} \|u\|_m^2 > \frac{1}{2} B_0$ . So the measure  $\mu_0^m$  differs from  $\mu_0$  (it is smoother). That is, introducing in Eq. (5.1) a hyperviscosity, we change the limiting, as  $\nu \rightarrow 0$ , distributions of velocity and vorticity.

### 5.2.5 Kicked Navier–Stokes system

We now discuss the existence of an inviscid limit for the kicked Navier–Stokes system on  $\mathbb{T}^2$  with a random force proportional to  $\sqrt{\nu}$  (cf. (3.74)):

$$\dot{u} + \nu Lu + B(u) = \sqrt{\nu} \sum_{k=1}^{\infty} \eta_k \delta(t - k). \tag{5.63}$$

Here  $\{\eta_k\}$  is a sequence of i.i.d. random variables satisfying the hypotheses imposed in Section 3.2.4. We assume, in addition, that

$$\mathbb{E} \xi_{sk} = 0 \text{ for all } s \in \mathbb{Z}_0^2, \quad \mathfrak{B}_1 = \sum_{s \in \mathbb{Z}_0^2} |s|^2 b_s^2 < \infty,$$

and introduce the quantities

$$D_0 = \sum_{s \in \mathbb{Z}_0^2} b_s^2 \mathbb{E} \xi_{sk}^2, \quad D_1 = \sum_{s \in \mathbb{Z}_0^2} b_s^2 |s|^2 \mathbb{E} \xi_{sk}^2.$$

*Exercise 5.2.20.* Let  $u(t)$ ,  $t \geq 0$ , be a stationary solution for the Markov semigroup associated with (5.63). Prove that

$$\int_{k-1}^k |\nabla u(t)|_2^2 dt = \frac{D_0}{2}, \quad \int_{k-1}^k |Lu(t)|_2^2 dt = \frac{D_1}{2} \quad \text{for any } k \geq 1.$$

Let  $C(\mathbb{R}_+ \setminus \mathbb{Z}_+; V)$  be the space of  $V$ -valued functions on  $\mathbb{R}_+$  that are continuous on each interval  $I_k = [k-1, k)$  with  $k \geq 1$  and have a limit as  $t \rightarrow k^-$ . This space is endowed with the family of seminorms

$$\|u\|_{C(I_k; V)} = \sup_{t \in I_k} \|u(t)\|_1, \quad k \geq 1.$$

It is easily seen that  $C(\mathbb{R}_+ \setminus \mathbb{Z}_+; V)$  is a Fréchet space. The following result can be proved with the help of the argument used to establish Theorem 5.2.2; we leave its proof to the reader as an exercise.

**Theorem 5.2.21.** *Under the above hypotheses, let  $\{\mu_\nu, \nu > 0\}$  be any family of stationary measures for the discrete-time Markov semigroup associated with (5.63) and let  $\{\mu_\nu\}$  be the corresponding family of measures on the space of trajectories. Then for any  $\varepsilon > 0$  the family  $\{\mu_\nu\}$  is tight in  $C(\mathbb{R}_+ \setminus \mathbb{Z}_+; V) \cap L_{\text{loc}}^2(\mathbb{R}_+; H^{2-\varepsilon})$ . Moreover, if  $\mu$  is a limiting point for  $\{\mu_\nu\}$  and  $u(t, x)$  is a random process whose law coincides with  $\mu$ , then the following properties hold.*

(i) *Almost every realisation of  $u(t, x)$  belongs to space*

$$L_{\text{loc}}^2(\mathbb{R}_+; V^2) \cap W_{\text{loc}}^{1,1}(\mathbb{R}_+; V) \cap W_{\text{loc}}^{1,\infty}(\mathbb{R}_+; L^p),$$

*for any  $p \in [1, 2)$ , and satisfies the Euler equation (5.24).*

(ii) *The process  $u(t, x)$  is 1-periodic in time. Moreover, the functions  $|u(t)|_2$  and  $|\nabla u(t)|_2$ , as well as quantities (5.42), are time-independent random constants.*

(iii) *We have the relations*

$$\begin{aligned} \mathbb{E} |\nabla u(t)|_2^2 &= \frac{D_0}{2} \quad \text{for all } t \geq 0, \\ \mathbb{E} |\nabla u(t)|_2^m &\leq C_m \quad \text{for all } t \geq 0, m \geq 1, \\ \int_{k-1}^k \|u(t)\|_2^2 &\leq \frac{D_1}{2} \quad \text{for all } k \geq 1, \end{aligned}$$

*where  $C_m > 0$  are some constants not depending on the limiting measure.*

The investigation of the inviscid limit was based essentially on the existence of two “good” integrals of motion for the limiting equation. In the case of Navier–Stokes equations, they are given by the energy and enstrophy. The additional set of infinitely many integrals given by the functions of vorticity enabled us to study further properties of the limiting measures, such as the

existence of density for finite dimensional functionals. A similar analysis can be carried out for other problems possessing those properties, e.g., the damped-driven Korteweg–de Vries equation. Moreover, due to complete integrability of the latter, it is possible to prove uniqueness of the limiting measure and to describe it as a stationary measure for a non-local stochastic PDE; see [KP08, Kuk10a, Kuk12].

### 5.2.6 Inviscid limit for the complex Ginzburg–Landau equation

The complex Ginzburg–Landau equation is an example for which we know the existence of only two integrals of motion, and this turns out to be sufficient to draw some nontrivial conclusions. They are, however, less complete than those obtained in the case of the Navier–Stokes system. We now briefly discuss the corresponding results.

Let us consider the following problem in a bounded domain  $Q \subset \mathbb{R}^d$ ,  $d \leq 4$ , with a smooth boundary  $\partial Q$ :

$$\dot{u} - (\nu + i)\Delta u + i\lambda|u|^2u = \sqrt{\nu}\eta(t, x), \quad u|_{\partial Q} = 0. \quad (5.64)$$

Here  $\nu$  and  $\lambda$  are positive parameters and  $\eta$  is a white noise force of the form (2.66), where  $\{e_j\}$  stands for a complete set of normalised real eigenfunctions for the Dirichlet Laplacian with eigenvalues  $\alpha_j$ ,  $\{\beta_j = \beta_j^+ + i\beta_j^-\}$  is a sequence of independent standard complex-valued Brownian motions, and  $\{b_j\} \subset \mathbb{R}$  is a sequence. In this case, as was mentioned in Section 3.5.5, the Cauchy problem for (5.64) is well posed in the space  $H_0^1(Q; \mathbb{C})$ , and almost every trajectory of solutions belongs to the space

$$L_{\text{loc}}^2(\mathbb{R}_+; H^2) \cap (W_{\text{loc}}^{1, \frac{4}{3}}(\mathbb{R}_+; L^{\frac{4}{3}}) + W_{\text{loc}}^{\alpha, 4}(\mathbb{R}_+; H^1)),$$

where  $\alpha \in (\frac{1}{4}, \frac{1}{2})$  is an arbitrary constant. Moreover, using the Bogolyubov–Krylov argument, one can prove that Eq. (5.64) has a stationary measure  $\mu_\nu$  for any  $\nu > 0$ ; cf. Sections 2.5. Let  $\{\mu_\nu\}$  be the corresponding family of measures on the space of trajectories.

Using essentially the same argument as in Sections 5.2.1 and 5.2.2, one can prove the following assertions.

**Tightness.** The family  $\{\mu_\nu\}$  is tight in  $H^{2-\varepsilon}$  and  $\{\mu_\nu\}$  is tight in the space

$$L_{\text{loc}}^2(\mathbb{R}_+; H^{2-\varepsilon}) \cap C(\mathbb{R}_+; H^{-\frac{d}{4}-\varepsilon}).$$

**Limiting process.** Any limiting measure  $\mu$  for the family  $\{\mu_\nu\}$  is concentrated on the space  $L_{\text{loc}}^2(\mathbb{R}_+; H^2) \cap W_{\text{loc}}^{\frac{4}{3}}(\mathbb{R}_+; L^{\frac{4}{3}})$ , and  $\mu$ -almost every curve  $v(t)$  of that space satisfies the nonlinear Schrödinger equation

$$\dot{v} - i\Delta v + i\lambda|v|^2v = 0.$$

Moreover, the quantities

$$\mathcal{H}_0(u) = \frac{1}{2} \int_D |u(x)|^2 dx, \quad \mathcal{H}_1(u) = \int_D \left( \frac{1}{2} |\nabla u(x)|^2 + \frac{\lambda}{4} |u(x)|^4 \right) dx$$

are preserved along any such trajectory, and their mean values admit some explicit bounds in terms of the coefficients  $b_j$ .

**Non-degeneracy.** Suppose that all numbers  $b_j$  are non-zero. Then the laws of  $\mathcal{H}_0(u)$  and  $\mathcal{H}_1(u)$  with respect to any limit point of the family  $\{\mu_\nu\}$  possess densities against the Lebesgue measure on the real line.

We refer the reader to the papers [KS04b, Shi11b] for an exact statement of the above results and their proofs and to [Kuk12] for some related results.

### 5.3 Relevance of the results for physics

Results of this chapter form a beginning of a rigorous mathematical theory of the space-periodic 2D turbulence in fluid stirred by a random force. Namely, here we study periodic flows of 2D fluid with high Reynolds numbers and finite energy. We are concerned with their stationary measures (i.e., the statistical equilibria). Denoting by  $\mu_\nu$  the unique stationary measure for fluid with viscosity  $\nu$ , we show that if  $\mu_\nu$  has a nontrivial limit as  $\nu \rightarrow 0$ , then the force must be proportional to  $\sqrt{\nu}$ . Accordingly, we take the force of the form<sup>10</sup> “ $\sqrt{\nu}$  times a random force  $\eta(t, x)$  independent of  $\nu$ ”:

$$\partial_t u + \nu Lu + B(u) = \sqrt{\nu} \eta(t, x), \quad \eta = \sum_{s \in \mathbb{Z}^2 \setminus 0} b_s \dot{\beta}_s(t) e_s(x). \quad (5.65)$$

Here  $\{e_s(x)\}$  is the usual trigonometric basis of the  $L_2$ -space of periodic solenoidal vector fields with zero mean value (see (2.29)),  $\dot{\beta}_s(t)$  are standard independent white noises, and  $b_s$  are real numbers such that

$$\mathcal{B}_0 = \sum b_s^2 < \infty, \quad \mathcal{B}_1 = \sum |s|^2 b_s^2 < \infty.$$

In this case, when  $\nu \rightarrow 0$ , the energy of the fluid stays of order one, while its Reynolds number grows like  $\nu^{-1}$ .

Equation (5.65) has a unique stationary measure  $\mu_\nu$ . The latter has limits as  $\nu$  goes to zero along sequences,  $\mu_0 = \lim_{\nu_j \rightarrow 0} \mu_{\nu_j}$ . We do not know if the limiting measure  $\mu_0$  is unique, but each  $\mu_0$  has a number of universal properties.

(i) The measure  $\mu_0$  is supported by the space  $H^2$ ,  $\mu_0(H^2) = 1$ , and

$$\mathbb{E}_{\mu_0} \|u\|_2^2 \leq \frac{1}{2} \mathcal{B}_1 < \infty. \quad (5.66)$$

It is space homogeneous if so is the force.

<sup>10</sup>Note that the Navier–Stokes equations with viscosity  $\nu$  and a random force  $\nu^a \eta(t, x)$  with any  $a \in \mathbb{R}$  reduces to (5.65) by suitable scaling of  $\nu$ ,  $u$ , and  $t$ ; see Section 5.2.4.

(ii) The  $\mu_0$ -averaged enstrophy  $\frac{1}{2}\mathbb{E}_{\mu_0} \int |\text{curl } u|^2 dx$  equals  $\frac{1}{4}\mathcal{B}_0$ , while the  $\mu_0$ -averaged energy  $\mathbb{E}_{\mu_0} \int |u|^2 dx$  has explicit lower and upper bounds. The random variables  $\exp(\kappa|u(x)|)$  and  $\exp(\kappa|\text{curl } u(x)|^{1/2})$  are  $\mu_0$ -integrable for any  $x$ , where the constant  $\kappa > 0$  does not depend on  $\mu_0$ .

(iii)  $\mu_0$  is an invariant measure for the free 2D Euler equation.

(iv) It is genuinely infinite-dimensional. Namely,  $\mu_0(K) = 0$  if  $K$  is a finite-dimensional set.

(v) It is  $\mu_0$ -unlikely that the energy  $E = \frac{1}{2} \int |u|^2 dx$  of the fluid is very small or very large. That is,

$$\mu_0\{E(u) \leq \delta\} \leq C\sqrt{\delta}, \quad \mu_0\{E(u) \geq \delta^{-1}\} \leq Ce^{-c\delta^{-1}} \quad \text{for all } \delta > 0,$$

where  $c$  and  $C$  are independent of  $\mu_0$ .

(vi) The stationary measures  $\mu_\nu, \nu > 0$ , satisfy infinitely many explicit algebraical relations. These relations are independent of  $\nu$  and depend only on two scalar characteristics  $\mathcal{B}_0$  and  $\mathcal{B}_1$  of the random force. We cannot pass to the limit in these relations to show that they remain true for  $\mu_0$ , but we use them to study  $\mu_0$ ; e.g., to derive the second assertion in (ii).

In Remark 5.2.19, we explained that the limiting measure(s)  $\mu_0$  will change if in Eq. (5.65) we replace the viscosity  $\nu Lu$  with a hyperviscosity  $\nu L^m u, m > 1$ .

Note that (iii) agrees with the belief (which goes back at least to Onsager [Ons49]) that the 2D Euler equation describes certain classes of 2D turbulence. Unfortunately, the equation does this in an implicit way since Euler-invariant measures form a large class and we do not know how to single out from it the measure  $\mu_0$ . Also note that (vi) indicates certain universality of the space-periodic 2D turbulence. Various universalities of 2D turbulence are often suggested in physical literature.

For  $k \geq 1$  denote by  $E_k$  the density of energy at a wave-number  $k$ . That is, for  $u(x) = \sum_{s \in \mathbb{Z}_0^2} u_s e_s(x)$  we define

$$E_k = (2C)^{-1} \sum_{k-C \leq |s| < k+C} \frac{1}{2} \mathbb{E}_{\mu_0} |u_s|^2,$$

where  $C \geq 1$  is of order one and we are the most interested in large  $k$ . Then

$$\mathbb{E}_{\mu_0} \|u\|_m^2 \sim \int_1^\infty E_k k^{2m} dk \quad \forall m. \tag{5.67}$$

The function  $k \mapsto E_k$  is called the *energy spectrum of turbulence*. Making bold assumption that  $E_k \sim k^{-r}$  for some  $r$ , we get from (i) that  $r \geq 5$ . We neglect the weak logarithmic divergence of the integral in (5.67) which occurs when  $r = 5$  (to remove the divergence one may correct the anzats for  $E_k$  by a suitable logarithmic factor). The exponent  $r = 5$  is distinguished due to relation (5.66),

which follows from the conservation of enstrophy in the 2D Euler equation. We conjecture that  $\mu_0(H^{2+\delta}) = 0$  for positive  $\delta$  and that

$$\mathbb{E}_{\mu_0} \|u\|_{2+\delta}^2 = \infty \quad \forall \delta > 0.$$

Then

$$E_k \sim k^{-5} \quad \text{for } k \gg 1, \quad (5.68)$$

modulo a possible logarithmic correction. This relation implies that periodic 2D turbulence exhibit a *direct cascade of energy*: even though the force  $\eta(t, x)$  in (5.65) is concentrated in low modes (and has a non-significant smooth component in high modes to guarantee the non-degeneracy), the energy spectrum decays only as  $k^{-5}$ .

It is interesting to compare our results and predictions with those of the heuristic theory of 2D turbulence originated by Batchelor [Bat82] (pp. 186–187), [Bat69] and Kraichnan [Kra67], and called below the BK theory. That theory is based on the assumption that the rate of dissipation of enstrophy  $\nu \mathbb{E}_{\mu_\nu} \|\text{curl } u\|_1^2$  converges to a finite non-zero limit  $\epsilon$  as  $\nu \rightarrow 0$ . Since  $\|\text{curl } u\|_1 = \|u\|_2$  for divergence-free vector fields, then for the space-periodic 2D turbulence we have  $\epsilon = 0 \cdot \mathbb{E}_{\mu_\nu} \|u\|_2^2$  which vanishes by (5.66). So the assumption fails. The explanation for this is simple: for the BK theory to hold, the 2D flow of the fluid should be not  $2\pi$ -periodic, but  $2\pi L$ -periodic, where  $\mathbb{N} \ni L \rightarrow \infty$  (it is not quite clear what relation should be imposed on  $\nu^{-1}$  and  $L$  when both of them go to infinity). The force  $\eta(t, x)$  remains essentially the same: most of its energy is supported by modes  $e_s(x)$ , where  $s$  is an integer of order one, while other modes  $e_{s'}(x)$ ,  $s' \in (\mathbb{Z}/L)^2$ , carry a tiny proportion of the energy, just to make the force non-degenerate with respect to the period  $2\pi L$ . In this large-volume setup the force  $\eta$  enters the right-hand side of the Navier–Stokes equations without any scaling factor, so the equations take the form

$$\partial_t u + \nu Lu + B(u) = \eta(t, x). \quad (5.69)$$

The BK theory predicts that for  $\nu \ll 1$  and  $L \gg 1$  the energy spectrum  $E_k$  behaves as follows:

$$E_k \sim k^{-3}, \quad 1 \ll k \ll k_+, \quad (5.70)$$

$$E_k \sim k^{-5/3}, \quad k_- \ll k \ll 1, \quad (5.71)$$

where  $k_+ \rightarrow \infty$  and  $k_- \rightarrow 0$  as  $\nu \rightarrow 0$  and  $L \rightarrow \infty$ . The theory also specifies  $\epsilon$ -dependent factors in front of  $k^{-3}$  and  $k^{-5/3}$  which we do not discuss. The asymptotics (5.70) differs from what we have in the space-periodic 2D turbulence, where the assumption  $E_k \sim k^{-r}$  implies that  $r \geq 5$  and we predict that  $r = 5$ . Presumably, this happens since the  $2\pi$ -periodic boundary conditions, imposed on the fluid, regularise it and make the direct cascade of energy weaker than the BK-cascade (5.70).<sup>11</sup>

<sup>11</sup>When asked if his theory of turbulence applies to 3D periodic flows, A. N. Kolmogorov said that he is not certain in that, since periodic flows may be “too regular” (we are thankful to A. V. Fursikov and M. I. Vishik for this recollection).

Relation (5.71), known as the *inverse cascade of energy*, has no analogy in the space-periodic turbulence, where  $k \geq 1$ . Some experts in turbulence believe that for limits (5.70) and (5.71) to exist when  $\nu \rightarrow 0$  and  $L \rightarrow \infty$ , the right-hand side of Eq. (5.69) has to be modified by the friction term  $-\tau u$  (sometimes called the *Ekman damping* in honor of Swedish oceanologist V. W. Ekman). The parameter  $\tau$  has to be sent to zero with  $\nu$  and  $L^{-1}$  (again, relation between these three quantities is not clear). See [Ber00] for an advanced version of the corresponding argument and see [BV11] for a detailed discussion.

So, in difference with the BK theory, the theory of space-periodic 2D turbulence, presented in this section, has a rigorous component (i) – (vi). It gives rise to a direct cascade of energy with an exponent  $\geq 5$ , which is predicted to be five, and differs from the BK-exponent 3 in (5.70). Further rigorous development of this theory by means of modern mathematical tools (including those presented in this book) seems to us more feasible than rigorous justification of the BK theory.

## Notes and comments

The existence of an inviscid limit of stationary measures (see Theorems 5.2.2 and 5.2.21) was proved by Kuksin [Kuk04]. The balance relations described in Section 5.1.2 were established by Kuksin and Penrose [KP05]. The co-area form of the balance relations (see Exercise 5.1.5), pointwise exponential estimates of Section 5.1.3, and the properties of the limiting measures described in Sections 5.2.1 – 5.2.4 were obtained by Kuksin [Kuk06b, Kuk08, Kuk07]. The inviscid limit for the complex Ginzburg–Landau equation was studied by Kuksin and Shirikyan [KS04b], and the local time technique used in Section 5.2.2 is taken from [Shi11b]. We refer the reader to the review papers [Kuk07, Kuk10b] for further references and a discussion of the inviscid limit for damped/driven integrable PDE's.





## Chapter 6

# Miscellanies

This chapter is mostly devoted to the 3D Navier–Stokes equations with random perturbations. We begin with the problem in thin domains and state a result on the convergence of the unique stationary distribution to a unique measure which is invariant under the flow of the limiting 2D Navier–Stokes system. We next turn to the 3D problem in an arbitrary bounded domain or a torus. We describe two different approaches for constructing Markov processes whose trajectories are concentrated on weak solutions of the Navier–Stokes system and investigate the large-time asymptotics of their trajectories. Finally, we discuss some qualitative properties of solutions in the case of perturbations of low dimension. Almost all the results of this chapter are presented without proofs.

### 6.1 3D Navier–Stokes system in thin domains

In this section, we present a result that justifies the study of 2D Navier–Stokes equations in the context of hydrodynamical turbulence. Namely, we study the 3D Navier–Stokes system in a thin domain and prove that, roughly speaking, if the domain is sufficiently thin, then the problem in question has a unique stationary measure, which attracts exponentially all solutions in a large ball and converges to a limiting measure invariant under the 2D dynamics. Moreover, when the width of the domain shrinks to zero, the law of a 3D solution converges to that of a 2D solution uniformly in time. The accurate formulation of these results requires some preliminaries from the theory of Navier–Stokes equations in thin domains. They are discussed in the first subsection. We next turn to the large-time asymptotics of solutions and the limiting behaviour of stationary measures and solutions.

### 6.1.1 Preliminaries on the Cauchy problem

Let  $Q_\varepsilon = \mathbb{T}^2 \times (0, \varepsilon) \subset \mathbb{R}^3$ , where  $\mathbb{T}^2$  stands for the two-dimensional torus<sup>1</sup> and  $\varepsilon > 0$  is a small parameter. We denote by  $x = (x_1, x_2, x_3)$  the space variable, so that  $(x_1, x_2)$  varies in  $\mathbb{T}^2$  while  $x_3$  belongs to the interval  $(0, \varepsilon)$ . Consider the 3D Navier–Stokes system

$$\dot{u} + \langle u, \nabla \rangle u - \nu \Delta u + \nabla p = \eta(t, x), \quad \operatorname{div} u = 0, \quad x \in Q_\varepsilon, \quad (6.1)$$

where  $u = (u_1, u_2, u_3)$  and  $p$  are unknown velocity field and pressure, and  $\eta$  denotes a random kick force of the form (2.65). Equations (6.1) are supplemented with the free boundary conditions in the vertical direction  $x_3$ :

$$u_3 = \partial_3 u_i = 0 \quad \text{for } x_3 = 0 \text{ and } \varepsilon, \quad (6.2)$$

where  $i = 1, 2$ . As in the 2D case with periodic or Dirichlet boundary conditions, problem (6.1), (6.2) can be reduced to an evolution equation in an appropriate functional space. Namely, denote by  $V_\varepsilon^2$  the space of divergence-free vector fields  $u \in H^2(Q_\varepsilon; \mathbb{R}^3)$  that satisfy the boundary conditions (6.2) and whose first and second components have zero mean value on  $Q_\varepsilon$ . Let  $H_\varepsilon$  and  $V_\varepsilon$  be the closures of  $V_\varepsilon^2$  in the spaces  $L^2(Q_\varepsilon)$  and  $H^1(Q_\varepsilon)$ , respectively, and let  $\Pi_\varepsilon$  be the orthogonal projection in  $L^2(Q_\varepsilon)$  to the closed subspace  $H_\varepsilon$ . The space  $H_\varepsilon$  is endowed with the usual  $L^2$  norm  $|\cdot|_\varepsilon$ , and  $V_\varepsilon$  is an Hilbert space with respect to the norm  $\|u\|_\varepsilon = |\nabla u|_\varepsilon$ , due to Poincaré's and Friedrichs' inequalities. Applying formally  $\Pi_\varepsilon$  to the first relation in (6.1), we get the nonlocal equation

$$\dot{u} + \nu L_\varepsilon u + B_\varepsilon(u) = \Pi_\varepsilon \eta(t, x). \quad (6.3)$$

Here  $L_\varepsilon$  is the Stokes operator in  $H_\varepsilon$  given on its domain  $V_\varepsilon^2$  by  $L_\varepsilon u = -\Pi_\varepsilon \Delta u$  and  $B_\varepsilon(u) = B_\varepsilon(u, u)$ , where  $B_\varepsilon : V_\varepsilon \rightarrow V_\varepsilon^*$  is a quadratic mapping defined for  $u, v \in V_\varepsilon^2$  by the relation  $B_\varepsilon(u, v) = \Pi_\varepsilon(\langle u, \nabla \rangle v)$ . Introduce the following orthogonal projections in  $L^2(Q_\varepsilon; \mathbb{R}^3)$ :

$$M_\varepsilon u = \left( \varepsilon^{-1} \int_0^\varepsilon u_1(x', y) dy, \varepsilon^{-1} \int_0^\varepsilon u_2(x', y) dy, 0 \right), \quad N_\varepsilon = \operatorname{Id} - M_\varepsilon,$$

where  $x' = (x_1, x_2)$ . It is straightforward to check that these projections also are continuous and orthogonal in  $V_\varepsilon$ , and their images are closed subspaces in both cases. Furthermore, recall that the spaces  $H$  and  $V$  were introduced at the end of Section 2.1.5 and note that they can be identified with the images of  $M_\varepsilon$  in  $H_\varepsilon$  and  $V_\varepsilon$ , respectively, by associating to each element  $(u_1, u_2)$  the vector function  $(u_1, u_2, 0)$ . The following result due to Raugel and Sell [RS93] (see also the paper [TZ96] by Temam and Ziane) enables one to construct global in time solutions of the homogeneous equation for initial data in a large ball, provided that  $\varepsilon > 0$  is sufficiently small.

<sup>1</sup>To simplify the presentation, we assume that the sides of the torus are equal. However, all the results remain true in the case of an arbitrary ratio of the sides.

**Theorem 6.1.1.** *Let  $R : (0, 1] \rightarrow \mathbb{R}_+$  be an arbitrary function such that*

$$\lim_{\varepsilon \rightarrow 0^+} \varepsilon^\theta R(\varepsilon) = 0 \quad \text{for some } \theta \in (0, \tfrac{1}{2}). \quad (6.4)$$

*Then there is  $\varepsilon_0 > 0$  depending on  $R(\cdot)$  such that, for any initial function  $u_0^\varepsilon \in V_\varepsilon$  satisfying the inequality  $\|u_0^\varepsilon\|_\varepsilon \leq R(\varepsilon)$ , Eq. (6.3) has a unique solution*

$$u^\varepsilon \in C(\mathbb{R}_+; V_\varepsilon) \cap L_{\text{loc}}^2(\mathbb{R}_+; V_\varepsilon^2)$$

*issued from  $u_0^\varepsilon$  at time  $t = 0$ . Moreover, if  $M_\varepsilon u_0^\varepsilon \rightarrow v_0$  as  $\varepsilon \rightarrow 0$  weakly in  $H$ , then for any  $T > 0$  we have*

$$\lim_{\varepsilon \rightarrow 0} M_\varepsilon u^\varepsilon = v \quad \text{in } C(0, T; H) \cap L^2(0, T; V), \quad (6.5)$$

*where  $v(t, x)$  is the solution of the 2D Navier–Stokes equation (2.19) with  $f \equiv 0$  issued from  $v_0$ .*

### 6.1.2 Large-time asymptotics of solutions

We first describe the random kick forces we are dealing with. For simplicity, we shall assume that  $\eta(t, \cdot)$  belongs to the image of the projection  $\Pi_\varepsilon$  for all  $t \geq 0$ , so that we can omit  $\Pi_\varepsilon$  from the right-hand side of (6.3). As was explained in Section 2.3, the Navier–Stokes system (6.1) with the random kick force (2.65) is equivalent to the discrete-time dynamical system

$$u_k = S_T^\varepsilon(u_{k-1}) + \eta_k, \quad (6.6)$$

where  $u_k = u(kT)$ ,  $S_t^\varepsilon$  denotes the resolving operator for the homogeneous problem (see Theorem 6.1.1), and the solutions are normalised by the condition of right-continuity at the points of the form  $kT$ . In the 3D case, there is an additional problem related to the fact that  $S_t$  may not be defined for all  $t \geq 0$ . However, it is possible to construct large subsets of  $V_\varepsilon$  on which the RDS (6.6) is well defined. Namely, for given positive functions  $a(\varepsilon)$  and  $b(\varepsilon)$ , we set

$$\mathcal{B}_\varepsilon = \{u \in V_\varepsilon : \|M_\varepsilon u\|_\varepsilon \leq a(\varepsilon), \|N_\varepsilon u\|_\varepsilon \leq b(\varepsilon)\}.$$

The following result is established in [CK08a] (see Proposition 4.1).

**Proposition 6.1.2.** *Let  $a(\varepsilon)$  and  $b(\varepsilon)$  be two positive functions such that (6.4) holds for  $R(\varepsilon) = a(\varepsilon) + b(\varepsilon)$  and*

$$\lim_{\varepsilon \rightarrow 0} \frac{\sqrt{\varepsilon} b^2(\varepsilon)}{a(\varepsilon)} = 0. \quad (6.7)$$

*Then for any  $T_0 > 0$  there are positive constants  $\varepsilon_0$  and  $c_0$  such that if*

$$\mathbb{P}\{\|M_\varepsilon \eta_k\|_\varepsilon \leq c_0 a(\varepsilon), \|N_\varepsilon \eta_k\|_\varepsilon \leq c_0 b(\varepsilon), k \geq 1\} = 1, \quad (6.8)$$

*then the RDS (6.6) is well defined on the set  $\mathcal{B}_\varepsilon$  for any  $T \geq T_0$  and  $\varepsilon \leq \varepsilon_0$ .*

This proposition enables one to define a family of Markov chains on  $\mathcal{B}_\varepsilon$  associated with (6.6). We denote by  $P_k^\varepsilon(u, \Gamma)$  its transition function and by  $\mathfrak{P}_k(\varepsilon)$  and  $\mathfrak{P}_k^*(\varepsilon)$  the corresponding Markov operators.

We now turn to the question of large-time behaviour of trajectories for (6.6) under some additional hypotheses on the kicks  $\eta_k$ . To formulate them, let us introduce a special orthogonal basis in the space  $V_\varepsilon$ . The Stokes operator  $L_\varepsilon$  is self-adjoint and has a compact resolvent, so the spaces  $H_\varepsilon$  and  $V_\varepsilon$  decompose into the direct sum of one-dimensional orthogonal subspaces spanned by the eigenfunctions of  $L_\varepsilon$ . Moreover, since the Stokes operator preserves the direct decomposition  $V_\varepsilon = M_\varepsilon V_\varepsilon \oplus N_\varepsilon V_\varepsilon$ , its eigenfunctions must belong to these subspaces, and it is easy to see that those belonging to  $M_\varepsilon V_\varepsilon$  are also eigenfunctions for the Stokes operator  $L$  on the 2D torus  $\mathbb{T}^2$  and are independent of  $\varepsilon$ . We denote them by  $e_j$  and write  $\lambda_j$  for the corresponding eigenvalues. As for the eigenfunctions belonging to  $N_\varepsilon V_\varepsilon$ , they do depend on  $\varepsilon$  and will be denoted by  $e_j^\varepsilon$ . We normalise the eigenfunctions by the condition

$$\|e_j\|_\varepsilon = \|e_j^\varepsilon\|_\varepsilon = \sqrt{\varepsilon} \quad \text{for all } j \geq 1, \quad (6.9)$$

whence it follows, in particular, that the norm in  $V$  of  $e_j$  considered as a vector field on  $\mathbb{T}^2$  is equal to 1.

We shall assume that the kicks  $\eta_k$  entering the random force  $\eta$  of the form (2.65) can be written as

$$\eta_k(x) = \eta_k^\varepsilon(x) = \sum_{j=1}^{\infty} b_j \xi_{jk} e_j(x) + \sum_{j=1}^{\infty} d_j \zeta_{jk} e_j^\varepsilon(x), \quad (6.10)$$

where  $b_j$  and  $d_j$  are non-negative numbers such that

$$\mathfrak{B} = \sum_{j=1}^{\infty} b_j^2 < \infty, \quad \mathfrak{D} = \sum_{j=1}^{\infty} d_j^2 < \infty,$$

and  $\{\xi_{jk}\}$  and  $\{\zeta_{jk}\}$  are two independent sequences of independent random variables whose laws possess densities  $p_j(r)$  and  $q_j(r)$  with respect to the Lebesgue measure on the real line. Furthermore, the functions  $p_j$  and  $q_j$  have bounded total variation and their supports are subsets of  $[-1, 1]$  containing the point  $r = 0$ . Note that, in view of (6.9), the  $V_\varepsilon$ -norm of the kicks satisfies the inequality

$$\|\eta_k\|_\varepsilon^2 \leq \sum_{j=1}^{\infty} (b_j^2 \|e_j\|_\varepsilon^2 + d_j^2 \|e_j^\varepsilon\|_\varepsilon^2) = \varepsilon(\mathfrak{B} + \mathfrak{D}).$$

Thus, by Proposition 6.1.2, the RDS (6.6) with the above hypotheses on  $\eta_k$  is well defined on  $\mathcal{B}_\varepsilon$ , provided that the functions  $a$  and  $b$  are minorised  $C\sqrt{\varepsilon}$  with a large constant  $C > 0$  and satisfy relations (6.4) and (6.7) in which  $R = a + b$ .

**Theorem 6.1.3.** *Let the above-mentioned hypotheses on  $\eta_k$  be satisfied and let  $a(\varepsilon)$  and  $b(\varepsilon)$  be two functions such that (6.7) holds and*

$$C_1\sqrt{\varepsilon} \leq a(\varepsilon) \leq C_2\sqrt{\varepsilon} \left(\ln \frac{1}{\varepsilon}\right)^\sigma, \quad C_1\sqrt{\varepsilon} \leq b(\varepsilon) \leq C_2\left(\ln \frac{1}{\varepsilon}\right)^{\sigma/2},$$

where  $0 \leq \sigma < \frac{1}{2}$  and  $C_1 > 0$  is sufficiently large. Then there is a constant  $\varepsilon_0 > 0$  and an integer  $N \geq 1$  not depending on  $\varepsilon$  such that the following assertions are true for  $0 < \varepsilon \leq \varepsilon_0$ , provided that

$$b_j \neq 0 \quad \text{for } j = 1, \dots, N : \quad (6.11)$$

**Existence and uniqueness.** *The RDS (6.6) has a unique stationary measure  $\mu_\varepsilon$ .*

**Exponential mixing.** *There are positive constants  $\alpha_\varepsilon$  and  $C_\varepsilon$  such that for any  $\lambda \in \mathcal{P}(\mathcal{B}_\varepsilon)$  we have*

$$\|\mathfrak{P}_k^*(\varepsilon)\lambda - \mu_\varepsilon\|_L^* \leq C_\varepsilon e^{-\alpha_\varepsilon k} \quad \text{for } k \geq 0,$$

where  $\|\cdot\|_L^*$  stands for the dual-Lipschitz norm on  $\mathcal{P}(V_\varepsilon)$ .

We refer the reader to Section 5 in [CK08a] for a proof of this result. Note also that if the random force  $\eta$  is more regular in  $x$ , then the stationary measure is supported by a space of smoother functions and the convergence to it holds in a stronger norm. Furthermore, under some additional assumptions, the constants  $\alpha_\varepsilon$  and  $C_\varepsilon$  may be chosen to be independent of  $\varepsilon$ .

### 6.1.3 The limit $\varepsilon \rightarrow 0$

We now turn to the main result of this section. Recall that the 2D Navier–Stokes system on the torus  $\mathbb{T}^2$  perturbed by a random force of the form (2.65) gives rise to a Markov chain in the space  $H$ ; see Section 2.3. In view of Theorem 3.2.9, if the kicks have the form (6.10) with  $d_j = 0$  for all  $j \geq 1$  and the non-degeneracy condition 6.11 is fulfilled, then the Markov chain has a unique stationary measure  $\mu$ . The following theorem established Chueshov and Kuksin [CK08a] shows that the averaged horizontal component of the unique stationary measure for (6.6) can be approximated by  $\mu$ .

**Theorem 6.1.4.** *Under the hypotheses of Theorem 6.1.3, if (6.11) holds with a sufficiently large  $N$ , then*

$$(M_\varepsilon)_* \mu_\varepsilon \rightarrow \mu \quad \text{as } \varepsilon \rightarrow 0, \quad (6.12)$$

where the convergence holds in the weak topology of  $\mathcal{P}(H)$ . If, in addition,  $\sum_j \lambda_j^2 b_j^2 < \infty$ , then (6.12) holds in the weak topology of  $\mathcal{P}(V^{2-\delta})$  for any  $\delta > 0$ .

Theorem 6.1.4 enables one to strengthen convergence (6.5) for solutions of (6.1), showing that it holds uniformly in time for an appropriate topology. Namely, we have the following result (cf. Corollary 4.3.2).

**Corollary 6.1.5.** *Under the hypotheses of Theorem 6.1.4, let  $\{u_0^\varepsilon\} \in V_\varepsilon$  be a family of initial functions such that  $\|u_0^\varepsilon\|_\varepsilon \leq C\sqrt{\varepsilon}$  for some constant  $C > 0$  and  $M_\varepsilon u_0^\varepsilon \rightarrow v_0$  weakly in  $H$ , where  $v_0 \in V$ . Let  $u^\varepsilon(t)$  be a solution of (6.1), (6.2)*

issued from  $u_0^\varepsilon$  and let  $v(t)$  be a solution of Eq. (2.76) with  $h \equiv 0$  on  $\mathbb{T}^2$  such that  $v(0) = v_0$ . Then

$$\sup_{t \geq 0} \|\mathcal{D}(M_\varepsilon u^\varepsilon(t)) - \mathcal{D}(v(t))\|_L^* \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0, \quad (6.13)$$

where  $\|\cdot\|_L^*$  stands for the dual-Lipschitz distance over  $H$ .

We note that the condition  $\|u_0^\varepsilon\|_\varepsilon \leq C\sqrt{\varepsilon}$  is not a smallness assumption for  $u_0^\varepsilon$  since it signifies not that the vector field  $u_0^\varepsilon(x)$  is small, but that the volume of the domain  $Q_\varepsilon$  is of order  $\varepsilon$ . A proof of the above results can be found in Section 5.3 of [CK08a]. Let us also mention that convergence (6.13) holds for the dual-Lipschitz distance over  $V^{2-\delta}$  if the coefficients  $b_j$  satisfy the additional decay condition of Theorem 6.1.4.

For the 3D stochastic Navier–Stokes system in  $Q_\varepsilon$  with a white in time random force  $\eta(t, x)$ , the argument of this section does not work, because Proposition 6.1.2 is no longer true, as domains of the form  $\mathcal{B}_\varepsilon$  are not invariant for the corresponding RDS. However, if we smooth out the stochastic system (6.1) by replacing there the bilinear term  $B(u)$  with  $B((I - \alpha\Delta)^{-1}u, u)$ , where  $\alpha > 0$  is arbitrary, then we get a well-posed stochastic PDE, and the corresponding RDS converges weakly to trajectories of (6.1) as  $\alpha \rightarrow 0^+$ . An analogue of Theorem 6.1.3 is valid for this system with any positive  $\alpha$ , and an analogue of convergence (6.12) holds true as  $\alpha \rightarrow 0^+$  and  $\varepsilon \rightarrow 0$  (in any order); see [CK08b].

### Relevance of the results for physics

The results of this section show that, in the thin domain  $\mathbb{T}^2 \times (0, \varepsilon)$ , non-isotropic 3D turbulence described by the Navier–Stokes system with a *not to big* vertical component may be well approximated by the 2D turbulence in  $\mathbb{T}^2$ . The same results holds also for the non-isotropic 3D turbulence in the thin spherical layer  $\mathbb{S}^2 \times (0, \varepsilon)$ .

Figure 6.1: Motion of the atmosphere above Madeira island. On the scales of order the island’s size, the motion clearly looks two-dimensional; cf. Figure 1 and Section 6.1.

Figure 6.2: Jupiter Red Spot. This is a huge 2D structure in the Jovian atmosphere (twice bigger than the Earth), persistent for more than 180 years. It is believed that the motion inside the spot is driven by the 2D turbulence, cf. Figure 1.

## 6.2 Ergodicity and Markov selection

The aim of this section is to discuss some results on 3D stochastic Navier–Stokes equations. Since the corresponding results are rather technical and certainly

far from their final form, to make them more transparent, we begin with the case of stochastic differential equations in  $\mathbb{R}^d$  (SDE). We use a simple example to explain informally two fundamental results: martingale problem and Markov selection. We next turn to 3D Navier–Stokes equations and discuss two different approaches for constructing Markov processes concentrated on solutions and proving a mixing property in the total variation distance.

### 6.2.1 Finite-dimensional stochastic differential equations

Let us consider the SDE

$$dx_t = b(x_t)dt + dw_t, \quad x_t \in \mathbb{R}^d, \quad (6.14)$$

where  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a continuous function and  $w_t$  is an  $\mathbb{R}^d$ -valued standard Brownian motion. If  $b$  is globally Lipschitz continuous, then for any  $x \in \mathbb{R}^d$  Eq. (6.14) has a unique solution defined on the positive half-line and satisfying the initial condition

$$x_0 = x, \quad (6.15)$$

and the family of all solutions form a Markov process  $(x_t, \mathbb{P}_x)$  in  $\mathbb{R}^d$  with the transition function  $P_t(x, \Gamma) = \mathbb{P}_x\{x_t \in \Gamma\}$ . Itô's formula implies that if  $f$  is a  $C^2$ -smooth function with bounded derivatives, then

$$f(x_t) = f(x_0) + \int_0^t Lf(x_r) dr + \int_0^t \langle \nabla f(x_r), dw_r \rangle, \quad (6.16)$$

where we set

$$L = \frac{1}{2}\Delta + \langle b(x), \nabla \rangle.$$

The operator  $L$  is called the *generator* corresponding to the SDE (6.14). Let us denote by  $\mathcal{F}_t$  the filtration associated with the Markov process and apply the condition expectation  $\mathbb{E}_x(\cdot | \mathcal{F}_s)$  to (6.16). Recalling that the stochastic integral defines a martingale, we obtain

$$\begin{aligned} \mathbb{E}_x \left( f(x_t) - \int_0^t Lf(x_r) dr \mid \mathcal{F}_s \right) &= f(x) + \mathbb{E}_x \left( \int_0^t \langle \nabla f(x_r), dw_r \rangle \mid \mathcal{F}_s \right) \\ &= f(x) + \int_0^s \langle \nabla f(x_r), dw_r \rangle \\ &= f(x_s) - \int_0^s Lf(x_r) dr, \end{aligned}$$

where the equality holds with  $\mathbb{P}_x$ -probability 1. Hence, we see that the process

$$M_t^f = f(x_t) - \int_0^t Lf(x_r) dr, \quad t \geq 0, \quad (6.17)$$

is a martingale with respect to the filtration  $\mathcal{F}_t$  and the probability  $\mathbb{P}_x$ , with an arbitrary  $x \in \mathbb{R}^d$ . The following proposition shows that the latter condition

is sufficient for the existence of a weak solution for (6.14), (6.15). Let  $\Omega$  be the Fréchet space  $C(\mathbb{R}_+; \mathbb{R}^d)$  with its Borel  $\sigma$ -algebra and let  $\{x_t(\omega)\}_{t \geq 0}$  be the canonical process.

**Proposition 6.2.1.** *Let  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a bounded continuous function and let  $\mathbb{P}$  be a probability measure on  $\Omega$  such that  $M_t^f$  is a continuous martingale with respect to  $\mathbb{P}$  for any  $C^2$ -function  $f$  with bounded derivatives. Then the process  $x_t - x_0 - \int_0^t b(x_r) dr$  is an  $\mathbb{R}^d$ -valued standard Brownian motion. In particular, problem (6.14), (6.15) has a weak solution defined on  $\mathbb{R}_+$ .*

We refer the reader to Section 5.4.B of the book [KS91] for the proof of a more general result. In what follows, any probability measure  $\mathbb{P}$  satisfying the hypotheses of Proposition 6.2.1 is called a *solution of the martingale problem associated with  $L$* . What has been said above implies that problem (6.14), (6.15) has a weak solution associated with an initial measure  $\mu \in \mathcal{P}(\mathbb{R}^d)$  if and only if the martingale problem associated with its generator  $L$  has a solution  $\mathbb{P}$  whose restriction at  $t = 0$  coincides with  $\mu$ . Moreover, the uniqueness of the law for a weak solution is equivalent to the uniqueness of a solution for the martingale problem with a given initial measure. We now study a sufficient condition for the latter property.

Let  $\mathbb{P}$  be a solution of the martingale problem associated with  $L$ , let  $x_t$  be the canonical process, and let  $w_t$  be the Brownian motion constructed in Proposition 6.2.1. Let  $u(t, x)$  be an arbitrary function having continuous partial derivatives  $\partial_t^j \partial_x^\alpha u$  for  $2j + |\alpha| \leq 2$  that are bounded in the strip  $[0, T] \times \mathbb{R}^d$  for some  $T > 0$ . Then, applying Itô's formula to the process  $y_t = u(T - t, x_t)$ , we derive

$$y_t - y_0 = \int_0^t (-\partial_t u + Lu)(T - r, x_r) dr + \int_0^t \langle u(T - r, x_r), dw_r \rangle.$$

Assuming now that  $u$  is the solution<sup>2</sup> of the Cauchy problem for *Kolmogorov's backward equation*

$$\partial_t u = Lu, \quad u(0, x) = f(x), \quad x \in \mathbb{R}^d, \quad t > 0, \quad (6.18)$$

where  $f \in C_0^\infty(\mathbb{R}^d)$  is a given function, we see that the process  $y_t$  is a martingale under  $\mathbb{P}$ . In particular, taking the expectation with respect to  $\mathbb{P}$ , we obtain

$$\mathbb{E} u(T - t, x_t) = \mathbb{E} y_t = \mathbb{E} y_0 = \mathbb{E} u(T, x_0) \quad \text{for any } t \in [0, T].$$

Setting  $t = T$ , we see that

$$\mathbb{E} f(y_T) = \mathbb{E} u(0, x_T) = \mathbb{E} f(x_T) = \mathbb{E} u(T, x_0).$$

The right-hand side of this relation depends only on the function  $f$  and the law of the restriction of  $\mathbb{P}$  at time  $t = 0$ . Since  $C_0^\infty(\mathbb{R}^d)$  is a determining

<sup>2</sup>Solutions with the above-mentioned regularity exist, for instance, if  $b$  is Hölder continuous; see Chapter 8 of [Kry96].



family for the Borel measures on  $\mathbb{R}^d$ , we conclude that the law of  $x_T$  under a solution of the martingale problem depends only on the initial law  $\mathbb{P}|_{t=0}$ . Further analysis shows that the law of the canonical process  $\{x_t\}$  under  $\mathbb{P}$  also depends only on  $\mathbb{P}|_{t=0}$ , provided that problem (6.18) has a solution satisfying the above-mentioned boundedness condition for any  $T > 0$ . More precisely, we have the following result due to Stroock and Varadhan; see Sections 5.4.E of the book [KS91].

**Theorem 6.2.2.** *Let  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a continuous function such that for any  $f \in C_0^\infty(\mathbb{R}^d)$  problem (6.18) has a solution  $u(t, x)$  whose derivatives  $\partial_t^j \partial_x^\alpha u$  with  $2j + |\alpha| \leq 2$  are bounded in the strip  $[0, T] \times \mathbb{R}^d$  for any  $T > 0$ . Then for every  $x \in \mathbb{R}^d$  the martingale problem associated with  $L$  has at most one solution  $\mathbb{P}$  such that  $\mathbb{P}|_{t=0} = \delta_x$ .*

It is a “natural” idea to try to apply the above approach to 3D Navier–Stokes equations, for which well-posedness is not known to hold. Unfortunately, the existing methods do not allow one to construct a sufficiently regular solution for the infinite-dimensional analogue of problem (6.18), and the possibility of carrying out the Stroock–Varadhan program for 3D Navier–Stokes equations is unclear. Nevertheless, one can prove that Kolmogorov’s equation possesses a solution with weak regularity properties, and this turns out to be sufficient for building a Markov family concentrated on trajectories of Navier–Stokes equations. This approach, together with an investigation of mixing properties, was realised by Da Prato, Debussche, and Odasso in the series of papers [DD03, DO06, Oda07] and is presented briefly in Section 6.2.2.

We now turn to the case in which the uniqueness of the law for the SDE (6.14) is unknown. In this situation, Krylov [Kry73] suggested a general scheme for constructing a *Markov selection*—a family of strong Markov processes concentrated on martingale solutions of (6.14). We now describe his approach following essentially the presentation in [SV79] and confining ourselves to the Markov property. The strong Markov property of the selection is more delicate and will not be discussed here.

As before, we assume that  $b(x)$  is a bounded continuous function and denote by  $L$  the generator corresponding to (6.14). Given  $s \geq 0$  and  $x \in \mathbb{R}^d$ , we shall say that a probability measure  $\mathbb{P}$  on the canonical space  $\Omega$  is a *solution of the martingale problem for  $L$  starting from  $(s, x)$*  if  $\mathbb{P}\{x_t = x \text{ for } 0 \leq t \leq s\} = 1$  and the process  $M_t^f$  defined by (6.17) is a martingale after time  $s$  for any  $f \in C_0^\infty(\mathbb{R}^d)$ . Let us denote by  $\mathcal{C}(s, x)$  the set of all solutions to the martingale problem for  $L$  starting from  $(s, x)$ . Thus,  $\mathcal{C}(s, x)$  is a subset in the space  $\mathcal{P}(\Omega)$ , which is endowed with the dual-Lipschitz distance and the topology of weak convergence. We shall write  $\mathcal{C}(x)$  instead of  $\mathcal{C}(0, x)$ . Since the drift coefficient  $b$  does not depend on time, it is clear that  $\mathcal{C}(s, x)$  can be obtained from the set  $\mathcal{C}(x)$  by “shifting” it to the half-line  $[s, +\infty)$  and “gluing” to the resulting measure the Dirac mass at the function equal to  $x$  for  $0 \leq t \leq s$ . To formulate accurately this and some other properties of  $\mathcal{C}(s, x)$ , we need the following lemma, whose proof in a more general setting can be found in [SV79] (see Section 6.1).

**Lemma 6.2.3.** *Let  $s \geq 0$  and let  $Q_\omega : \Omega \rightarrow \mathcal{P}(\Omega)$  be a random probability measure such that  $Q_\omega(\Gamma)$  is  $\mathcal{F}_s$ -measurable for any  $\Gamma \in \mathcal{B}(\Omega)$  and*

$$Q_\omega(\{x_t = \omega(s) \text{ for } 0 \leq t \leq s\}) = 1 \quad \text{for all } \omega \in \Omega.$$

*Then for any  $P \in \mathcal{P}(\Omega)$  there is a unique measure  $P \otimes_s Q_\omega \in \mathcal{P}(\Omega)$  such that*

$$(P \otimes_s Q_\omega)(B) = P(B), \quad (6.19)$$

$$(P \otimes_s Q_\omega)(B \cap \Gamma) = \int_B Q_\omega(\Gamma) P(d\omega) \quad (6.20)$$

*for any  $B \in \mathcal{F}_s$  and  $\Gamma \in \mathcal{F}_s^+$ , where  $\mathcal{F}_s^+$  stands for the  $\sigma$ -algebra generated by the canonical process on  $[s, +\infty)$ .*

In other words, if  $Q_\omega$  is a random probability measure that is measurable with respect to  $\mathcal{F}_s$  and whose restriction to  $\mathcal{F}_s$  coincides with the Dirac mass at the function identically equal to  $\omega(s)$  on the interval  $[0, s]$ , then there is a unique measure in  $\mathcal{P}(\Omega)$  such that its restriction to  $\mathcal{F}_s$  coincides with  $P$  (see (6.19)) and the restriction to  $\mathcal{F}_s^+$  of its regular conditional probability given  $\mathcal{F}_s$  coincides with  $Q_\omega|_{\mathcal{F}_s^+}$  (see (6.20)). Note also that the uniqueness is a straightforward consequence of relation (6.20), the monotone class technique, and the fact that  $\mathcal{F}$  is generated by the  $\pi$ -system of the sets of the form  $B \cap \Gamma$  with  $B \in \mathcal{F}_s$  and  $\Gamma \in \mathcal{F}_s^+$ .

The time-independence of  $b$  imply that if  $P \in \mathcal{C}(s, x)$  for some  $s > 0$ , then  $(\theta_s)_*P \in \mathcal{C}(x)$ , where  $\theta_s$  stands for the shift operator on  $\Omega$ . Conversely, if  $P \in \mathcal{C}(x)$ , then  $\delta_x \otimes_s ((\theta_{-s})_*P) \in \mathcal{C}(s, x)$ , where  $\delta_x$  is the Dirac mass at the function identically equal to  $x$  and, with a slight abuse of notation, we denote by  $(\theta_{-s})_*P$  the measure defined by  $((\theta_{-s})_*P)(\Gamma) = P(\theta_s\Gamma)$ . That is, the mapping

$$P \mapsto \delta_x \otimes_s ((\theta_{-s})_*P) \in \mathcal{C}(s, x)$$

defines a bijection between  $\mathcal{C}(x)$  and  $\mathcal{C}(s, x)$ .

The proposition below establishes some further properties of the sets  $\mathcal{C}(s, x)$ . It is the first of the two key ingredients for the construction of a Markov selection for (6.14).

**Proposition 6.2.4.** *The family  $\{\mathcal{C}(s, x), s \geq 0, x \in \mathbb{R}^d\}$  possesses the following properties.*

- (a) *The set  $\mathcal{C}(s, x)$  is compact in  $\mathcal{P}(\Omega)$  for any  $s \geq 0$  and  $x \in \mathbb{R}^d$ .*
- (b) *The mapping  $(s, x) \mapsto \mathcal{C}(s, x)$  is lower-semicontinuous. That is, if  $s_n \rightarrow s$ ,  $x_n \rightarrow x$ ,  $P_n \in \mathcal{C}(s_n, x_n)$ , and  $P_n \rightarrow P$  in  $\mathcal{P}(\Omega)$ , then  $P \in \mathcal{C}(s, x)$ . In particular, the mapping  $(s, x) \mapsto \mathcal{C}(s, x)$  defines a random compact set in  $\mathcal{P}(\Omega)$ , with the underlying space  $\mathbb{R}_+ \times \mathbb{R}^d$ .*
- (c) *Let  $P \in \mathcal{C}(x)$  and let  $P_s(\omega, \Gamma)$  be the regular conditional probability of  $P$  given  $\mathcal{F}_s$ . Then there is a  $P$ -null set  $N \in \mathcal{F}_s$  such that<sup>3</sup>*

$$\delta_{\omega(s)} \otimes_s P_s(\omega, \cdot) \in \mathcal{C}(s, \omega(s)) \quad \text{for } \omega \notin N. \quad (6.21)$$

<sup>3</sup>Note that the second factor  $P_s(\omega, \cdot)$  in the product  $\otimes_s$  of (6.21) is considered as a constant measure.

- (d) Let  $P \in \mathcal{C}(x)$  and let  $Q : \Omega \rightarrow \mathcal{P}(\Omega)$  be an  $\mathcal{F}_s$ -measurable function such that  $\delta_{\omega(s)} \otimes_s Q_{\omega(s)} \in \mathcal{C}(s, \omega(s))$  for all  $\omega \in \Omega$ . Then  $P \otimes_s Q \in \mathcal{C}(x)$ .

We do not give a proof of this proposition, leaving it to the reader as an exercise or referring to [SV79, Lemma 12.2.1] for a more general result. Let us, however, clarify informally the meaning of properties (a) – (d) and give some hints for their proof.

The compactness of  $\mathcal{C}(s, x)$  is almost obvious because, due to the boundedness of  $b$ , any measure in  $\mathcal{C}(s, x)$  is concentrated on a set of functions whose  $C^1$ -norm is bounded with high probability on any finite interval. The lower-semicontinuity of  $\mathcal{C}(s, x)$  means essentially that the limit of a sequence of solutions is again a solution. Since the noise is additive, this property is a straightforward consequence of Skorohod's embedding theorem. Property (c) says, roughly speaking, that if  $\{y_t(\omega), t \geq 0\}$  is a weak solution of (6.14) starting from  $(0, x)$ , then its restriction to the half-line  $[s, +\infty)$  is a weak solution of (6.14) starting from  $(s, y_s(\omega))$ . Finally, property (d) is a formalisation of the claim that if we cut a solution starting from  $(0, x)$  at a time  $s$  and then continue it as another solution, then the resulting process is again a solution starting from  $(0, x)$ , provided that the two solutions agree at time  $s$ .

The second key ingredient for constructing a Markov selection is the following theorem, which says that one can choose elements in  $\mathcal{C}(x)$  that depend on  $x$  in a measurable manner and satisfy the Markov property.

**Theorem 6.2.5.** *Let  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a bounded continuous function, let  $\mathcal{C}(x)$  be the set of solutions for martingale problem starting from  $(0, x)$ , and let  $\{x_t\}_{t \geq 0}$  be the canonical process on  $\Omega$ . Then there exists a measurable mapping  $x \mapsto \mathbb{P}_x$  from  $\mathbb{R}^d$  to  $\mathcal{P}(\Omega)$  such that  $\mathbb{P}_x \in \mathcal{C}(x)$  for all  $x$  and  $(x_t, \mathbb{P}_x)$  is a homogeneous family of Markov processes.*

We refer the reader to Krylov's original paper [Kry73] or to Chapter 12 of the book [SV79] for a proof of this theorem and some further results. Note that the Markov process constructed in Theorem 6.2.5 may not satisfy the Feller property. For instance, the one-dimensional equation  $\dot{x} = \text{sgn}(x)|x|^{1/2}$  has no Feller selection; see Exercise 12.4.2 in [SV79]. Furthermore, not every selection  $x \mapsto \mathcal{C}(x)$  will produce a Markov family, as is easily seen by studying the equation  $\dot{x} = |x|^{1/2}$ .

The above scheme can be carried out in the case of 3D Navier–Stokes equations with additive noise. There is, however, a number of essential difficulties, which result in that one can establish only almost sure Markov property. This question, as well as the problem of mixing under a non-degeneracy assumption on the noise, was analysed in the papers of Flandoli and Romito [FR08, Rom08] and is discussed in Section 6.2.3.

## 6.2.2 Da Prato–Debussche–Odasso theorem

Let us consider the 3D Navier–Stokes equations with a white noise perturbation. To simplify the presentation, we assume that the space variable  $x$  belongs to

the 3D torus, even though most of the results in this and next subsections remain valid for any bounded domain with smooth boundary. Projecting to the space  $H$ , we get the equation

$$\dot{u} + \nu Lu + B(u) = h(x) + \partial_t \zeta(t, x), \quad \zeta(t, x) = \sum_{s \in \mathbb{Z}_0^3} b_s \beta_s(t) e_s(x), \quad (6.22)$$

where  $h \in H$  is a deterministic function,  $b_s \in \mathbb{R}$  are some constants such that

$$\mathfrak{B} = \sum_{s \in \mathbb{Z}_0^3} b_s^2 < \infty, \quad (6.23)$$

$\{\beta_s\}$  is a family of independent standard Brownian motions, and  $\{e_s\}$  is the normalised trigonometric basis in  $H$  (in the 3D case); cf. Sections 2.1.4, 2.1.5, and 2.4. Using the methods of the deterministic theory and treating the equation pathwise, one can prove the existence of a global weak solution for (6.22) satisfying the initial condition

$$u(0, x) = u_0(x), \quad (6.24)$$

where  $u_0 \in H$  is a given function. Indeed, denoting by  $z(t)$  a solution of the stochastic Stokes equation (2.100) issued from zero and writing  $u = v + z$ , we see that  $v$  must be a solution of problem (2.111), (2.112). The regularity of  $z$  guaranteed by Proposition 2.4.2 turns out to be sufficient to prove the existence of a weak solution  $v$  with probability 1. The resulting process  $u$  is a weak solution of the original problem, and further analysis shows that an energy inequality holds for it. We thus obtain the following result established in [BT73, VF88]; see also [CC92, FG95] for the case of a multiplicative noise.

**Proposition 6.2.6.** *Under the above hypotheses on  $h$  and  $\eta$ , for any  $u_0 \in H$  there is a random process  $u : \mathbb{R}_+ \rightarrow H$  whose almost every trajectory belongs to the space*

$$L_{\text{loc}}^\infty(\mathbb{R}_+; H) \cap L_{\text{loc}}^2(\mathbb{R}_+; V) \cap C(\mathbb{R}_+, V^{-r}) \quad (6.25)$$

with any  $r > 0$  and satisfies Eq. (2.99), where the equality holds in  $C(\mathbb{R}_+; V^*)$ . Moreover,

$$\mathbb{E} \left( |u(t)|_2^2 + \nu \int_0^t |\nabla u(s)|_2^2 ds \right) \leq |u_0|_2^2 + (\mathfrak{B} + C|h|_2^2) t, \quad t \geq 0, \quad (6.26)$$

where  $C > 0$  is a constant depending only on  $\nu$ .

The question of uniqueness of a weak solution remains open. However, under strong non-degeneracy hypotheses on the noise, it is possible to construct a Markov process concentrated on weak solutions of (6.22) and to prove for it the property of exponential mixing. To state the corresponding results, we first reformulate the above proposition on the existence of a solution in different terms.

For technical reasons, we consider solutions of weaker regularity than those constructed in Proposition 6.2.6. Let us fix a small constant  $r > 0$  and introduce the Fréchet space (cf. (6.25))

$$\Omega = L_{\text{loc}}^2(\mathbb{R}_+; V^{1-r}) \cap C(\mathbb{R}_+, V^{-r}), \quad (6.27)$$

endowed with a natural metric. Let  $\mathcal{F}$  be the Borel  $\sigma$ -algebra on  $\Omega$  and let  $\mathcal{F}_t$  be the natural filtration on  $\Omega$ . The law  $\mathbb{P}$  of the weak solution for (6.22), (6.24) constructed in Proposition 6.2.6 is a probability measure on  $(\Omega, \mathcal{F})$  possessing the following properties:

- (i)  $\mathbb{P}(\{\omega \in \Omega : \omega(0) = u_0\}) = 1$ ;
- (ii) The  $\mathbb{P}$ -law of the process<sup>4</sup>

$$M_t = \omega(t) - \omega(0) + \int_0^t (\nu L\omega + B(\omega)) ds - ht, \quad t \geq 0, \quad (6.28)$$

coincides with that of  $\zeta$ .

In what follows, any probability measure  $\mathbb{P}$  on  $(\Omega, \mathcal{F})$  satisfying these properties is called a *martingale solution* for problem (6.22), (6.24). Similarly, a probability measure  $\mathbb{P}$  on  $(\Omega, \mathcal{F})$  is called a *stationary martingale solution* for Eq. (6.22) if the canonical process is stationary under  $\mathbb{P}$  and property (ii) mentioned above holds for it.

Let us assume that the coefficients of the random force  $\zeta$  entering Eq. (6.22) satisfy the inequality

$$c|s|^{-3+\delta} \leq |b_s| \leq C|s|^{-\frac{5}{2}-\delta} \quad \text{for all } s \in \mathbb{Z}_0^3, \quad (6.29)$$

where  $C$ ,  $c$ , and  $\delta$  are positive constants. This means, roughly speaking, that the trajectories of  $\zeta$  have regularity higher than  $H^{5/4}$ , but lower than  $H^{3/2}$ . The following result shows that it is possible to construct a family of martingale solutions which satisfy a Markov property.

**Theorem 6.2.7.** *Let  $h \in V$  and let  $\zeta$  be a random force for which (6.29) holds. Then there is a measurable mapping  $u_0 \mapsto \mathbb{P}_{u_0}$  from  $V^2$  to the space of probability measures on  $(\Omega, \mathcal{F})$  with the weak topology such that the following properties hold.*

- (i) *For any  $u_0 \in V^2$ , the measure  $\mathbb{P}_{u_0}$  is a martingale solution for (6.22), (6.24) such that  $P_t(u_0, V^2) = 1$ , where  $P_t(u_0, \cdot)$  stands for the law of  $\omega(t)$  under  $\mathbb{P}_{u_0}$ :*

$$P_t(u_0, \Gamma) = \mathbb{P}_{u_0}\{\omega(t) \in \Gamma\}, \quad \Gamma \in \mathcal{B}(V^{-r}). \quad (6.30)$$

---

<sup>4</sup>It is easy to see that the trajectories of  $M_t$  belongs to  $C(\mathbb{R}_+; V^{-2})$  with  $\mathbb{P}$ -probability 1. Thus, the law of  $M$  is a probability measure on that space.

(ii) For any  $u_0 \in V^2$  and  $t > 0$ , with  $\mathbb{P}_{u_0}$ -probability 1 we have

$$\mathbb{P}_{u_0} \{ \omega(t + \cdot) \in \Gamma \mid \mathcal{F}_t \} = \mathbb{P}_{\omega(t)}(\Gamma), \quad (6.31)$$

where  $\Gamma \in \mathcal{F}$  is an arbitrary cylindrical set,

$$\Gamma = \{ \omega \in \Omega : \omega(t_i) \in B_i \text{ for } i = 1, \dots, n \},$$

with  $0 \leq t_1 < \dots < t_n$  and  $B_i \in \mathcal{B}(V^2)$ .

(iii) The function  $t \mapsto P_t(u_0, \cdot)$  from  $V^2$  to  $\mathcal{P}(V^2)$  is continuous in the weak topology.

We refer the reader to the original work [DO06] for a proof, confining ourselves to some comments on this result. Its formulation in [DO06] is slightly different: weak solutions are constructed there on an abstract probability space depending on the initial condition. Nevertheless, it is easy to see in the case of an additive noise that the law of a weak solution for (6.22) considered as a probability measure on the canonical space (6.27) is a martingale solution. Thus, we consider a (measurable) mapping taking an initial condition  $u_0$  to its law  $\mathbb{P}_{u_0}$ . Property (i) means that, at any time  $t \geq 0$ , the solution is  $H^2$  smooth. Note, however, that this property is not sufficient to ensure that the trajectories lie in  $H^2$ . Assertion (ii) is a weak form of the Markov property: in relation (6.31), one can take only cylindrical subsets  $\Gamma$  that are defined with the help of some Borel sets in  $V^2$ .<sup>5</sup> Finally, the last property means that the Markov semigroup  $\mathfrak{P}_t$  associated with the transition function  $P_t(u_0, \Gamma)$  is stochastically continuous.

Once there is a Markov semigroup associated with the 3D Navier–Stokes equations, a natural question is whether it is ergodic. The following theorem gives a positive answer to it.

**Theorem 6.2.8.** *Under the hypotheses of Theorem 6.2.7, the Markov semigroup  $\{\mathfrak{P}_t\}$  constructed there is irreducible and possesses the strong Feller property in  $V^2$ . In particular, it has a unique stationary distribution  $\mu \in \mathcal{P}(V^2)$ , and for any  $\lambda \in \mathcal{P}(V^2)$  we have*

$$\|\mathfrak{P}_t^* \lambda - \mu\|_{\text{var}} \rightarrow 0 \quad \text{as } t \rightarrow \infty, \quad (6.32)$$

where  $\|\cdot\|_{\text{var}}$  stands for the total variation norm in  $V^2$ . Furthermore, if  $\|h\|_1$  is sufficiently small, then convergence (6.32) is exponentially fast for any measure  $\lambda$  with finite second moment on  $H$ .

The proof of the strong Feller and irreducibility properties can be found in [DO06], while the exponential convergence in the total variation norm is established in [Oda07]. In conclusion, let us note that Theorem 6.2.8 does not imply uniqueness of a stationary distribution for the 3D Navier–Stokes equations, since different Markov semigroups may have different stationary measures.

<sup>5</sup>The set  $G = G_\Gamma \subset \Omega$  on which (6.31) holds is a full-measure event depending a priori on  $\Gamma$ , and we cannot rule out that the intersection of all  $G_\Gamma$ , with  $\Gamma$  varying in  $\mathcal{F}$ , is empty.

### 6.2.3 Flandoli–Romito theorem

The results discussed in the foregoing subsection show that it is possible to construct martingale solutions for the Navier–Stokes equations that satisfy a weak form of the Markov property, provided that the spatially regular white noise is sufficiently rough. A different approach inspired by Krylov’s paper [Kry73] (see the discussion in Section 6.2.1) was taken by Flandoli and Romito, who proved the existence of a (weak) Markov selection without any non-degeneracy assumption on the noise. To formulate their results, we shall need some auxiliary concepts.

**Definition 6.2.9.** Let  $X$  be a separable Banach space and let  $\Omega = C(\mathbb{R}_+; X)$  be the canonical space, which is endowed with its Borel  $\sigma$ -algebra  $\mathcal{F}$  and the natural filtration  $\{\mathcal{F}_t, t \geq 0\}$ . We shall say that a family of probability measures  $\{\mathbb{P}_u, u \in X\}$  possesses the *a. s. Markov property* if the mapping  $u \mapsto \mathbb{P}_u$  is measurable from  $X$  to the space  $\mathcal{P}(\Omega)$  with the weak topology, and for any  $u \in X$  there is a subset  $T_u \subset (0, +\infty)$  of zero Lebesgue measure such that, with  $\mathbb{P}_u$ -probability 1, we have

$$\mathbb{E}_u(f(\omega(t)) | \mathcal{F}_r) = \mathbb{E}_{\omega(r)}f(\omega(t-r)) \quad \text{for } r \notin T_u \text{ and all } t \geq r, \quad (6.33)$$

where  $f : X \rightarrow \mathbb{R}$  is an arbitrary bounded measurable function and  $\omega(t)$  stands for the canonical process.

For the purposes of this section, we need a concept of a martingale solution for (6.22) which is slightly different from (and is stronger than) that used in the foregoing subsection. To simplify the presentation, we shall always assume that  $h \equiv 0$  and the coefficients  $b_j$  satisfy (6.23). Let  $\Omega$  be the Fréchet space defined by (6.27) and endowed with its natural filtration and let  $z$  be a solution of the Stokes equations that vanishes at  $t = 0$  (see (2.101)).

**Definition 6.2.10.** A probability measure  $\mathbb{P}$  on  $\Omega$  is called a *admissible martingale solution* of (6.22) with an initial measure  $\mu \in \mathcal{P}(H)$  if it possesses the following properties:

- (i) The restriction of  $\mathbb{P}$  to  $t = 0$  coincides with  $\mu$  and

$$\mathbb{P}(L_{\text{loc}}^\infty(\mathbb{R}_+; H) \cap L_{\text{loc}}^2(\mathbb{R}_+, V)) = 1;$$

- (ii) The law of process (6.28) with  $h \equiv 0$  coincide with that of  $\zeta$ ;  
 (iii) There is a set  $T \subset (0, +\infty)$  of zero Lebesgue measure such that

$$\mathbb{P}\{\mathcal{E}_t(v, z) \leq \mathcal{E}_r(v, z)\} = 1 \quad \text{for } r \notin T \text{ and all } t \geq r, \quad (6.34)$$

where we set  $v(t) = \omega(t) - z(t)$  and

$$\mathcal{E}_t(v) = \frac{1}{2}|v(t)|_2^2 + \int_0^t (\nu|\nabla v|_2^2 + \langle B(v+z, z), v \rangle) ds.$$

In what follows, we call (6.34) the *energy inequality* for solutions of (6.22). The main difference between admissible martingale solutions and the solutions considered in the preceding section is that the former satisfy the energy inequality. This property is crucial for proving the following weak-strong uniqueness result (see [Soh01] for the deterministic problem and [Shi07b] for the stochastic case).

**Proposition 6.2.11.** *Let  $\mathbb{P}$  be an admissible martingale solution for (6.22) issued from a point  $u_0 \in V$  and let  $\{u(t), t \in [0, T]\}$  be a random process with continuous  $V$ -valued trajectories that satisfy (6.22) and coincide with  $u_0$  at  $t = 0$  with probability 1. Then the restriction to  $[0, T]$  of  $\mathbb{P}$  coincides with the law of  $\{u(t), t \in [0, T]\}$ .*

We now turn to the problem of a Markov selection among admissible martingale solutions of (6.22). The following result established in [FR08, Theorem 4.1] is an analogue of Theorem 6.2.5 in the case of Navier–Stokes equations.

**Theorem 6.2.12.** *Assume that the coefficients  $b_s$  satisfy (6.23). Then there is family  $\{\mathbb{P}_{u_0}, u_0 \in H\}$  of measures on  $\Omega$  possessing the a. s. Markov property such that, for any  $u_0 \in H$ ,  $\mathbb{P}_{u_0}$  is an admissible martingale solution for (6.22) with the initial measure  $\delta_{u_0}$ .*

In what follows, any such family will be called a *Markov selection* for (6.22). Theorem 6.2.12 is true without any non-degeneracy hypothesis on the noise. In particular, it holds in the case  $\zeta \equiv 0$ . If we assume that the noise is very rough, then the Markov property will be true for any time  $s \geq 0$ , and the Markov family  $(\omega(t), \mathbb{P}_{u_0})$  will be ergodic. More precisely, let us assume that the coefficients  $b_s$  satisfy the inequality (cf. (6.29))

$$c|s|^{-\frac{11}{6}-\delta} \leq |b_s| \leq C|s|^{-\frac{11}{6}-\delta} \quad \text{for all } s \in \mathbb{Z}_0^3, \quad (6.35)$$

where  $C > c > 0$  are some constants. The following theorem is established in [FR08, Rom08].

**Theorem 6.2.13.** *Assume that (6.35) holds. Then there is  $\theta \geq 0$  depending on  $\delta$  such that the assertions below hold for the Markov selection  $\{\mathbb{P}_{u_0}, u_0 \in H\}$  constructed in Theorem 6.2.12.*

**Markov property.** *The family  $(\omega(t), \mathbb{P}_{u_0})$  is a Markov process. That is, relation (6.33) holds for all  $t \geq r \geq 0$  and any bounded measurable function  $f : H \rightarrow \mathbb{R}$ .*

**Strong Feller property.** *The family  $(\omega(t), \mathbb{P}_{u_0})$  is strong Feller in the topology of  $V^{1+\theta}$ . That is, the function  $u_0 \mapsto P_t(u_0, \Gamma)$  is continuous on  $V^{1+\theta}$  for any  $\Gamma \in \mathcal{B}(H)$ .*

**Stationary distribution.** *There exists a unique stationary distribution  $\mu$  for  $(\omega(t), \mathbb{P}_{u_0})$ .*



**Exponential mixing.** *Convergence (6.32) holds exponentially fast for any initial measure  $\lambda \in \mathcal{P}(H)$  with finite second moment.*

A striking consequence of the existence of a Markov selection is that the local existence of a strong solution on an arbitrary small interval implies the global existence, provided that the selection possesses the Feller property on  $V^{1+\theta}$  with  $\theta > 0$  and is irreducible. In particular, the result is true under the hypotheses of Theorem 6.2.13.

**Proposition 6.2.14.** *Let  $\{\mathbb{P}_u, u \in H\}$  be an arbitrary Markov selection for the 3D Navier–Stokes equations (6.22) which possesses the following properties for some  $\theta \geq 0$ .*

**Feller property.** *The Markov semigroup associated with  $\mathbb{P}_u$  is Feller on  $V^{1+\theta}$ . That is,  $\mathfrak{P}_t f \in C_b(V^{1+\theta})$  for any  $f \in C_b(V^{1+\theta})$ .*

**Irreducibility.** *For any  $t > 0$  and  $u \in V^{1+\theta}$ , we have  $P_t(u, G) > 0$  for all open subsets  $G \subset V^{1+\theta}$ .*

*Let us assume that Eq. (6.22) possesses a martingale solution  $\mathbb{P}$  on an interval  $[0, \tau]$  such that  $\mathbb{P}(C(0, \tau; V^{1+\theta})) = 1$ . Then,  $\mathbb{P}_u(C(\mathbb{R}_+; V^{1+\theta})) = 1$  for any  $u \in V^{1+\theta}$ . In particular, the stochastic Navier–Stokes equations (6.22) are well-posed in  $V^{1+\theta}$ .*

Let us emphasise that no initial condition is imposed on  $\mathbb{P}$ . So  $\mathbb{P}|_{t=0}$  may be any measure on  $V^{1+\theta}$ , e.g., a Dirac mass at zero.

*Sketch of the proof.* We follow essentially the argument used by Flandoli and Romito [FR08, Section 6.4]. To simplify formulas, we shall assume that  $\theta = 0$ . Let  $\mu$  be the projection of  $\mathbb{P}$  to time  $t = 0$ :

$$\mu(\Gamma) = \mathbb{P}(\{\omega \in \Omega : \omega(0) \in \Gamma\}), \quad \Gamma \in \mathcal{B}(H).$$

We define  $\mathbb{P}_\mu = \int_H \mathbb{P}_u \mu(du)$  and note that  $\mathbb{P}_\mu$  is an admissible martingale solution for (6.22) with the initial measure  $\mu$ . By Proposition 6.2.11, we see that the restriction of  $\mathbb{P}_\mu$  to  $[0, \tau]$  coincides with  $\mathbb{P}$ . The Markov property of  $\mathbb{P}_u$  now implies that<sup>6</sup>

$$\begin{aligned} 1 &= \mathbb{P}_\mu(C(0, \tau; V)) = \mathbb{E}_\mu \mathbb{P}_\mu\{C(0, \tau; V) \mid \mathcal{F}_\theta\} \\ &= \mathbb{E}_\mu\{\mathbb{I}_{C(0, \theta; V)} \mathbb{P}_{\omega(\theta)}(C(0, \tau - \theta; V))\}, \end{aligned}$$

where we set  $\theta \in (0, \tau)$  is an arbitrary point for which the Markov property holds. Since the non-negative function under the second expectation above does not exceed 1, we conclude that

$$\mathbb{P}_{\omega(\theta)}(C(0, \tau - \theta; V)) = 1 \quad \text{for } \mathbb{P}_\mu\text{-almost every } \omega \in \Omega.$$

<sup>6</sup>With a slight abuse of notation, we use the same symbol for a measure on  $\Omega$  and its restriction to the interval  $[0, \tau]$ .

The irreducibility property implies that there is a dense subset  $A \subset V$  such that  $\mathbb{P}_v(C(0, \tau - \theta; V)) = 1$  for any  $v \in A$ . We claim that this relation is true for any  $v \in V$ . If this is established, then the Markov property will imply by iteration that  $\mathbb{P}_u(C(\mathbb{R}_+; V)) = 1$  for any  $u \in V$ . Since a solution of the Navier–Stokes system is unique in the class  $C(\mathbb{R}_+; V)$ , we conclude that it is well-posed in  $V$ .

We now prove  $\mathbb{P}_v(C(0, \theta_1; V)) = 1$  for any  $v \in V$ , where  $\theta_1 = \tau - \theta$ . For any integer  $N \geq 1$ , denote by  $\mathbb{Q}_N(\theta_1)$  the set of all rational numbers  $2^{-N}p \in [0, \theta_1]$ , where  $p$  is integer. Then

$$C(0, \theta_1; V) = \bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{N=1}^{\infty} G(N, k, m), \quad (6.36)$$

where we set

$$G(N, k, m) = \left\{ \omega \in \Omega : \|\omega(t) - \omega(s)\|_V \leq \frac{1}{m} \text{ for } t, s \in \mathbb{Q}_N(\theta_1), |t - s| \leq \frac{1}{k} \right\}.$$

Note that each subset  $G(N, k, m)$  depends on the values of  $\omega \in \Omega$  at finitely many points  $t_1, \dots, t_l \in [0, \theta_1]$ , where  $l \geq 1$  is an integer, and can be represented in the form

$$G(N, k, m) = \{ \omega \in \Omega : (\omega(t_1), \dots, \omega(t_l)) \in \Gamma(N, k, m) \},$$

where  $\Gamma(N, k, m)$  is a closed set in the direct product of  $l$  copies of  $V$ .

Now let  $v \in V$  and let  $\{v_n\} \subset A$  be a sequence converging to  $v$  in  $V$ . Then  $\mathbb{P}_{v_n}(C(0, \theta_1; V)) = 1$ , whence it follows that

$$\mathbb{P}_{v_n}(G(N, k, m)) \geq 1 - \varepsilon_k \quad \text{for all } m, n, N,$$

where  $\varepsilon_k \rightarrow 0$  as  $k \rightarrow \infty$ . The Feller and Markov properties and the above mentioned structure of  $G(N, k, m)$  imply that

$$\mathbb{P}_v(G(N, k, m)) \geq \limsup_{n \rightarrow \infty} \mathbb{P}_{v_n}(G(N, k, m)) \geq 1 - \varepsilon_k \quad \text{for all } m, N.$$

Combining this with (6.36), we arrive at the required result, which completes the proof of the proposition.  $\square$

In conclusion, let us note the irreducibility condition of Proposition 6.2.14 implies that the noise  $\eta$  is non-zero, and therefore it cannot be satisfied for the deterministic problem. On the other hand, the energy inequality (for admissible martingale solutions) and the Feller property make sense also for the deterministic Navier–Stokes system. The energy inequality is true for any solution obtained by a constructive procedure (such as Galerkin approximations or regularisation of the nonlinear term), while the Feller property means, roughly speaking, that any sequence of strong solutions converge to a strong solution, provided that the initial conditions converge in  $V^{1+\theta}$ . This is a very strong property and would immediately imply global existence of a strong solution in the deterministic case.

### 6.3 Navier–Stokes equations with a very degenerate noise

The understanding of the 3D Navier–Stokes dynamics with a smooth stochastic forcing is rather poor, and only some partial results showing its mixing character are available. They are based on a detailed study of controllability properties of the Navier–Stokes system. The counterparts of these results in the 2D case are much more complete and have independent interest. In this section, we first discuss some results for 2D Navier–Stokes equations and then turn to their generalisations for the 3D case. We confine ourselves to the case of periodic boundary conditions, even though some of the results remain true for other geometries.

#### 6.3.1 2D Navier–Stokes equations: controllability and mixing properties

Let us consider controlled<sup>7</sup> Navier–Stokes equations on a 2D torus :

$$\dot{u} + \nu Lu + B(u) = h + \eta(t). \quad (6.37)$$

To simplify the presentation, we shall consider only the case of a square torus  $\mathbb{T}^2$ ; however, all the results are true for a rectangular torus. We assume that  $h \in H$  is a given function and  $\eta(t)$  is a *control* taking values in a finite-dimensional space  $E \subset V^2$ . Recall that  $S_t(u_0, h + \eta)$  denotes the solution of (6.37) issued from  $u_0 \in H$ . We shall need the following two concepts of controllability.

**Definition 6.3.1.** We shall say that Eq. (6.37) is *approximately controllable at a time  $T > 0$  by an  $E$ -valued control* if for any  $u_0, \hat{u} \in H$  and any  $\varepsilon > 0$  there is  $\eta \in C^\infty(0, T; E)$  such that

$$\|S_T(u_0, h + \eta) - \hat{u}\|_2 < \varepsilon. \quad (6.38)$$

We shall say that Eq. (6.37) is *solidly controllable in projections at a time  $T > 0$  by an  $E$ -valued control* if for any  $u_0 \in H$ , any  $R > 0$ , and any finite-dimensional subspace  $F \subset H$  there is a compact set  $X \subset L^2(0, T; E)$  and a constant  $\varepsilon > 0$  such that for any continuous mapping  $\Phi : X \rightarrow F$  satisfying the inequality  $\sup_{\xi \in X} \|\Phi(\xi) - S_T(u_0, \xi)\| \leq \varepsilon$  we have

$$P_F \Phi(X) \supset B_F(R), \quad (6.39)$$

where  $P_F : H \rightarrow H$  denotes the orthogonal projection in  $H$  onto  $F$ .

Thus, when studying the controlled Navier–Stokes system (6.37), we assume that the function  $\eta$  is at our disposal, and we try to make sure that at the final time  $T$  the solutions possess some particular properties.

<sup>7</sup>The meaning of this term is clarified below.

To formulate the main result on controllability, we introduce some notations. Given a finite-dimensional subspace  $G \subset V^2$ , we denote by  $\mathcal{E}(G)$  the largest vector space in  $V^2$  whose elements can be represented in the form

$$\eta_1 = \eta - \sum_{j=1}^N B(\zeta^j),$$

where  $\eta, \zeta^1, \dots, \zeta^N \in G$  and  $N$  is an integer depending on  $\eta_1$ . Let us define an increasing sequence  $\{E_k\}$  of subspaces in  $V^2$  by the rule

$$E_0 = E, \quad E_k = \mathcal{E}(E_{k-1}) \quad \text{for } k \geq 1,$$

and denote by  $E_\infty$  the union of  $E_k$ ,  $k \geq 1$ . We shall say that  $E$  is *saturating* if  $E_\infty$  is dense in  $H$ . The following result is due to Agrachev and Sarychev [AS05, AS06].

**Theorem 6.3.2.** *Let  $h \in H$  and let  $E \subset V^2$  be a saturating subspace. Then for any  $T > 0$  Eq. (6.37) is approximately controllable and solidly controllable at time  $T$  by an  $E$ -valued control.*

A necessary and sufficient condition for a subspace to be saturating can be found in [AS06, Theorem 4]. In particular, if  $E$  contains the basis vectors  $e_{(s_1, s_2)}$  with  $|s_1| \vee |s_2| \leq 1$ , then it is saturating.

We now fix  $T > 0$  and assume that the external force  $\eta$  in (6.37) is a random process on  $[0, T]$  whose trajectories belong to  $L^2(0, T; E)$  almost surely. In this case, the dynamics of (6.37) is well defined because we can apply the corresponding deterministic result to any realisation of  $\eta$  (see Theorem 2.1.13). We wish to investigate qualitative properties of solutions for (6.37) at time  $T$ . To this end, we impose the following *decomposability* assumption on the law of  $\eta$ . Let  $\lambda$  be the law of the process  $\eta$  regarded as a random variable in  $L^2(0, T; E)$ .

**Definition 6.3.3.** We shall say that  $\lambda$  is *decomposable* if there is an orthonormal basis in  $L^2(0, T; E)$  such that the measure  $\lambda$  can be written as the direct product of its projections to the straight lines spanned by the basis vectors.

Given a decomposable measure  $\lambda$ , we shall denote by  $\lambda_j$  its projection to the straight line spanned by the  $j^{\text{th}}$  basis vector. Thus, we can write  $\lambda = \otimes_{j=1}^\infty \lambda_j$ . Combining controllability properties of the Navier–Stokes system with a result on the image of probability measures under analytic mappings, Agrachev, Kuksin, Sarychev, and Shirikyan [AKSS07] established the following theorem.

**Theorem 6.3.4.** *Let  $T > 0$ , let  $h \in H$ , let  $E \subset V^2$  be a finite-dimensional subspace, and let  $\eta$  be a random process whose law  $\lambda$  is a decomposable measure in  $L^2(0, T; E)$  such that the support of  $\lambda$  coincides with the entire space, the second moment  $\mathfrak{m}_2(\lambda)$  is finite, and the one-dimensional projections  $\lambda_j$  possess continuous densities with respect to the Lebesgue measure. Then for any  $u_0 \in H$  the solution  $u(t, u_0)$  of (6.37) issued from  $u_0$  possesses the following properties.*

**Non-degeneracy.** *The support of the law of  $u(T; u_0)$  coincides with  $H$ .*

**Regularity.** For any finite-dimensional subspace  $F \subset H$ , the law of  $\mathbf{P}_{Fu}(T; u_0)$  possesses a density with respect to the Lebesgue measure and continuously depends on  $u_0 \in H$  in the total variation norm.

Note that a similar result is true when the random perturbation  $\eta$  is a spatially regular white noise or a kick force. In this case, the above properties can be reformulated in terms of the transition function  $P_t(u, \Gamma)$ . Namely, the non-degeneracy is equivalent to the relation  $\text{supp } P_t(u, \cdot) = H$ , while the regularity means that the finite-dimensional projections of  $P_t(u, \cdot)$  are absolutely continuous with respect to the Lebesgue measure and continuously depend on  $u$  in the total variation norm. These properties implies, in particular, that if  $\mu$  is a stationary measure, then its support coincides with  $H$  and any finite-dimensional projection is absolutely continuous.

### 6.3.2 3D Navier–Stokes equations with a degenerate noise

The Cauchy problem for the Navier–Stokes system is not known to be well posed in the 3D case. Nevertheless, the control problem still makes sense, and an analogue of Theorem 6.3.2 is true. We shall not give here an exact formulation, referring the reader to the papers [Shi06a, Shi07a]. Theorem 6.3.4 can be extended partially to the 3D case, and we now describe the corresponding result.

To this end, we first introduce the concept of an *admissible weak solution* for 3D Navier–Stokes equation; cf. Definition 6.2.10. We consider Eq. (6.37) on the 3D torus, assuming that  $h \in H$  is a deterministic function and  $\eta$  is a spatially regular white noise defined on a filtered probability space  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  satisfying the usual hypotheses; see Section 1.2.1.

**Definition 6.3.5.** We shall say that a random process  $u(t)$  is an *admissible solution* for (6.37) if it can be written in the form  $u(t) = z(t) + v(t)$ , where  $z$  is the solution of the stochastic Stokes equation constructed in Proposition 2.4.2 and  $v$  is a random process satisfying the following properties.

- (i) The process  $v$  is adapted to the filtration  $\mathcal{F}_t$ , and its almost every trajectory belongs to the space  $L_{\text{loc}}^\infty(\mathbb{R}_+; H) \cap L_{\text{loc}}^2(\mathbb{R}_+; V)$  and satisfies Eq. (2.111).
- (ii) The following energy inequality holds for all  $t \geq 0$  with probability 1:

$$\frac{1}{2}|v(t)|_2^2 + \nu \int_0^t |\nabla v|_2^2 ds + \int_0^t (B(v+z, z)) ds \leq \frac{1}{2}|v(0)|_2^2 + \int_0^t (h, v) ds.$$

We shall say that  $\tilde{u}$  is an *admissible weak solution* for (6.37) if there is a spatially regular white noise  $\tilde{\eta}$  distributed as  $\eta$  such that  $\tilde{u}$  is an admissible solution for Eq. (6.37) with  $\eta$  replaced by  $\tilde{\eta}$ .

A proof of the existence of an admissible weak solution for (6.37) can be found in [VF88, CG94, FG95]. Moreover, it is proved in [VF88, FG95] that

one can construct an admissible weak solution with a time-independent law  $\mu$  satisfying the condition

$$\mathfrak{m}_2(\mu) = \int_H \|u\|_V^2 \mu(du) < \infty.$$

In this case, we call  $\mu$  a *stationary measure* for (6.37). Given two measures  $\mu_1$  and  $\mu_2$  on a Polish space  $X$ , we write  $\mu_1 \leq \mu_2$  if  $\mu_1(\Gamma) \leq \mu_2(\Gamma)$  for any  $\Gamma \in \mathcal{B}(X)$ . The following result is proved by Shirikyan [Shi07b].

**Theorem 6.3.6.** *Let  $h \in H$  and let  $\eta$  be a spatially regular white noise of the form (2.66). Then there is an integer  $N \geq 1$  not depending on  $\nu$  such that if  $b_j \neq 0$  for  $1 \leq j \leq N$ , then the following two assertions hold for any stationary measure  $\mu$  for (6.37) that satisfies the inequality  $\mathfrak{m}_2(\mu) \leq \mathfrak{m}$ , where  $\mathfrak{m} > 0$  is a constant.*

**Non-degeneracy.** *For any ball  $Q \subset V$  there is a constant  $p = p(Q, \mathfrak{m}) > 0$  such that  $\mu(B) \geq p(Q, \mathfrak{m})$ .*

**Weak regularity.** *Let  $F \subset H$  be a finite-dimensional subspace and let  $\mu_F$  be the projection of  $\mu$  to  $F$ . Then there is a function  $\rho_F \in C^\infty(F)$  depending only on  $\mathfrak{m}$  such that  $\mu_F \geq \rho_F \ell_F$ , where  $\ell_F$  stands for the Lebesgue measure on  $F$ .*

In conclusion, let us mention that the non-degeneracy property remains true in the case of a bounded domain, provided that the noise is non-zero in all Fourier modes. This result was established by Flandoli [Fla97].

# Chapter 7

## Appendix

### 7.1 Monotone class theorem

Let  $\Omega$  be a set and let  $\mathcal{M}$  be a family of subsets of  $\Omega$ . We shall say that  $\mathcal{M}$  is a *monotone class* if it contains  $\Omega$  and possesses the following properties:

- if  $A, B \in \mathcal{M}$  and  $A \cap B = \emptyset$ , then  $A \cup B \in \mathcal{M}$ ;
- if  $A, B \in \mathcal{M}$  and  $A \subset B$ , then  $B \setminus A \in \mathcal{M}$ ;
- if  $A_i \in \mathcal{M}$  for  $i = 1, 2, \dots$ , and  $A_1 \subset A_2 \subset \dots$ , then  $\bigcup_i A_i \in \mathcal{M}$ .

It is clear that any  $\sigma$ -algebra is a monotone class, but not vice versa. The following well-known theorem gives a sufficient condition ensuring that the minimal monotone class containing a family of subsets coincides with the  $\sigma$ -algebra generated by that family. Recall that a family  $\mathcal{C}$  of subsets of  $\Omega$  is called a  *$\pi$ -system* if  $A \cap B \in \mathcal{C}$  for any  $A, B \in \mathcal{C}$ .

**Theorem 7.1.1.** *Let  $\Omega$  be a set and let  $\mathcal{C}$  be a  $\pi$ -system of subsets of  $\Omega$ . Then the minimal monotone class containing  $\mathcal{C}$  coincides with the  $\sigma$ -algebra generated by  $\mathcal{C}$ .*

*Proof.* It suffices to show that the minimal monotone class  $\mathcal{M}$  containing  $\mathcal{C}$  is a  $\sigma$ -algebra. This will be established if we prove that the intersection of any two sets in  $\mathcal{M}$  is also an element of  $\mathcal{M}$ .

Let us fix an arbitrary set  $A \in \mathcal{C}$  and define the family

$$\mathcal{M}_A = \{B \in \mathcal{M} : A \cap B \in \mathcal{M}\}.$$

It is clear that  $\mathcal{M}_A$  is a monotone class and  $\mathcal{M}_A \supset \mathcal{C}$ . Therefore, by the definition of  $\mathcal{M}$ , we must have  $\mathcal{M}_A \supset \mathcal{M}$ . We have thus shown that, for any  $A \in \mathcal{C}$  and  $B \in \mathcal{M}$ , the intersection  $A \cap B$  belongs to  $\mathcal{M}$ .

We now fix  $A \in \mathcal{M}$  and consider the family  $\mathcal{M}_A$ . Repeating literally the above argument, we can show that  $\mathcal{M}_A \supset \mathcal{M}$ . This completes the proof of the lemma.  $\square$

The following two results provide important applications of the monotone class technique.

**Corollary 7.1.2.** *Let  $(\Omega, \mathcal{F})$  be a measurable space and let  $\mathcal{C} \subset \mathcal{F}$  be a  $\pi$ -system generating the  $\sigma$ -algebra  $\mathcal{F}$ . Suppose that  $\mu_1$  and  $\mu_2$  are two probability measures on  $(\Omega, \mathcal{F})$  such that*

$$\mu_1(\Gamma) = \mu_2(\Gamma) \quad \text{for any } \Gamma \in \mathcal{C}. \quad (7.1)$$

*Then  $\mu_1 = \mu_2$ .*

*Proof.* The  $\sigma$ -additivity of measures imply that if (7.1) holds, then  $\mu_1$  and  $\mu_2$  are equal on the minimal monotone class  $\mathcal{M}$  containing  $\mathcal{C}$ . In view of Theorem 7.1.1, the  $\sigma$ -algebra  $\mathcal{F}$  coincides with  $\mathcal{M}$ .  $\square$

**Corollary 7.1.3.** *Let  $X$  be a Polish space with its Borel  $\sigma$ -algebra  $\mathcal{B}(X)$ . Then there is a countable family  $\{f_j, j \geq 1\}$  of bounded Lipschitz functions such that any two measures  $\mu_1, \mu_2 \in \mathcal{P}(X)$  satisfying the following relation are equal:*

$$(f_j, \mu_1) = (f_j, \mu_2) \quad \text{for any } j \geq 1. \quad (7.2)$$

*Proof.* For any closed subset  $F \subset X$ , there is a sequence of Lipschitz functions  $\{g_k\}$  such that  $0 \leq g_k \leq 1$  and  $\{g_k\}$  converges pointwise to the indicator function of  $F$ . For instance, we can take

$$g_k(u) = \frac{\text{dist}_X(u, G_{1/k})}{\text{dist}_X(u, G_{1/k}) + \text{dist}_X(u, F)},$$

where  $G_\varepsilon$  stands for the complement of the  $\varepsilon$ -neighbourhood for  $F$ . Now note that if  $(g_k, \mu_1) = (g_k, \mu_2)$  for all  $k$ , then passing to the limit as  $k \rightarrow \infty$ , we obtain  $\mu_1(F) = \mu_2(F)$ .

Let  $\mathcal{C} \subset \mathcal{F}$  be a countable  $\pi$ -system of closed subsets that generates  $\mathcal{F}$ . For any  $F \in \mathcal{C}$ , we can construct a sequence  $\{g_k\}$  that satisfies the above properties. Taking the union of all such sequences, we obtain a countable family  $\{f_j\}$  such that, for any two measures  $\mu_1, \mu_2 \in \mathcal{P}(X)$  satisfying (7.2), we have

$$\mu_1(F) = \mu_2(F) \quad \text{for any } F \in \mathcal{C}.$$

Since  $\mathcal{C}$  is a  $\pi$ -system generating  $\mathcal{F}$ , we see that  $\mu_1 = \mu_2$ .  $\square$

## 7.2 Standard measurable spaces

Recall that two measurable spaces  $(\Omega_i, \mathcal{F}_i)$ ,  $i = 1, 2$ , are said to be *isomorphic* if there is a bijective map  $f : \Omega_1 \rightarrow \Omega_2$  such that  $f(\Gamma) \in \mathcal{F}_2$  if and only if  $\Gamma \in \mathcal{F}_1$ .

**Definition 7.2.1.** Let  $(\Omega, \mathcal{F})$  be a measurable space. We shall say that  $(\Omega, \mathcal{F})$  is a *standard measurable space* if it is isomorphic to one of the following Polish spaces endowed with their Borel  $\sigma$ -algebras:



- the finite set  $\{1, \dots, N\}$  with the discrete topology;
- the set of positive integers  $\mathbb{N}$  with the discrete topology;
- the closed interval  $[0, 1]$  with its standard metric.

A proof of the following theorem can be found in [Dud02] (see Chapter 13).

**Theorem 7.2.2.** *Any Polish space endowed with its Borel  $\sigma$ -algebra is a standard measurable space.*

An important property of a standard measurable space is that its  $\sigma$ -algebra is *countably generated*. Namely, we have the following result.

**Proposition 7.2.3.** *Let  $(\Omega, \mathcal{F})$  be a standard measurable space. Then there is a  $\pi$ -system  $\mathcal{C} \subset \mathcal{F}$  that is at most countable and generates the  $\sigma$ -algebra  $\mathcal{F}$ . Moreover, if  $\Omega$  is a Polish space endowed with its Borel  $\sigma$ -algebra, then there is a  $\pi$ -system  $\mathcal{C}$  of closed subsets that satisfies the above properties.*

*Proof.* We shall confine ourselves to the proof of the assertion concerning the Polish spaces, because the first part of the theorem follows from it and the definition of a standard measurable space. Let  $\{\omega_j\} \subset \Omega$  be a countable dense subset and let  $\mathcal{C}_0$  be the countable family of all closed balls centred at  $\omega_j$  with rational radii. We claim that the family  $\mathcal{C}$  of all finite intersections of the elements in  $\mathcal{C}_0$  satisfies the required properties. Indeed, the construction implies that  $\mathcal{C}$  is an at most countable  $\pi$ -system. Furthermore, any open set in  $\Omega$  can be represented as a countable union of the elements of  $\mathcal{C}_0$ . It follows that the Borel  $\sigma$ -algebra is the minimal one that contains  $\mathcal{C}$ .  $\square$

## 7.3 Projection theorem

Let  $(\Omega, \mathcal{F})$  be a measurable space and let  $X$  be a Polish space endowed with its Borel  $\sigma$ -algebra  $\mathcal{B}(X)$ . Consider the measurable space  $(\Omega \times X, \mathcal{F} \otimes \mathcal{B}(X))$ , and the natural projection  $\Pi_\Omega : \Omega \times X \rightarrow \Omega$ . It is well known that the image under  $\Pi_\Omega$  of a measurable subset of  $\Omega \times X$  does not necessarily belong<sup>1</sup> to  $\mathcal{F}$ . Recall that a subset  $A \subset \Omega$  is said to be *universally measurable* if it belongs to the completion of  $\mathcal{F}$  with respect to any probability measure on  $(\Omega, \mathcal{F})$ . A proof of the following result can be found in [CV77] (see Theorem III.23).

**Theorem 7.3.1. (projection theorem)** *For any set  $A \in \mathcal{F} \otimes \mathcal{B}(X)$ , its projection  $\Pi_\Omega(A)$  is universally measurable.*

As an immediate consequence of the above theorem, we establish the following result (see Chapter 2 in [Cra02]). Recall that a function  $g : \Omega \rightarrow \mathbb{R}$  is said to be *universally measurable* if so is the set  $\{\omega \in \Omega : g(\omega) > a\}$  for any  $a \in \mathbb{R}$ .

<sup>1</sup>This property is not true even in the case when  $\Omega$  is the closed interval  $[0, 1]$ , so that the map  $\Pi_\Omega$  is continuous.

**Corollary 7.3.2.** *Let  $f : \Omega \times X \rightarrow \mathbb{R}$  be a measurable function. Then the functions*

$$\bar{f}(\omega) = \sup_{u \in X} f(\omega, u), \quad \underline{f}(\omega) = \inf_{u \in X} f(\omega, u)$$

*are universally measurable.*

*Proof.* Note that

$$\{\omega \in \Omega : \bar{f}(\omega) > a\} = \{\omega \in \Omega : f(\omega, u) > a \text{ for some } u \in X\} = \Pi_\Omega(A),$$

where we set

$$A = \{(\omega, u) \in \Omega \times X : f(\omega, u) > a\}.$$

It remains to note that the set  $A$  belongs to  $\mathcal{F} \otimes \mathcal{B}(X)$ , and therefore, by Theorem 7.3.1, its projection is universally measurable. The universal measurability of  $\underline{f}$  can be established by a similar argument.  $\square$

## 7.4 Gaussian random variables

Let us recall that a one-dimensional normal distribution  $N(m, \sigma)$ , with mean value  $m \in \mathbb{R}$  and variance  $\sigma^2 > 0$ , is given by its density against the Lebesgue measure:

$$p_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

In the case  $\sigma = 0$ , the normal distribution  $N(m, 0)$  coincides with the Dirac measure concentrated at  $x = m$ . The distribution function of  $N(0, \sigma)$  is denoted by  $\Phi_\sigma$ :

$$\Phi_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2\sigma^2} dy.$$

Recall that the *characteristic function* of a measure  $\mu \in \mathcal{P}(\mathbb{R})$  is defined as  $\varphi_\mu(t) = \int_{\mathbb{R}} e^{itx} \mu(dx)$ . A measure  $\mu$  is a normal distribution  $N(m, \sigma)$  if and only if its characteristic function is given by

$$\varphi_{m,\sigma}(t) = \exp(imt - \sigma^2 t^2/2), \quad t \in \mathbb{R}. \quad (7.3)$$

Let  $X$  be a separable Banach space and let  $X^*$  be its dual space. A probability measure  $\mu$  on  $(X, \mathcal{B}(X))$  is said to be *Gaussian* if  $\ell_*(\mu)$  has a normal distribution for any  $\ell \in X^*$ . A Gaussian measure  $\mu$  is said to be *centred* if the mean value of  $\ell_*(\mu)$  is zero for any  $\ell \in X^*$ . An  $X$ -valued random variable is said to be *Gaussian* if its distribution is a Gaussian probability measure. A comprehensive study of Gaussian measures on topological vector spaces can be found in [Bog98]. The following proposition establishes some simple properties that were used in the main text.

**Proposition 7.4.1.** (i) *Let  $X$  be a separable Banach space and let  $\{\mu_n\}$  be a sequence of Gaussian measures that converges weakly to  $\mu \in \mathcal{P}(X)$ . Then  $\mu$  is a Gaussian measure.*

(ii) A finite linear combination of independent Gaussian random variables is also Gaussian.

*Proof.* (i) It suffices to consider the one-dimensional case. Let us denote by  $m_n$  and  $\sigma_n$  the mean value and the variance for  $\mu_n$ . Then the characteristic function of  $\mu_n$  is given by  $\varphi_{\mu_n}(t) = \exp(im_n t - \sigma_n^2 t^2 / 2)$ . Furthermore, since  $\mu_n$  converges weakly to  $\mu$ , the corresponding characteristic functions must converge pointwise. It follows that the sequences  $\{m_n\}$  and  $\{\sigma_n\}$  have some limits  $m \in \mathbb{R}$  and  $\sigma \geq 0$ , respectively, and the characteristic function of  $\mu$  coincides with (7.3). We thus conclude that  $\mu$  is a normal distribution on  $\mathbb{R}$ .

(ii) This property is a straightforward consequence of the fact that the characteristic function of the sum of independent random variables is equal to the product of the corresponding characteristic functions.  $\square$

We now consider an important example of a Gaussian random variable. Let  $\beta(t)$  be a standard Brownian motion. For a constant  $\gamma \in \mathbb{R}$ , we define the stochastic process

$$z(t) = \int_0^t e^{\gamma(t-s)} d\beta(s), \quad t \in [0, 1].$$

It is straightforward to see that almost every trajectory of  $z$  belongs to the space  $\dot{C}(0, 1)$  of real-valued continuous functions on  $[0, 1]$  vanishing at zero.

**Proposition 7.4.2.** *The process  $\{z(t), 0 \leq t \leq 1\}$  regarded as a random variable with range in  $\dot{C}(0, 1)$  has a centred Gaussian distribution whose support coincides with the entire space.*

*Sketch of the proof.* To prove that  $z$  is a centred Gaussian random variable, we need to show that if  $\ell : \dot{C}(0, 1) \rightarrow \mathbb{R}$  is a continuous linear functional, then  $\ell(z)$  has a normal distribution with zero mean value. By the Riesz theorem (e.g., see Theorem 7.4.1 in [Dud02]), there is a function of bounded variation  $\varphi : [0, 1] \rightarrow \mathbb{R}$  such that

$$\ell(g) = \int_0^1 g(t) d\varphi(t) \quad \text{for any } g \in \dot{C}(0, 1).$$

It follows that

$$\ell(z) = \int_0^1 \left( \int_0^t e^{\gamma(t-s)} d\beta(s) \right) d\varphi(t).$$

We wish to change the order of integration. To this end, we use the following lemma whose proof is left to the reader as an exercise.

**Lemma 7.4.3.** *Let  $\beta$  be a standard Brownian motion, let  $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$  be a continuously differentiable function, and let  $\varphi : [0, 1] \rightarrow \mathbb{R}$  be a function of bounded variation. Then, with probability 1, we have*

$$\int_0^1 \left( \int_0^t f(t, s) d\beta(s) \right) d\varphi(t) = \int_0^1 \left( \int_s^1 f(t, s) d\varphi(t) \right) d\beta(s). \quad (7.4)$$

Relation (7.4) implies that

$$\ell(z) = \int_0^1 h(s) d\beta(s), \quad h(s) = e^{-\gamma s} \int_s^1 e^{\gamma t} d\varphi(t).$$

The mean value of a stochastic integral is always zero. Thus, the required assertion will be proved if we show that the stochastic integral of any square-integrable deterministic function  $h$  has a normal distribution.

If  $h$  is a piecewise constant function equal to  $c_k$  on  $[t_{k-1}, t_k)$ ,  $k = 1, \dots, N$ , then

$$\int_0^1 h(s) d\beta(s) = \sum_{k=1}^N c_k (\beta(t_k) - \beta(t_{k-1})).$$

Assertion (ii) of Proposition 7.4.1 implies that this random variable has a normal distribution. In the general case, it suffices to approximate  $h$  by piecewise constant functions and to use assertion (i) of Proposition 7.4.1.

We now prove that the support of the distribution for  $\{z(t), 0 \leq t \leq 1\}$  coincides with the space  $\dot{C}(0, 1)$ . To this end, it suffices to show that for any function  $g \in C^1(0, 1)$  vanishing at zero and any  $\varepsilon > 0$  we have

$$\mathbb{P}\left\{\sup_{0 \leq t \leq 1} |z(t) - g(t)| < \varepsilon\right\} > 0. \quad (7.5)$$

Let  $f \in C^1(0, 1)$  be a function vanishing at zero such that

$$g(t) = \int_0^t e^{\gamma(t-s)} f'(s) ds.$$

Integrating by parts, we can write

$$|z(t) - g(t)| \leq |\beta(t) - f(t)| + |\gamma| \int_0^t e^{\gamma(t-s)} |\beta(s) - f(s)| ds.$$

It follows that if

$$\omega \in A_\delta := \left\{ \sup_{0 \leq t \leq 1} |\beta(t) - f(t)| < \delta \right\},$$

then  $|z(t) - g(t)| \leq \delta(1 + e^{|\gamma|})$  for any  $t \in [0, 1]$ . Hence, the required inequality (7.5) will be established if we show that  $\mathbb{P}(A_\delta) > 0$  for any  $\delta > 0$ . This property is an immediate consequence of the fact that the support of the law for  $\{\beta(t), 0 \leq t \leq 1\}$  is the space  $\dot{C}(0, 1)$ .  $\square$

*Exercise 7.4.4.* Prove Lemma 7.4.3. *Hint:* If  $\varphi$  is continuously differentiable, then (7.4) is obvious. In the general case, construct a sequence of  $C^1$ -functions with uniformly bounded variations that converges pointwise to  $\varphi$  and apply Helly's first theorem (see [KF75]).

## 7.5 Weak convergence of random measures

Measurable isomorphisms between Polish spaces (see Section 7.2) need not to preserve the topology. In particular, the properties of continuity of functions and of weak convergence of measures are not invariant under such isomorphisms. The following result shows that any Polish space is homeomorphic to a subset of a compact metric space.

**Theorem 7.5.1.** *Let  $X$  be a Polish space. Then there is a metric  $d$  on  $X$  such that the topology defined by  $d$  coincides with the original one and  $(X, d)$  is relatively compact, that is, the completion  $\bar{X}$  of  $(X, d)$  is compact. Moreover,  $\mathcal{B}(X)$  is formed by intersections with  $X$  of the sets from  $\mathcal{B}(\bar{X})$ .*

*Proof.* For the first assertion, see Theorem 2.8.2 in [Dud02]. To prove the second, denote by  $\mathcal{C}$  the intersections with  $X$  of the sets from  $\mathcal{B}(\bar{X})$ . This is a  $\sigma$ -algebra. Denote by  $i : X \rightarrow \bar{X}$  the natural embedding. For any closed set  $F \subset X$ , we have  $i^{-1}(i(\bar{F})) = F$ . That is, the  $\sigma$ -algebra  $\mathcal{C}$  contains all closed sets. So it equals  $\mathcal{B}(X)$ .  $\square$

The well-known *Alexandrov theorem* says that if a sequence of measures  $\{\mu_k\}$  on a Polish space  $X$  is such that  $(f, \mu_k)$  converges for any bounded continuous function  $f : X \rightarrow \mathbb{R}$ , then the sequence  $\{\mu_k\}$  weakly converges to a limit (see [Ale43] or the corollary of Theorem 1 in [GS80, Section VI.1]). The following theorem due to Berti, Pratelli, and Rigo [BPR06] establishes a similar result for random probability measures.

**Theorem 7.5.2.** *Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space and let  $\{\mu_\omega^k, \omega \in \Omega\}$  be a sequence of random probability measures on a Polish space  $X$ , endowed with the Borel  $\sigma$ -algebra, such that the sequence  $\{(f, \mu_\omega^k)\}$  converges almost surely<sup>2</sup> for any  $f \in C_b(X)$ . Then there is a random probability measure  $\{\mu_\omega\}$  such that*

$$\mu_\omega^k \rightarrow \mu_\omega \quad \text{as } k \rightarrow \infty \text{ for almost every } \omega \in \Omega. \quad (7.6)$$

*Proof. Step 1.* In view of Theorem 7.5.1, there is no loss of generality in assuming that  $X$  is a subset of a compact metric space  $Y$  such that  $\Gamma \cap X \in \mathcal{B}(X)$  for any  $\Gamma \in \mathcal{B}(Y)$ . Let us define a sequence of random measures on  $Y$  by the relation

$$\hat{\mu}_\omega^k(\Gamma) = \mu_\omega^k(\Gamma \cap X), \quad \Gamma \in \mathcal{B}(Y).$$

Since  $Y$  is compact, for any  $\omega \in \Omega$  there is a sequence  $k_n \rightarrow \infty$  and a measure  $\hat{\mu}_\omega \in \mathcal{P}(Y)$  such that  $\hat{\mu}_\omega^{k_n} \rightarrow \hat{\mu}_\omega$ . It is well known that the space  $C_b(Y)$  is separable (see Exercise 1.2.4 or Corollary 11.2.5 in [Dud02]). Let  $\{\hat{f}_j\}$  be a dense sequence in  $C_b(Y)$ . Then a sequence of measures  $\nu_k \in \mathcal{P}(Y)$  converges weakly to  $\nu \in \mathcal{P}(Y)$  if and only if  $(\hat{f}_j, \nu_k) \rightarrow (\hat{f}_j, \nu)$  as  $k \rightarrow \infty$  for any  $j \geq 1$ .

Let us denote by  $f_j$  the restriction of  $\hat{f}_j$  to  $X$ , and let  $\Omega_0$  be the set of  $\omega \in \Omega$  for which the sequence  $\{(f_j, \mu_\omega^k)\}_{k \geq 1}$  converges for any  $j \geq 1$ . By assumption,

<sup>2</sup>We emphasise that the set of convergence may depend on  $f$ .

we have  $\mathbb{P}(\Omega_0) = 1$ . Since

$$(\hat{f}_j, \hat{\mu}_\omega^k) = (f_j, \mu_\omega^k) \quad \text{for any } j, k \geq 1,$$

we conclude that

$$\hat{\mu}_\omega^k \rightarrow \hat{\mu}_\omega \quad \text{as } k \rightarrow \infty \text{ for any } \omega \in \Omega_0. \quad (7.7)$$

Suppose we have constructed  $\mathcal{C} \subset X$  and a full-measure set  $\Omega_1 \subset \Omega_0$  such that  $\mathcal{C} \in \mathcal{B}(Y)$  and

$$\hat{\mu}_\omega(\mathcal{C}) = 1 \quad \text{for } \omega \in \Omega_1. \quad (7.8)$$

In this case, setting

$$\mu_\omega = \begin{cases} \hat{\mu}_\omega(\cdot \cap X) & \text{for } \omega \in \Omega_1, \\ \gamma & \text{otherwise,} \end{cases}$$

where  $\gamma \in \mathcal{P}(X)$  is an arbitrary measure, we see that (7.7) implies (7.6).

*Step 2.* To prove (7.8), note that, by the Lebesgue theorem on dominated convergence, for any  $f \in C_b(X)$ , the sequence  $\mathbb{E}(f, \mu_\omega^k)$  converges. By Alexandrov's theorem, it follows that the measures  $\mathbb{E} \mu_\omega^k$  converge weakly in  $\mathcal{P}(X)$ . On the other hand, in view of (7.7), we have  $\mathbb{E} \hat{\mu}_\omega^k \rightarrow \mathbb{E} \mu_\omega$  as  $k \rightarrow \infty$ .

Let us fix any constant  $\varepsilon > 0$  and use Prokhorov's compactness criterion to find a compact set  $\mathcal{C}_\varepsilon \subset X$  such that

$$\mathbb{E} \mu_\omega^k(\mathcal{C}_\varepsilon) \geq 1 - \varepsilon \quad \text{for any } k \geq 1.$$

Since  $\mathcal{C}_\varepsilon$  is closed in  $Y$ , we obtain

$$\mathbb{E} \hat{\mu}_\omega(X) \geq \mathbb{E} \hat{\mu}_\omega(\mathcal{C}_\varepsilon) \geq \limsup_{k \rightarrow \infty} \mathbb{E} \hat{\mu}_\omega^k(\mathcal{C}_\varepsilon) = \limsup_{k \rightarrow \infty} \mathbb{E} \mu_\omega^k(\mathcal{C}_\varepsilon) \geq 1 - \varepsilon.$$

Recalling that  $\varepsilon > 0$  was arbitrary and defining  $\mathcal{C} = \cup_n \mathcal{C}_{1/n}$ , we conclude that  $\mathbb{E} \hat{\mu}_\omega(\mathcal{C}) = 1$ , whence follows (7.8). The proof is complete.  $\square$

## 7.6 Gelfand triple and Yosida approximation

Let  $H$  and  $V$  be two separable real Hilbert<sup>3</sup> spaces such that  $V$  is continuously and densely embedded into  $H$ . We denote by  $(\cdot, \cdot)_H$  and  $(\cdot, \cdot)_V$  the scalar products in  $H$  and  $V$ , respectively, and by  $\|\cdot\|_H$  and  $\|\cdot\|_V$  the corresponding norms. The space  $H$  can be regarded as a dense subspace of the dual space  $V^*$ . Indeed, each element  $u \in H$  defines a continuous linear functional  $\ell_u : V \rightarrow \mathbb{R}$  by the relation

$$\langle \ell_u, v \rangle = (u, v)_H \quad \text{for } v \in V, \quad (7.9)$$

where  $\langle \cdot, \cdot \rangle$  stands for the pairing between a Hilbert space and its dual. Since the embedding  $V \subset H$  is continuous, we have

$$\|\ell_u\|_{V^*} = \sup_{\|v\|_V \leq 1} |(u, v)_H| \leq C \|u\|_H.$$

<sup>3</sup>The construction below also applies if  $V$  is a reflexive Banach space.

Furthermore, if  $u_1 \neq u_2$ , then  $\ell_{u_1} \neq \ell_{u_2}$ , because  $V$  is dense in  $H$ . Finally, to show that  $H$  is dense in  $V^*$ , assume that  $f : V^* \rightarrow \mathbb{R}$  is a continuous linear functional vanishing on  $H$ . Then

$$\langle f, \ell_u \rangle = \langle \ell_u, v_f \rangle = (u, v_f)_H = 0 \quad \text{for any } u \in H, \quad (7.10)$$

where  $v_f \in V$  is the element corresponding to  $f$  in the natural isomorphism between  $V^{**}$  and  $V$ . Relation (7.10) implies that  $v_f = 0$  and, hence,  $f = 0$ .

Thus, we have dense and continuous embeddings

$$V \subset H \subset V^*. \quad (7.11)$$

In what follows, any separable Hilbert spaces satisfying these conditions are called a *Gelfand triple* and  $H$  is called a *rigged Hilbert space*.

We now describe a general construction, called Yosida approximation, that enables one to approach the elements of  $H$  and  $V^*$  by those of  $V$  in a regular way. By the Riesz representation theorem, there is a unique isometry  $L : V \rightarrow V^*$  such that

$$\langle Lu, v \rangle = \langle Lv, u \rangle = (u, v)_V \quad \text{for any } u, v \in V. \quad (7.12)$$

**Proposition 7.6.1.** *The operator  $I + \varepsilon L$  taking  $v \in V$  to  $v + \varepsilon Lv \in V^*$  is invertible for any  $\varepsilon > 0$ , and its inverse is continuous. Moreover, for  $X = H, V$ , or  $V^*$  and  $u \in X$ , we have*

$$\|(I + \varepsilon L)^{-1}u\|_X \leq \|u\|_X \quad \text{for } \varepsilon > 0, \quad (7.13)$$

$$(I + \varepsilon L)^{-1}u \rightarrow u \quad \text{in } X \text{ as } \varepsilon \rightarrow 0^+. \quad (7.14)$$

*Proof.* Relations (7.9) and (7.12) imply that

$$\langle (I + \varepsilon L)u, u \rangle = \|u\|_H^2 + \varepsilon \|u\|_V^2 \quad \text{for any } u \in V, \quad (7.15)$$

and therefore the operator  $I + \varepsilon L$  is injective. To prove that it is surjective, we fix  $f \in V^*$  and consider the equation  $(I + \varepsilon L)u = f$  for  $u \in V$ . In view of (7.12), it is equivalent to

$$(u, v)_\varepsilon := (u, v)_H + \varepsilon (u, v)_V = \langle f, v \rangle \quad \text{for any } v \in V. \quad (7.16)$$

Endowing the space  $V$  with the new scalar product  $(\cdot, \cdot)_\varepsilon$ , we can use the Riesz representation theorem to find, for any  $\varepsilon > 0$ , a unique element  $u_\varepsilon \in V$  satisfying (7.16). Thus,  $I + \varepsilon L$  is a bijection of  $V$  onto  $V^*$ . The Banach inverse mapping theorem implies that  $(I + \varepsilon L)^{-1} : V^* \rightarrow V$  is continuous.

We now prove (7.13) and (7.14) for  $X = H$ ; the proof in the two other cases is similar. Let  $u \in H$  and  $u_\varepsilon = (I + \varepsilon L)^{-1}u$ . Then  $(I + \varepsilon L)u_\varepsilon = u$ . Applying both sides of this relation to  $u_\varepsilon$  and using (7.15), we obtain

$$\|u_\varepsilon\|_H^2 + \varepsilon \|u_\varepsilon\|_V^2 = \langle u, u_\varepsilon \rangle = (u, u_\varepsilon)_H \leq \|u\|_H \|u_\varepsilon\|_H,$$

whence (7.13) follows immediately. To prove convergence (7.14), we first note that it suffices to consider the case in which  $u \in V$ ; the general case can be

treated by a standard approximation argument combined with (7.13). Let  $u \in V$  and let

$$\delta_\varepsilon := (I + \varepsilon L)^{-1}u - u = -\varepsilon(I + \varepsilon L)^{-1}Lu.$$

Setting  $g = Lu$  and  $g_\varepsilon = (I + \varepsilon L)^{-1}g$  and using again relation (7.15), we derive

$$\|g_\varepsilon\|_H^2 + \varepsilon \|g_\varepsilon\|_V^2 = \langle g, g_\varepsilon \rangle \leq \|g\|_{V^*} \|g_\varepsilon\|_V \leq \frac{1}{4\varepsilon} \|g\|_{V^*}^2 + \varepsilon \|g_\varepsilon\|_V^2,$$

whence it follows that

$$\|g_\varepsilon\|_H^2 \leq \frac{1}{4\varepsilon} \|g\|_{V^*}^2.$$

Recalling the definition of  $\delta_\varepsilon$ , we obtain

$$\|\delta_\varepsilon\|_H = \varepsilon \|g_\varepsilon\|_H \leq \frac{\sqrt{\varepsilon}}{2} \|g\|_{V^*} \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0^+.$$

This completes the proof of the proposition.  $\square$

*Exercise 7.6.2.* Prove (7.13) and (7.14) for  $X = V$  and  $X = V^*$ . *Hint:* To prove (7.13), apply both sides of the relation  $(I + \varepsilon L)u_\varepsilon = u$  regarded as elements of  $V^*$  to  $Lu_\varepsilon$  and  $L^{-1}u_\varepsilon$ . To establish (7.14), modify the argument used in the above proof.

## 7.7 Itô formula in Hilbert spaces

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space with a right-continuous filtration  $\{\mathcal{F}_t, t \geq 0\}$ . We assume that  $\mathcal{F}_t$  is augmented with respect to  $(\mathcal{F}, \mathbb{P})$ , that is, the  $\sigma$ -algebra  $\mathcal{F}_t$  contains all  $\mathbb{P}$ -null sets of  $\mathcal{F}$ . Let  $H$  be a separable Hilbert space, let  $\{u_t, t \geq 0\}$  be a random process in  $H$ , and let  $\{\beta_j(t), j \geq 1\}$  be a sequence of independent standard Brownian motions. We assume that all these processes are defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  and are  $\mathcal{F}_t$ -adapted, and that almost all trajectories of  $u_t$  are continuous in time.

**Definition 7.7.1.** We shall say that  $\{u_t\}$  is an *Itô process in  $H$  with constant diffusion* if it can be represented in the form

$$u_t = u_0 + \int_0^t f_s ds + \sum_{j=1}^{\infty} \beta_j(t) g_j, \quad t \geq 0, \quad (7.17)$$

where the equality holds almost surely,  $\{g_j\} \subset H$  is a sequence satisfying the condition

$$\sum_{j=1}^{\infty} \|g_j\|_H^2 < \infty, \quad (7.18)$$

and  $f_t$  is an  $\mathcal{F}_t$ -progressively measurable  $H$ -valued process such that

$$\mathbb{P} \left\{ \int_0^T \|f_t\|_H dt < \infty \text{ for any } T > 0 \right\} = 1. \quad (7.19)$$



Before formulating a first result, we make some comments on the above definition. Conditions (7.18) and (7.19) ensure that the right-hand side of (7.17) makes sense. Indeed, the integral is almost surely finite due to (7.19), and the series converges in  $L^2(\Omega; H)$  for any fixed  $t > 0$ , because

$$\mathbb{E} \left\| \sum_{j=m}^n \beta_j(t) g_j \right\|_H^2 = \mathbb{E} \sum_{j=m}^n \|g_j\|_H^2 \beta_j^2(t) = t \sum_{j=m}^n \|g_j\|_H^2 \rightarrow 0 \quad \text{as } m, n \rightarrow \infty.$$

Moreover, the Doob–Kolmogorov inequality (7.55) implies that

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq T} \left\| \sum_{j=m}^{\infty} \beta_j(t) g_j \right\|_H > \varepsilon \right\} \leq \varepsilon^{-2} T \sum_{j=m}^{\infty} \|g_j\|_H^2 \rightarrow 0 \quad \text{as } m \rightarrow \infty,$$

and therefore the series on the right-hand side of (7.17) converges in probability in the space  $C(0, T; H)$ . It defines an  $H$ -valued process whose almost all trajectories are continuous in time.

**Theorem 7.7.2.** *Let  $J = [0, T]$ , let  $H$  be a separable real Hilbert space, let  $F : J \times H \rightarrow \mathbb{R}$  be a twice continuously differentiable function that is bounded and uniformly continuous together with its derivatives, and let  $\{u_t, t \geq 0\}$  be an Itô process in  $H$  with constant diffusion. Then*

$$F(t, u_t) = F(0, u_0) + \int_0^t A(s) ds + \sum_{j=1}^{\infty} \int_0^t B_j(s) d\beta_j(s), \quad t \in J, \quad (7.20)$$

where we set

$$A(t) = (\partial_t F)(t, u_t) + (\partial_u F)(t, u_t; f_t) + \frac{1}{2} \sum_{j=1}^{\infty} (\partial_u^2 F)(t, u_t; g_j), \quad (7.21)$$

$$B_j(t) = (\partial_u F)(t, u_t; g_j), \quad (7.22)$$

and  $(\partial_u F)(u; v)$  and  $(\partial_u^2 F)(u; v)$  denote the values of the first and second derivatives of  $F$  with respect to  $u$  at a point  $v \in H$ .

This result can be established by literal repetition of the derivation of Itô's formula in the finite-dimensional case. A detailed proof of a more general result can be found in [DZ92].

*Exercise 7.7.3.* Show that all the terms on the right-hand side of (7.20) are almost surely continuous functions of time.

As in the case of finite-dimensional Itô processes, relation (7.20) remains valid for a more general class of functionals. Namely, let us assume that  $F(t, u)$  is twice continuously differentiable on  $\mathbb{R}_+ \times H$ , and  $F$  is uniformly continuous, together with its derivatives, on any bounded subset of  $\mathbb{R}_+ \times H$ . In this case, relation (7.20) holds in the following sense. Let us introduce the stopping times<sup>4</sup>

$$\tau_n = \inf \{t \geq 0 : \|u_t\|_H > n\}, \quad (7.23)$$

<sup>4</sup>Note that, since the trajectories of  $u_t$  are continuous, the stopping times  $\tau_n$  go to  $+\infty$  for almost all  $\omega$ .

where the infimum of an empty set is equal to  $+\infty$ . Then

$$F(t \wedge \tau_n, u_{t \wedge \tau_n}) = F(0, u_0) + \int_0^{t \wedge \tau_n} A(s) ds + \sum_{j=1}^{\infty} \int_0^{t \wedge \tau_n} B_j(s) d\beta_j(s), \quad t \geq 0. \quad (7.24)$$

It is not difficult to see that all the terms on the right-hand side of (7.24) make sense. We refer the reader to [Kry02, Chapter III] and [DZ92] for a justification of (7.24) in the finite- and infinite-dimensional cases, respectively.

The above version of Itô's formula is not applicable to study solutions of nonlinear stochastic PDE's. Indeed, these solutions are random processes in a function space and Itô processes in some bigger space, while very often one needs to study the behaviour of functionals defined on the former space. In literature, there are versions of Itô's formula that apply to some nonlinear PDE's (e.g., see [Roz90]). However, they are not flexible enough to treat all the situations encountered in this book.

Below we suggest a version of Itô formula sufficient for our purposes. To state the result, we assume that  $H$  is a rigged Hilbert space. More precisely, let  $V \subset H \subset V^*$  be a Gelfand triple and let  $\{u_t, t \geq 0\}$  be an  $H$ -valued random process. We assume that it satisfies the following hypothesis.

**Condition 7.7.4.** Almost all trajectories of  $\{u_t\}$  are continuous with range in  $H$  and locally square integrable with range in  $V$ . Moreover,  $\{u_t\}$  is an Itô process in  $V^*$  (rather than in  $H$ ) with constant diffusion (see Definition 7.7.1) such that  $g_j \in H$  for any  $j \geq 1$ , inequality (7.18) holds, and (7.19) is replaced by the stronger condition

$$\mathbb{P} \left\{ \int_0^T \|f_t\|_{V^*}^2 dt < \infty \right\} = 1. \quad (7.25)$$

Note that, due to (7.18), the diffusion term in (7.17) is a continuous  $H$ -valued function of  $t \geq 0$  for almost all  $\omega \in \Omega$ . Combining this with the continuity of  $u_t$ , we see that the integral  $\int_0^t f_s ds$  is also continuous with range in  $H$ . However, it is not absolutely continuous in general, and the function  $f_t$  is locally (square) integrable only with range in  $V^*$ .

**Theorem 7.7.5.** *Let  $F : \mathbb{R}_+ \times H \rightarrow \mathbb{R}$  be a twice continuously differentiable function that is uniformly continuous, together with its derivatives, on any bounded subset of  $\mathbb{R}_+ \times H$ . Assume also that  $F$  satisfies the following two conditions.*

- (i) *For any  $T > 0$  there is a positive continuous function  $K_T(r)$ ,  $r \geq 0$ , such that<sup>5</sup>*

$$|(\partial_u F)(t, u; v)| \leq K_T(\|u\|_H) \|u\|_V \|v\|_{V^*} \quad \text{for } t \in [0, T], u \in V, v \in V^*. \quad (7.26)$$

<sup>5</sup>Inequality (7.26) implies that the derivative  $(\partial_u F)(t, u)$  defined initially on  $H$  admits a continuous extension to  $V^*$  for any  $u \in V$ .

(ii) For any sequence  $\{w_k\} \subset V$  converging to  $w \in V$  in the topology of  $V$  and any  $t \in \mathbb{R}_+$  and  $v \in V^*$ , we have

$$(\partial_u F)(t, w_k; v) \rightarrow (\partial_u F)(t, w; v) \quad \text{as } k \rightarrow \infty.$$

Then for an arbitrary random process  $\{u_t, t \geq 0\}$  satisfying Condition 7.7.4 relation (7.24) holds for almost every  $\omega \in \Omega$ .

Before proving this result, let us show that all the terms on the right-hand side of (7.24) make sense and for almost every  $\omega \in \Omega$  are continuous in time. Indeed, approximating  $f_t$  by piecewise constant functions and using (7.26), we easily show that  $(\partial_u F)(t, u_t; f_t)$  is a measurable function of  $(t, \omega)$ . Moreover, it follows from (7.21), (7.25), and (7.26) that

$$|A(t)| \leq C_1 + K_T(R_T)\|u_t\|_V\|f_t\|_{V^*} + C_2 \sum_{j=1}^{\infty} \|g_j\|_H^2 \leq C_3(1 + \|u_t\|_V^2 + \|f_t\|_{V^*}^2),$$

where  $R_T = \sup\{\|u_t\|_H, 0 \leq t \leq T\}$  and  $C_i$  are almost surely finite random constants. Hence, the function  $A$  is locally square integrable in time for almost all  $\omega \in \Omega$ , and its integral is continuous in time. Furthermore, if  $\|u_t\|_H \leq n$ , then  $|B_j(t)| \leq C_4(n)\|g_j\|_H$  for any  $j \geq 1$ , whence it follows that

$$\mathbb{E} \left| \sum_{j=1}^{\infty} \int_0^{t \wedge \tau_n} B_j(s) d\beta_j(s) \right|^2 \leq C_4^2(n) t \sum_{j=1}^{\infty} \|g_j\|_H^2 < \infty.$$

Thus, for any  $T > 0$ , the series on the right-hand side of (7.24) converges in  $L^2(\Omega)$  uniformly with respect to  $t \in [0, T]$ . Moreover, by the Doob–Kolmogorov inequality (7.55), for any  $\varepsilon > 0$  we have

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq T} \left| \sum_{j=m}^{\infty} \int_0^{t \wedge \tau_n} B_j(s) d\beta_j(s) \right| \geq \varepsilon \right\} \leq \varepsilon^{-2} C_4^2(n) T \sum_{j=m}^{\infty} \|g_j\|_H^2 \rightarrow 0$$

as  $m \rightarrow \infty$ , whence we conclude that the series of (7.24) converges in probability in the space  $C(0, T)$  for any  $T > 0$ , and its limit is almost surely continuous in time.

*Proof of Theorem 7.7.5.* Both left- and right-hand sides of (7.24) are continuous functions of time for almost all  $\omega \in \Omega$ . Therefore, if we show that, for any fixed  $t \geq 0$ , the equality in (7.24) holds almost surely, then a standard argument based on continuity will imply that (7.24) takes place for all  $t \geq 0$  outside a universal  $\mathbb{P}$ -null set.

Let us fix a constant  $\varepsilon > 0$  and define the vectors  $g_j^\varepsilon = (I + \varepsilon L)^{-1} g_j$  and the processes  $u_t^\varepsilon = (I + \varepsilon L)^{-1} u_t$  and  $f_t^\varepsilon = (I + \varepsilon L)^{-1} f_t$ , where  $L : V \rightarrow V^*$  is the operator satisfying (7.12). It follows from Proposition 7.6.1 that  $\{u_t^\varepsilon\}$  is an Itô process in  $H$  with constant diffusion. So, Itô's formula (7.24) holds for it:

$$F(t \wedge \tau_n^\varepsilon, u_{t \wedge \tau_n^\varepsilon}^\varepsilon) = F(0, u_0^\varepsilon) + \int_0^{t \wedge \tau_n^\varepsilon} A^\varepsilon(s) ds + \sum_{j=1}^{\infty} \int_0^{t \wedge \tau_n^\varepsilon} B_j^\varepsilon(s) d\beta_j(s), \quad t \geq 0, \quad (7.27)$$

where  $\tau_n^\varepsilon = \inf\{t \geq 0 : \|u_t^\varepsilon\|_H > n\}$ , and the processes  $A^\varepsilon$  and  $B_j^\varepsilon$  are defined by relations (7.21) and (7.22), where  $u_t$ ,  $f_t$ , and  $g_j$  are replaced by  $u_t^\varepsilon$ ,  $f_t^\varepsilon$ , and  $g_j^\varepsilon$ , respectively. We wish to pass to the limit in (7.27) as  $\varepsilon \rightarrow 0^+$ .

*Step 1.* Proposition 7.6.1 and the continuity of trajectories of  $u_t$  imply that, for almost all  $\omega \in \Omega$ , we have

$$\tau_n^\varepsilon \geq \tau_n \quad \text{for } \varepsilon > 0, \quad \tau_n^\varepsilon \rightarrow \tau_n \quad \text{as } \varepsilon \rightarrow 0^+. \quad (7.28)$$

Since  $F$  is a continuous function of its arguments, we conclude that for any  $s \geq 0$  and almost all  $\omega$ , we have

$$F(s \wedge \tau_n^\varepsilon, u_{s \wedge \tau_n^\varepsilon}^\varepsilon) \rightarrow F(s \wedge \tau_n, u_{s \wedge \tau_n}) \quad \text{as } \varepsilon \rightarrow 0^+. \quad (7.29)$$

It follows that we can pass to the limit in the left-hand side and the first term of the right-hand side of (7.27).

*Step 2.* We now study the first integral in (7.27). In view of Proposition 7.6.1, for almost all  $\omega$ , we have

$$\begin{aligned} \sup_{0 \leq t \leq T} |(\partial_t F)(t, u_t^\varepsilon) - (\partial_t F)(t, u_t)| &\rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0^+, \\ \sup_{0 \leq t \leq T} |(\partial_u^2 F)(t, u_t^\varepsilon; g_j^\varepsilon) - (\partial_u^2 F)(t, u_t; g_j)| &\leq C_5(\varepsilon) \|g_j\|_H^2 + C_6 \|g_j^\varepsilon - g_j\|_H^2, \end{aligned} \quad (7.31)$$

where  $C_5(\varepsilon)$  is a random constant going to zero as  $\varepsilon \rightarrow 0^+$ . Proposition 7.6.1, inequality (7.18), and Lebesgue's theorem on dominated convergence imply that

$$\sum_{j=1}^{\infty} \|g_j^\varepsilon - g_j\|_H^2 \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0^+. \quad (7.32)$$

Combining this with (7.31) and (7.28), we see that

$$\sum_{j=1}^{\infty} \left| \int_0^{t \wedge \tau_n^\varepsilon} (\partial_u^2 F)(s, u_s^\varepsilon; g_j^\varepsilon) ds - \int_0^{t \wedge \tau_n} (\partial_u^2 F)(s, u_s; g_j) ds \right| \rightarrow 0 \quad \text{a. s. as } \varepsilon \rightarrow 0^+.$$

Furthermore, it follows from inequality (7.26) that

$$\begin{aligned} &|(\partial_u F)(t, u_t^\varepsilon; f_t^\varepsilon) - (\partial_u F)(t, u_t; f_t)| \\ &\leq C_7 \|u_t^\varepsilon\|_V \|f_t^\varepsilon - f_t\|_{V^*} + |(\partial_u F)(t, u_t^\varepsilon; f_t) - (\partial_u F)(t, u_t; f_t)|. \end{aligned}$$

In view of (7.25) and Proposition 7.6.1, for any  $T > 0$ , we have

$$\int_0^T \|u_s^\varepsilon\|_V \|f_s^\varepsilon - f_s\|_{V^*} ds \leq \left( \int_0^T \|u_s^\varepsilon\|_V^2 ds \right)^{1/2} \left( \int_0^T \|f_s^\varepsilon - f_s\|_{V^*}^2 ds \right)^{1/2} \rightarrow 0$$

as  $\varepsilon \rightarrow 0^+$ . Using the Lebesgue theorem on dominated convergence and property (ii) of the function  $F$ , we derive

$$\int_0^{t \wedge \tau_n} |(\partial_u F)(t, u_t^\varepsilon; f_t) - (\partial_u F)(t, u_t; f_t)| ds \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0^+.$$

Finally, it follows from (7.28) and Proposition 7.6.1 that

$$\int_{t \wedge \tau_n}^{t \wedge \tau_n^\varepsilon} |(\partial_u F)(s, u_s^\varepsilon; f_s^\varepsilon)| ds \leq C_8(n) \int_{t \wedge \tau_n}^{t \wedge \tau_n^\varepsilon} \|u_s\|_V \|f_s\|_{V^*} ds \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0^+.$$

Combining all what has been said above, we obtain

$$\int_0^{t \wedge \tau_n^\varepsilon} A^\varepsilon(s) ds \rightarrow \int_0^{t \wedge \tau_n} A(s) ds \quad \text{as } \varepsilon \rightarrow 0^+ \text{ for almost all } \omega \in \Omega. \quad (7.33)$$

*Step 3.* It remains to study the sum on the right-hand side of (7.27). By Itô's isometry, we have

$$\begin{aligned} \mathbb{E} \left| \int_{t \wedge \tau_n}^{t \wedge \tau_n^\varepsilon} B_j^\varepsilon(s) d\beta_j(s) \right|^2 &\leq \int_0^t \mathbb{E} \mathbb{I}_{[\tau_n, \tau_n^\varepsilon]}(s) |B_j^\varepsilon(s)|^2 ds \\ &\leq C_9(n) \|g_j^\varepsilon\|_H^2 \mathbb{E} |t \wedge \tau_n^\varepsilon - t \wedge \tau_n|. \end{aligned}$$

Combining this with (7.28) and inequality (7.13) with  $X = H$ , and using the Lebesgue theorem on dominate convergence, we obtain

$$\mathbb{E} \left| \sum_{j=1}^{\infty} \int_{t \wedge \tau_n}^{t \wedge \tau_n^\varepsilon} B_j^\varepsilon(s) d\beta_j(s) \right|^2 \leq C_9(n) (\mathbb{E} |t \wedge \tau_n^\varepsilon - t \wedge \tau_n|) \sum_{j=1}^{\infty} \|g_j\|_H^2 \rightarrow 0 \quad (7.34)$$

as  $\varepsilon \rightarrow 0^+$ . Furthermore, since  $\partial_u F$  is uniformly continuous on bounded subsets, we have

$$\begin{aligned} \sup_{0 \leq s \leq t \wedge \tau_n} |B_j^\varepsilon(s) - B_j(s)| &\leq \|g_j\|_H \sup_{0 \leq s \leq t \wedge \tau_n} \|(\partial_u F)(s, u_s^\varepsilon) - (\partial_u F)(s, u_s)\|_{\mathcal{L}(H, \mathbb{R})} \\ &\quad + \|g_j^\varepsilon - g_j\|_H \sup_{0 \leq s \leq t \wedge \tau_n} \|(\partial_u F)(s, u_s)\|_{\mathcal{L}(H, \mathbb{R})} \\ &= C_{10}(n, \varepsilon) \|g_j\|_H + C_{11}(n) \|g_j^\varepsilon - g_j\|_H, \end{aligned}$$

where  $C_{10}(n, \varepsilon)$  is a random constant going to zero as  $\varepsilon \rightarrow 0^+$ . It follows that

$$\begin{aligned} \mathbb{E} \left| \sum_{j=1}^{\infty} \int_0^{t \wedge \tau_n} (B_j^\varepsilon(s) - B_j(s)) d\beta_j \right|^2 &= \sum_{j=1}^{\infty} \int_0^t \mathbb{E} \mathbb{I}_{[0, t \wedge \tau_n]}(s) |B_j^\varepsilon(s) - B_j(s)|^2 ds \\ &\leq 2C_{10}^2(n, \varepsilon) \sum_{j=1}^{\infty} \|g_j\|_H^2 + 2C_{11}^2(n) \sum_{j=1}^{\infty} \|g_j^\varepsilon - g_j\|_H^2. \end{aligned}$$

In view of (7.32), the right-hand side of this inequality goes to zero as  $\varepsilon \rightarrow 0^+$ . Combining this with (7.34), we find a sequence  $\varepsilon_k \rightarrow 0^+$  such that, for almost all  $\omega \in \Omega$ ,

$$\sum_{j=1}^{\infty} \int_0^{t \wedge \tau_n^\varepsilon} B_j^\varepsilon(s) d\beta_j(s) \rightarrow \sum_{j=1}^{\infty} \int_0^{t \wedge \tau_n} B_j(s) d\beta_j(s) \quad \text{as } \varepsilon = \varepsilon_k \rightarrow 0^+. \quad (7.35)$$

Comparing (7.29), (7.33), and (7.35) and passing to the limit as  $\varepsilon = \varepsilon_k \rightarrow 0^+$  in relation (7.27), we obtain (7.24). The proof of Theorem 7.7.5 is complete.  $\square$

**Corollary 7.7.6.** *Suppose that the conditions of Theorem 7.7.5 are satisfied, and*

$$\sum_{j=1}^{\infty} \mathbb{E} \int_0^t |B_j(s)|^2 ds < \infty \quad \text{for any } t \geq 0. \quad (7.36)$$

*Then the sum of stochastic integrals*

$$M_t = \sum_{j=1}^{\infty} \int_0^t B_j(s) d\beta_j(s), \quad t \geq 0,$$

*defines a square-integrable martingale with almost surely continuous trajectories, and with probability 1, we have*

$$F(t, u_t) = F(u_0) + \int_0^t A(s) ds + M_t, \quad t \geq 0. \quad (7.37)$$

*Proof.* The fact that (7.36) is a square-integrable martingale with continuous trajectories is a standard assertion in the theory of Itô's integral; e.g., see Chapter 3 in [Øks03]. When proving (7.37), as in the case of Theorem 7.7.5, it suffices to show that the equality holds almost surely for any fixed  $t \geq 0$ . We wish to pass to the limit in (7.24) as  $n \rightarrow \infty$ . Since  $\tau_n \rightarrow \infty$  as  $n \rightarrow \infty$  and  $u_t$  is continuous in time with range in  $H$ , for almost every  $\omega \in \Omega$  we have

$$F(t \wedge \tau_n, u_{t \wedge \tau_n}) \rightarrow F(t, u_t) \quad \text{as } n \rightarrow \infty. \quad (7.38)$$

Furthermore, it follows from (7.18) and (7.21) that  $A(\cdot)$  is a continuous function of time, and therefore,

$$\int_0^{t \wedge \tau_n} A(s) ds \rightarrow \int_0^t A(s) ds \quad \text{with probability 1 as } n \rightarrow \infty. \quad (7.39)$$

Finally, condition (7.36) and Itô's isometry (see [Øks03]) imply that

$$\begin{aligned} \varepsilon_n &= \mathbb{E} \left| \sum_{j=1}^{\infty} \int_0^{t \wedge \tau_n} B_j(s) d\beta_j(s) - \sum_{j=1}^{\infty} \int_0^t B_j(s) d\beta_j(s) \right|^2 \\ &= \mathbb{E} \int_0^t \sum_{j=1}^{\infty} |\mathbb{I}_{[0, t \wedge \tau_n]}(s) - 1|^2 |B_j(s)|^2 ds. \end{aligned}$$

Applying the Lebesgue theorem on dominated convergence, we see that  $\varepsilon_n \rightarrow 0$  as  $n \rightarrow \infty$ . Combining this with (7.38) and (7.39), we arrive at (7.37).  $\square$

## 7.8 Local time for continuous Itô processes

In this section, we discuss distributions of convex functionals evaluated on one-dimensional Itô processes. Recall that a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is said to be *convex* if

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \quad \text{for any } x, y \in \mathbb{R}, \alpha \in [0, 1]$$

It is well known that a convex function has left- and right-hand derivatives  $\partial^- f$  and  $\partial^+ f$  at any point  $x \in \mathbb{R}$ . Furthermore, these derivatives are non-decreasing functions, they satisfy the inequality  $\partial^- f(x) \leq \partial^+ f(x)$  for any  $x \in \mathbb{R}$ , and the equality breaks on a set which is at most countable. Finally, the second derivative  $\partial^2 f(x)$  in the sense of distributions is a positive measure, which satisfies the relation

$$\mu_f([a, b]) = \partial^- f(b) - \partial^+ f(a) \quad \text{for all } a < b.$$

The following result is a straightforward consequence of Theorem 7.1 in [KS91, Chapter 3].

**Theorem 7.8.1.** *Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space with filtration  $\mathcal{F}_t$ ,  $t \geq 0$ , let  $\{\beta_j\}$  be a sequence of independent Brownian motions with respect to  $\mathcal{F}_t$ , and let  $y_t$  be a scalar Itô process of the form*

$$y_t = y_0 + \int_0^t x_s ds + \sum_{j=1}^{\infty} \int_0^t \theta_s^j d\beta_j(s), \quad (7.40)$$

where  $x_t$  and  $\theta_t^j$  are  $\mathcal{F}_t$ -adapted processes such that

$$\mathbb{E} \int_0^t \left( |x_s| + \sum_{j=1}^{\infty} |\theta_s^j|^2 \right) ds < \infty \quad \text{for any } t > 0. \quad (7.41)$$

Then there is a random field  $\Lambda_t(a, \omega)$ ,  $t \geq 0$ ,  $a \in \mathbb{R}$ ,  $\omega \in \Omega$ , such that the following properties hold.

- (i) *The mapping  $(t, a, \omega) \mapsto \Lambda_t(a, \omega)$  is measurable, and for any  $a \in \mathbb{R}$  the process  $t \mapsto \Lambda_t(a, \omega)$  is  $\mathcal{F}_t$ -adapted, continuous and non-decreasing. Moreover, for every  $t \geq 0$  and almost every  $\omega \in \Omega$  the function  $a \mapsto \Lambda_t(a, \omega)$  is right-continuous.*
- (ii) *For any non-negative Borel-measurable function  $g : \mathbb{R} \rightarrow \mathbb{R}$ , with probability 1 we have*

$$\int_0^t g(y_s) \left( \sum_{j=1}^{\infty} |\theta_s^j|^2 \right) ds = 2 \int_{-\infty}^{\infty} g(a) \Lambda_t(a, \omega) da, \quad t \geq 0. \quad (7.42)$$

- (iii) *For any convex function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , with probability 1 we have*

$$\begin{aligned} f(y_t) &= f(y_0) + \sum_{j=1}^{\infty} \int_0^t \partial^- f(y_s) \theta_s^j d\beta_j \\ &\quad + \int_0^t \partial^- f(y_s) x_s ds + \int_{-\infty}^{\infty} \Lambda_t(a, \omega) \partial^2 f(da), \quad t \geq 0. \end{aligned} \quad (7.43)$$

The random field  $\Lambda_t(a, \omega)$  constructed in the above theorem is called a *local time* for  $y_t$ , and relation (7.43) is usually referred to as the *change of variable formula*. This is an analogue of Itô's formula (7.20) for one-dimensional processes and non-smooth convex functionals. Let us comment on properties (ii) and (iii). Taking  $g(x) = \mathbb{I}_{[\alpha, \beta]}(x)$  in (7.42), we obtain

$$\int_0^t \mathbb{I}_{[\alpha, \beta]}(y_s) \left( \sum_{j=1}^{\infty} |\theta_s^j|^2 \right) ds = 2 \int_{\alpha}^{\beta} \Lambda_t(a, \omega) da. \quad (7.44)$$

Assuming that the quadratic variation of the diffusion term is bounded from above and separated from zero, we see that

$$\ell(\{s \in [0, t] : \alpha \leq y_s \leq \beta\}) \sim \int_{\alpha}^{\beta} \Lambda_t(a, \omega) da.$$

Thus, one can say roughly that  $\Lambda_t(a, \omega)$  measures the fraction of time spent by the process  $y_s$ ,  $0 \leq s \leq t$ , in the vicinity of  $a$ .

Furthermore, relation (7.43) enables one to find an explicit formula for  $\Lambda_t(a, \omega)$ . Indeed, taking  $f(x) = (x - a)_+ = \max(0, x - a)$ , we obtain

$$(y_t - a)_+ = (y_0 - a)_+ + \sum_{j=1}^{\infty} \int_0^t \mathbb{I}_{[a, +\infty)}(y_s) \theta_s^j d\beta_j + \int_0^t \mathbb{I}_{[a, +\infty)}(y_s) x_s ds + \Lambda_t(a, \omega). \quad (7.45)$$

## 7.9 Krylov's estimate

There is no analogue of the concept of local time for multidimensional Itô processes, so a multidimensional version of formula (7.44) is not available. In many cases, however, it is possible to get rather sharp estimates for a process' law, using Itô's formula and martingales inequalities. This type of results were obtained by Krylov; see [Kry74, Kry86] and the book [Kry80]. In Chapter 5, we use a simplified version of such an estimate. Namely, consider a  $d$ -dimensional Itô process written in the form (7.40), where  $x_t$  and  $\theta_t^j$  are adapted  $d$ -dimensional processes satisfying condition (7.41), and  $\beta_j$  are independent Brownian motions. We shall denote by  $\theta_t^{jk}$ ,  $k = 1, \dots, d$ , the components of the vector function  $\theta_t^j$  and by  $\sigma_t$  the diffusion, which is a  $d \times d$  symmetric matrix with the entries

$$\sigma_t^{kl} = \sum_{j=1}^{\infty} \theta_t^{jk} \theta_t^{jl}, \quad k, l = 1, \dots, d.$$

Note that  $\sigma_t$  is non-negative for any  $t \geq 0$ . To simplify the presentation, we shall assume that  $y_t$  is a stationary process; this is the situation considered in Chapter 5.



**Theorem 7.9.1.** *Under the above hypotheses, there is a constant  $C_d > 0$  depending only on the dimension  $d$  such that, for any bounded measurable function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , we have*

$$\mathbb{E} \int_0^1 (\det \sigma_t)^{1/d} f(x_t) dt \leq C_d |f|_d \mathbb{E} \int_0^1 |x_s| ds, \quad (7.46)$$

where  $|f|_d$  stands for the  $L^d$  norm of  $f$ .

*Proof.* A standard argument based on the monotone class theorem and Fatou's lemma shows that it suffices to establish inequality (7.46) for smooth functions  $f \geq 0$  with compact support. The general idea of the proof is rather simple: we first apply Itô's formula to the stationary semimartingale  $\psi(y_t)$  and derive an identity for it and then, choosing a particular function  $\psi$ , we obtain an expression minorised by the left-hand side of (7.46).

Let  $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$  be a twice continuously differentiable function with bounded derivatives. By Itô's formula, we have

$$\psi(y_t) = \psi(y_0) + \int_0^t \left( (\nabla \psi(y_s), dy_s) + \frac{1}{2} \sum_{k,l=1}^d \sigma_s^{kl} (\partial_k \partial_l \psi)(y_s) ds \right).$$

Taking the mean value, using the stationarity of  $y_t$ , and setting  $t = 1$ , we get

$$2 \mathbb{E} \int_0^1 (\nabla \psi(y_s), x_s) ds + \mathbb{E} \int_0^1 \sum_{k,l=1}^d \sigma_s^{kl} (\partial_k \partial_l \psi)(y_s) ds = 0. \quad (7.47)$$

We now need the following proposition, whose proof<sup>6</sup> can be found in [Kry87] (see Theorem III.2.3).

**Proposition 7.9.2.** *For any  $\lambda > 0$  and any non-negative function  $f \in C_0^\infty(\mathbb{R}^d)$  there is a function  $\psi \in C^2(\mathbb{R}^d)$  with bounded derivatives such that*

$$\sum_{k,l=1}^d a_{kl} (\partial_k \partial_l \psi)(x) - \lambda (\text{Tr } a) \psi(x) + (\det a)^{1/d} f(x) \leq 0, \quad (7.48)$$

$$|\nabla \psi(x)| \leq \lambda^{1/2} \psi(x), \quad \psi(x) \leq A_d \lambda^{-1/2} |f|_1, \quad (7.49)$$

where  $a = (a_{kl})$  is an arbitrary non-negative symmetric  $d \times d$  matrix,  $A_d$  is a constant depending only on  $d$ , and the inequalities hold for all  $x \in \mathbb{R}^d$ .

Let us take for  $\psi$  in (7.47) the function constructed in Proposition 7.9.2. Recalling that  $\sigma_t$  is a non-negative matrix for any  $t \geq 0$  and taking into account (7.48), we derive

$$\mathbb{E} \int_0^1 (\det \sigma_t)^{1/d} f(y_s) ds \leq 2 \mathbb{E} \int_0^1 (\nabla \psi(y_s), x_s) ds + \lambda \mathbb{E} \int_0^1 \psi(y_s) (\text{Tr } \sigma_s) ds. \quad (7.50)$$

<sup>6</sup>Note that Proposition 7.9.2 is obvious for  $d = 1$ . In this case, it suffices to take for  $\psi$  the unique bounded solution of the equation  $\psi'' - \psi + f = 0$ .

Using inequalities (7.49), we see that the right-hand side of (7.50) can be estimated from above by

$$C_1|f|_1\mathbb{E}\int_0^1|x_s|ds+C_2\sqrt{\lambda}|f|_1\mathbb{E}\int_0^1\text{Tr}\sigma_s ds,$$

where  $C_1$  and  $C_2$  are some constants depending only on  $d$ . Substituting this expression into (7.50) and passing to the limit as  $\lambda \rightarrow 0^+$ , we arrive at the required inequality (7.46).  $\square$

## 7.10 Girsanov theorem

In this section, we deal with processes defined on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with a right-continuous filtration  $\{\mathcal{F}_t, t \geq 0\}$ . The following result is a simple consequence of Girsanov's theorem (see [Øks03]).

**Theorem 7.10.1.** *Let  $y$  be an  $\mathbb{R}^N$ -valued process of the form*

$$y(t) = \sum_{j=1}^N b_j \beta_j(t) e_j, \quad t \geq 0,$$

where  $\{e_j\}$  is an orthonormal basis in  $\mathbb{R}^N$ ,  $\beta_j$  are independent standard Brownian motions, and  $b_j > 0$  are some constants. Let  $\tilde{y}$  be another process given by

$$\tilde{y}(t) = y(t) + \int_0^t a(s) ds,$$

where  $a = (a_1, \dots, a_N)$  is a progressively measurable process such that

$$\mathbb{E} \exp\left(C \int_0^\infty |a(t)|^2 dt\right) < \infty \quad \text{for any } C > 0. \quad (7.51)$$

Then the distributions of the processes  $y$  and  $\tilde{y}$  regarded as random variables in  $C(\mathbb{R}_+, \mathbb{R}^N)$  satisfy the inequality

$$\|\mathcal{D}(y) - \mathcal{D}(\tilde{y})\|_{\text{var}} \leq \frac{1}{2} \left( \left( \mathbb{E} \exp\left\{6 \int_0^\infty \sum_{j=1}^N b_j^{-2} |a_j(t)|^2 dt\right\} \right)^{1/2} - 1 \right)^{1/2}. \quad (7.52)$$

*Proof.* Let us consider an exponential supermartingale  $M_t$  corresponding to the stochastic integral  $\xi_t = -\int_0^t \sum_{j=1}^N b_j^{-1} a_j(t) d\beta_j(t)$ . That is,

$$M_t(\omega) = \exp\left(-\int_0^t \sum_{j=1}^N b_j^{-1} a_j(t) d\beta_j(t) - \frac{1}{2} \int_0^t \sum_{j=1}^N b_j^{-2} |a_j(t)|^2 dt\right), \quad t \geq 0.$$

By Girsanov's theorem (see Theorem 8.6.5 in [Øks03]), condition (7.51) implies that this is a martingale. So  $\mathbb{E} M_t = \mathbb{E} M_0 = 1$  for each  $t \geq 0$ . Since  $M_t \geq 0$ , by

the *convergence theorem* of Section 7.11, there exists a limit  $\lim_{t \rightarrow \infty} M_t = M$ , and  $\mathbb{E}M = 1$ . Let us consider the probability measure  $\mathbb{Q} = M(\omega)\mathbb{P}$  on the space  $(\Omega, \mathcal{F})$ . Due to Girsanov's theorem (see [Øks03]), the  $\mathbb{P}$ -law of  $y$  coincides with the  $\mathbb{Q}$ -law of  $\tilde{y}$ . That is,  $\mathcal{D}(y) = y_*(\mathbb{P}) = \tilde{y}_*(\mathbb{Q})$  and  $\mathcal{D}(\tilde{y}) = \tilde{y}_*(\mathbb{P})$ . By Exercise 1.2.13 (i), the left-hand side of (7.52) can be estimated by  $\|\mathbb{P} - \mathbb{Q}\|_{\text{var}}$ . Proposition 1.2.7 implies that the latter quantity does not exceed

$$\frac{1}{2}\mathbb{E}|M - 1| \leq \frac{1}{2}(\mathbb{E}M^2 - 1)^{1/2}. \quad (7.53)$$

To estimate the right-hand side, let us note that

$$\mathbb{E}M_t^2 = \mathbb{E}e^{2\xi_t - \langle \xi \rangle_t} \leq (\mathbb{E}e^{4\xi_t - 8\langle \xi \rangle_t})^{1/2} (\mathbb{E}e^{6\langle \xi \rangle_t})^{1/2} \leq (\mathbb{E}e^{6\langle \xi \rangle_t})^{1/2}$$

for any  $t \geq 0$ , where we used that  $\exp(4\xi_t - 8\langle \xi \rangle_t)$  is the exponential supermartingale corresponding to  $\xi_t$ . By the above inequality and the monotone convergence theorem, we have  $\mathbb{E}M^2 \leq (\mathbb{E}e^{6\langle \xi \rangle_\infty})^{1/2}$ . Substituting this inequality into (7.53), we arrive at the required result.  $\square$

## 7.11 Martingales, submartingales, and supermartingales

The results of this section are valid for both discrete and continuous time processes. To simplify the presentation, we mostly restrict ourselves to the case of continuous time. We assume that the corresponding processes are continuous in  $t$  for a.e.  $\omega$  since in the main text we do not use discontinuous processes.

Let  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  be a filtered probability space satisfying the usual hypotheses. Recall that a (continuous) random process  $\{\xi_t, t \geq 0\}$  valued in a Hilbert space  $H$  is called a *martingale* with respect to  $\{\mathcal{F}_t\}$  if it is adapted to  $\mathcal{F}_t$ , each random variable  $\xi_t$  is integrable, and for  $t \geq s \geq 0$  we have  $\mathbb{E}(\xi_t | \mathcal{F}_s) = \xi_s$  with probability 1. A martingale  $\{\xi_t\}$  is said to be *square integrable* if  $\mathbb{E}\|\xi_t\|^2 < \infty$  for any  $t \geq 0$ . A *quadratic variation* of an  $H$ -valued square-integrable martingale  $\{\xi_t\}$  is a real-valued random process  $\langle \xi \rangle_t$  adapted to  $\{\mathcal{F}_t\}$  such that  $\langle \xi \rangle_0 = 0$  almost surely, almost every trajectory of  $\langle \xi \rangle_t$  is non-decreasing, and the process  $\|\xi_t\|_H^2 - \langle \xi \rangle_t$  is a real-valued martingale.

Important examples of continuous martingales are given by stochastic integrals. Namely, consider a Hilbert space  $K$  with an orthonormal basis  $\{e_j\}$ , standard independent  $\mathcal{F}_t$ -adapted Brownian motions  $\{\beta_j(t)\}$ , and an  $\mathcal{F}_t$ -adapted process  $B_t$  valued in  $\mathcal{L}(K, H)$  such that

$$\mathbb{E} \int_0^T \|B_t\|_{\text{HS}}^2 dt < \infty \quad \text{for any } T > 0.$$

Here  $\|C\|_{\text{HS}}$  stands for the Hilbert–Schmidt norm of an operator  $C \in \mathcal{L}(K, H)$  (that is,  $\|C\|_{\text{HS}}^2 = \sum_j \|Ce_j\|^2$ ). Then the stochastic integral

$$\xi_t = \int_0^t \sum_{j=1}^{\infty} B_t e_j d\beta_j(t) \quad (7.54)$$

is a continuous square-integrable martingale in  $H$ ; see Section 5.2 in [DZ96]. Moreover, its quadratic variation is given by

$$\langle \xi \rangle_t = \int_0^t \|B_s\|_{\text{HS}}^2 ds.$$

*Example 7.11.1.* The  $H$ -valued process  $\zeta(t)$  defined in (2.98) can be written in the form (7.54), where  $K = H$  and  $B$  is an operator in  $H$  such that  $Be_j = b_j e_j$  for any  $j \geq 1$ . In this case,  $\|B\|_{\text{HS}}^2 = \mathfrak{B}$  and  $\langle \zeta \rangle_t = \mathfrak{B}t$ .

A continuous random process  $\{M_t, t \geq 0\}$  in  $\mathbb{R}$  is called a *submartingale* (or *supermartingale*) with respect to  $\{\mathcal{F}_t\}$  if it is integrable and adapted to  $\mathcal{F}_t$ , and for all  $t \geq s \geq 0$  we have

$$\mathbb{E}(M_t | \mathcal{F}_s) \geq M_s \quad (\text{or } \mathbb{E}(M_t | \mathcal{F}_s) \leq M_s).$$

If  $\xi_t$  is a martingale in  $H$  and  $f : H \rightarrow \mathbb{R}$  is a continuous convex function, then  $f(\xi_t)$  is a submartingale, while  $-f(\xi_t)$  is a supermartingale; see [DZ92].

We now formulate without proof several results on martingales, submartingales and supermartingales.

**Doob–Kolmogorov inequality.** Let  $\{M_t, 0 \leq t \leq T\}$  be a non-negative submartingale. Then

$$\mathbb{P}\left\{\sup_{0 \leq t \leq T} M_t > c\right\} \leq \frac{1}{c} \mathbb{E} M_T. \quad (7.55)$$

In particular, if  $\xi_t$  is an  $H$ -valued martingale, then (7.55) is true with  $M_t = \|\xi_t\|_H^p$  for any  $p \geq 1$ .

**Doob’s moment inequality.** Let  $\{M_t, t \geq 0\}$  be a non-negative submartingale. Then, for any  $p \in (1, \infty)$ , we have

$$\mathbb{E}\left(\sup_{t \geq 0} M_t^p\right) \leq \left(\frac{p}{p-1}\right)^p \lim_{t \rightarrow \infty} \mathbb{E} M_t^p; \quad (7.56)$$

A similar inequality is true on any finite interval  $[0, T]$ , in which case the supremum should be taken over  $t \in [0, T]$  and the limit of  $\mathbb{E} M_t^p$  should be replaced by  $\mathbb{E} M_T^p$ .

**Exponential supermartingale.** Let  $\{\xi_t, t \geq 0\}$  be a martingale whose quadratic variation  $\langle \xi \rangle_t$  is a. s. finite for all  $t \geq 0$  and let  $\zeta_t := \xi_t - \frac{1}{2} \langle \xi \rangle_t$ . Then the process  $\exp(\zeta_t)$  is a supermartingale, and we have the inequality

$$\mathbb{P}\left\{\sup_{t \geq 0} \zeta_t \geq c\right\} \leq e^{-c} \quad \text{for any } c > 0. \quad (7.57)$$

A similar result holds on a finite interval  $[0, T]$ , in which case the supremum should be taken over  $t \in [0, T]$ .

**Doob's optional sampling theorem.** Let  $\{\xi_t, t \geq 0\}$  be a submartingale and let  $\sigma \leq \tau$  be two almost surely finite stopping times for the filtration  $\{\mathcal{F}_t\}$ . Then  $\mathbb{E}\xi_\sigma = \mathbb{E}\xi_\tau$ .

A proof of (7.55) and (7.56) can be found in Section III.3 of [Kry02]. The fact that the process  $\exp \zeta_t$  is a supermartingale is established in [RY99, Section 4.3], and (7.57) is the classical supermartingale inequality; see [Kry02, Theorem III.6.11] or [Mey66, Theorem VI.T1]. Finally, Doob's optional sampling theorem is proved in Section 1.3.C of [KS91].

The assertion below is a special case of a more general theorem due to Doob. Its proof can be found in [Dud02]; see Theorems 10.5.1 and 10.5.4 and Corollary 10.5.2.

**Convergence theorem.** Let  $\{M_k, k \geq 0\}$  be a real-valued martingale such that  $\sup_{k \geq 0} \mathbb{E}|M_k| < \infty$ . Then there is an integrable random variable  $M_\infty$  such that

$$M_k \rightarrow M_\infty \text{ almost surely, } \mathbb{E}|M_k - M_\infty| \rightarrow 0 \text{ as } k \rightarrow \infty.$$

## 7.12 Limit theorems for discrete-time martingales

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space, let  $\{\mathcal{F}_k, k \geq 0\}$  be a discrete-time filtration on it, and let  $\{M_k, k \geq 1\}$  be a martingale with respect to  $\mathcal{F}_k$ . The following theorem establishes a strong law of large numbers. Its proof can be found in [Fel71]; see Theorem 3 in Section VII.9.

**Theorem 7.12.1.** *Let  $\{M_k, k \geq 1\}$  be a zero-mean square-integrable martingale and let  $\{c_k\}$  be an increasing sequence going to  $+\infty$  such that*

$$\sum_{k=1}^{\infty} c_k^{-2} \mathbb{E} X_k^2 < \infty,$$

where  $X_k = M_k - M_{k-1}$  and  $M_0 = 0$ . Then

$$\mathbb{P}\{c_k^{-1} M_k \rightarrow 0 \text{ as } k \rightarrow \infty\} = 1.$$

We now turn to the law of iterated logarithm (LIL). Given a square integrable martingale  $\{M_k\}$ , define its conditional variance by the formula

$$V_k^2 = \sum_{l=1}^k \mathbb{E}(X_l^2 | \mathcal{F}_{l-1}). \quad (7.58)$$

The following result on the LIL for martingales with identically distributed increments is due to Heyde [Hey73].

**Theorem 7.12.2.** *Let a zero-mean martingale  $\{M_k\}$  be such that the martingale differences  $X_k$  are identically distributed,  $\mathbb{E} X_1^2 = \sigma^2 > 0$ , and*

$$k^{-1}V_k^2 \rightarrow \sigma^2 \quad \text{in probability as } k \rightarrow \infty. \quad (7.59)$$

Then for almost every  $\omega \in \Omega$  we have

$$\limsup_{k \rightarrow \infty} \frac{M_k}{\sqrt{2k \ln \ln k}} = \sigma, \quad \liminf_{k \rightarrow \infty} \frac{M_k}{\sqrt{2k \ln \ln k}} = -\sigma. \quad (7.60)$$

Finally, let us discuss the central limit theorem (CLT). We define

$$s_k^2 = \mathbb{E} M_k^2 = \sum_{l=1}^k \mathbb{E} X_l^2.$$

In what follows, we assume that

$$\mathbb{E} X_k^2 \rightarrow \sigma^2 \quad \text{as } k \rightarrow \infty, \quad (7.61)$$

$$s_k^{-2}V_k \rightarrow \sigma^2 \quad \text{in probability as } k \rightarrow \infty, \quad (7.62)$$

where  $\sigma \geq 0$  is a constant, and that the martingale differences  $X_k$  satisfy *Lindeberg's condition*, that is,

$$s_n^{-2} \sum_{k=1}^n \mathbb{E} (X_k^2 \mathbb{I}_{\{|X_k| > \varepsilon s_n\}}) \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad (7.63)$$

where  $\varepsilon > 0$  is arbitrary. Note that, if  $\sigma > 0$ , then  $s_k^2 \sim k\sigma^2$  as  $k \rightarrow \infty$ , and Lindeberg's condition holds if and only if, for any  $\varepsilon > 0$ ,

$$n^{-1} \sum_{k=1}^n \mathbb{E} (X_k^2 \mathbb{I}_{\{|X_k| > \varepsilon \sqrt{n}\}}) \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (7.64)$$

This condition is not very restrictive. For instance, it holds if  $\mathbb{E} X_k^4 \leq C < \infty$  for all  $k \geq 1$ . Indeed, using Schwarz's and Chebyshev's inequalities, we derive

$$\begin{aligned} \mathbb{E}(X_k^2 \mathbb{I}_{\{|X_k| > \varepsilon \sqrt{n}\}}) &\leq (\mathbb{E} X_k^4)^{1/2} \mathbb{P}_u\{|X_k| > \varepsilon \sqrt{n}\}^{1/2} \\ &\leq \varepsilon^{-2} n^{-1} \mathbb{E} X_k^4 \leq C \varepsilon^{-2} n^{-1}, \end{aligned}$$

whence it follows that (7.63) holds. The following theorem due to Brown [Bro71] establishes the CLT for martingales.

**Theorem 7.12.3.** *Let  $\{M_k\}$  be a zero-mean martingale such that the martingale differences  $X_k$  satisfy (7.61), (7.62) with some  $\sigma \geq 0$  and (7.63). Then*

$$\mathcal{D}(k^{-1/2}M_k) \rightarrow N(0, \sigma) \quad \text{as } k \rightarrow \infty.$$

That is (see Lemma 1.2.16),

$$\lim_{k \rightarrow \infty} \mathbb{P}\{k^{-1/2}M_k \leq x\} = \Phi_\sigma(x) \quad \text{for any } x \in \mathbb{R}.$$

## 7.13 Martingale approximation for Markov processes

Let  $X$  be a separable Banach space and let  $(u_t, \mathbb{P}_u)$  be a continuous-time Markov process in  $X$  defined on a measurable space  $(\Omega, \mathcal{F})$  with filtration  $\{\mathcal{F}_t, t \geq 0\}$ . We denote by  $P_t(u, \Gamma)$  the transition function for  $(u_t, \mathbb{P}_u)$  and by  $\mathfrak{P}_t$  and  $\mathfrak{P}_t^*$  the corresponding Markov semigroups (see Section 1.3.3). Given a weight function  $w$  (see Section 3.3.1) and a constant  $\gamma \in (0, 1]$ , we denote by  $C^\gamma(X, w)$  the space of continuous functions  $f : X \rightarrow \mathbb{R}$  for which

$$|f|_{w, \gamma} = \sup_{u \in X} \frac{|f(u)|}{w(\|u\|)} + \sup_{0 < \|u-v\|_X \leq 1} \frac{|f(u) - f(v)|}{\|u-v\|^\gamma (w(\|u\|_X) + w(\|v\|_X))} < \infty.$$

This is a subspace of the space of Hölder-continuous functions on  $X$ .

**Definition 7.13.1.** A Markov process  $(u_t, \mathbb{P}_u)$  is said to be *uniformly mixing* for the class  $C^\gamma(X, w)$  if it has a unique stationary distribution  $\mu \in \mathcal{P}(X)$ , and for any  $f \in C^\gamma(X, w)$ , we have<sup>7</sup>

$$|\mathfrak{P}_t f(u) - (f, \mu)| \leq \alpha(t) |f|_{w, \gamma} w_1(\|u\|_X), \quad t \geq 0, \quad (7.65)$$

where  $\alpha$  is a non-increasing integrable function on  $\mathbb{R}_+$ ,  $w_1$  is a weight function, and both of them do not depend on  $f$ .

The exercise below shows that an analogue of inequality (7.65) is true for random initial conditions.

*Exercise 7.13.2.* Let  $(u_t, \mathbb{P}_u)$  be a Markov process in  $X$  that is uniformly mixing for the class  $C^\gamma(X, w)$  and let  $\lambda \in \mathcal{P}(X)$  be a measure such that

$$(w_1, \lambda) := \int_X w_1(\|u\|_X) \lambda(du) < \infty, \quad (7.66)$$

where  $w_1$  is the function from (7.65). Show that, for any  $f \in C^\gamma(X, w)$ , we have

$$|\mathbb{E}_\lambda f(u_t) - (f, \mu)| \leq \alpha(t) |f|_{w, \gamma} (w_1, \lambda), \quad t \geq 0. \quad (7.67)$$

*Hint:* Integrate (7.65) with respect to  $\lambda(du)$ .

The following construction due to Gordin [Gor69] is very useful when studying limit theorems for stochastic processes. In the context of uniformly mixing Markov processes, it enables one to reduce the problem to the case of martingales.

Let us take any function  $f \in C^\gamma(X, w)$  with  $(f, \mu) = 0$  and a measure  $\lambda \in \mathcal{P}(X)$  and set

$$M_t^\lambda = \int_0^\infty (\mathbb{E}_\lambda(f(u_s) | \mathcal{F}_t) - \mathbb{E}_\lambda(f(u_s) | \mathcal{F}_0)) ds, \quad t \geq 0, \quad (7.68)$$

<sup>7</sup>We assume, in particular, that any function  $f \in C^\gamma(X, w)$  is integrable with respect to  $\mu$ .

where  $\mathbb{E}_\lambda$  stands for the mean value with respect to the measure  $\mathbb{P}_\lambda$  defined by

$$\mathbb{P}_\lambda(\Gamma) = \int_\Omega \mathbb{P}_u(\Gamma) \lambda(du), \quad \Gamma \in \mathcal{F}.$$

**Proposition 7.13.3.** *Let  $(u_t, \mathbb{P}_u)$  be a Markov process in  $X$  uniformly mixing for the class  $C^\gamma(X, w)$ . Assume, in addition, that*

$$\mathbb{E}_u w_1(\|u_t\|_X) \leq C(1 + w_1(\|u\|_X)) \quad \text{for any } t \geq 0, u \in X, \quad (7.69)$$

where  $C > 0$  is independent of  $t \geq 0$  and  $u \in X$ . Then, for any measure  $\lambda \in \mathcal{P}(X)$  satisfying (7.66) and any function  $f \in C^\gamma(X, w)$  with zero mean value, the process  $\{M_t^\lambda, t \geq 0\}$  is well defined and forms a zero-mean martingale with respect to the filtration  $\{\mathcal{F}_t\}$  and the probability  $\mathbb{P}_\lambda$ . Moreover, we have, with  $\mathbb{P}_\lambda$ -probability 1,

$$\int_0^t f(u_s) ds = M_t^\lambda - g(u_t) + g(u_0), \quad t \geq 0, \quad (7.70)$$

where  $g(v) = \int_0^\infty \mathfrak{P}_s f(v) ds$ .

*Proof.* To simplify notations, we shall write  $M_t$  instead of  $M_t^\lambda$ . We first show that

$$\mathbb{E}_\lambda |M_t| \leq (|f|, \mu)t + C_1(1 + (w_1, \lambda)), \quad (7.71)$$

where  $C_1 > 0$  is a constant not depending on  $\lambda$  and  $t \geq 0$ . Indeed, the Markov property and inequality (7.67) imply that, with  $\mathbb{P}_\lambda$ -probability 1, we have

$$|\mathbb{E}_\lambda(f(u_s) | \mathcal{F}_t)| \leq |\mathfrak{P}_{s-t} f(u_t)| \leq \alpha(s-t) |f|_{w, \gamma} w_1(\|u_t\|_X), \quad s \geq t. \quad (7.72)$$

Integrating (7.69) with respect to  $\lambda(du)$ , we see that

$$\mathbb{E}_\lambda w_1(\|u_t\|_X) \leq C(1 + (w_1, \lambda)) \quad \text{for any } t \geq 0. \quad (7.73)$$

Taking the mean value  $\mathbb{E}_\lambda$  of inequality (7.72) and using (7.73), for  $s \geq t$  we obtain

$$\mathbb{E}_\lambda |\mathbb{E}_\lambda(f(u_s) | \mathcal{F}_t)| \leq C\alpha(s-t) |f|_{w, \gamma} (1 + (w_1, \lambda)). \quad (7.74)$$

On the other hand, since  $u_s$  is  $\mathcal{F}_t$ -measurable for  $t \geq s$ , inequality (7.67) with  $f$  replaced by  $|f|$  implies that

$$\mathbb{E}_\lambda |\mathbb{E}_\lambda(f(u_s) | \mathcal{F}_t)| = \mathbb{E}_\lambda |f(u_s)| \leq (|f|, \mu) + \alpha(s) |f|_{w, \gamma} (w_1, \lambda), \quad s \leq t.$$

Combining this with (7.74) and using the fact that  $\int_0^\infty \alpha(t) dt < \infty$ , we obtain the required inequality (7.71).

The fact that  $M_t$  is a martingale can be checked easily by applying the conditional expectation  $\mathbb{E}_\lambda(\cdot | \mathcal{F}_r)$  to  $M_t$  with  $t \geq r$ . Furthermore, it is straightforward to check that  $\mathbb{E}_\lambda M_t = 0$  for all  $t \geq 0$ . Thus, it remains to establish (7.70).



To this end, we use the Markov property and write

$$\begin{aligned} M_t &= \int_0^t f(u_s) ds + \int_t^\infty \mathbb{E}_\lambda(f(u_s) | \mathcal{F}_t) ds - \int_0^\infty \mathbb{E}_\lambda(f(u_s) | \mathcal{F}_0) ds \\ &= \int_0^t f(u_s) ds + \int_t^\infty (\mathfrak{P}_{s-t}f)(u_t) ds - \int_0^\infty (\mathfrak{P}_s f)(u_0) ds. \end{aligned}$$

This implies the required relation (7.70).  $\square$

## 7.14 Generalised Poincaré inequality

**Proposition 7.14.1.** *For any integer  $d \geq 1$  there is a constant  $C_d > 0$  such that, for any  $p \geq 1$ , we have*

$$\int_{\mathbb{T}^d} |u|^{2p} dx \leq C_d \int_{\mathbb{T}^d} |\nabla(|u|^{2p})|^2 dx = C_d p^2 \int_{\mathbb{T}^d} |u|^{2(p-1)} |\nabla u|^2 dx, \quad (7.75)$$

where  $u \in H^1(\mathbb{T}^d)$  is an arbitrary function with zero mean value for which the right-hand side is finite.

*Proof.* We argue <sup>8</sup> by contradiction. If (7.75) does not hold, then we can construct a sequence of functions  $u_k \in H^1(\mathbb{T}^d)$  with zero mean value and some constants  $p_k \geq 1$  such that <sup>9</sup>

$$\int_{\mathbb{T}^d} |u_k|^{2p_k} dx = (2\pi)^d, \quad p_k^2 \int_{\mathbb{T}^d} |u_k|^{2(p_k-1)} |\nabla u_k|^2 dx \leq k^{-1}, \quad k \geq 1. \quad (7.76)$$

Let us set  $v_k = |u_k|^{p_k-1} u_k$ . Then inequalities (7.76) imply that

$$\int_{\mathbb{T}^d} |v_k|^2 dx = (2\pi)^d, \quad \int_{\mathbb{T}^d} |\nabla v_k|^2 dx \leq k^{-1}, \quad k \geq 1. \quad (7.77)$$

Thus,  $\{v_k\}$  is a bounded sequence in  $H^1$ . Since the embedding  $H^1 \subset L^2$  is compact (see (1.5)), there is no loss of generality in assuming that  $v_k \rightarrow v$  in  $L^2$ , where  $v$  is a function in  $H^1$ . Moreover, it follows from the second inequality in (7.77) that  $\nabla v = 0$  almost everywhere, and therefore  $v \equiv C$ . On the other hand, the first relation in (7.77) implies that  $|v|_2^2 = (2\pi)^d$ , whence it follows that  $C = 1$ . We have thus shown that

$$\int_{\mathbb{T}^d} ||u_k|^{p_k-1} u_k - 1|^2 dx \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Combining this with the elementary inequality  $|y-1| \leq 2||y|^{p-1}y-1|$ , valid for any  $y \in \mathbb{R}$  and  $p \geq 1$ , we see that  $u_k \rightarrow 1$  in  $L^2$ . This contradicts the fact that the mean value of  $u_k$  is zero.  $\square$

<sup>8</sup>We thank S. Brandle and M. Struwe for communicating us this proof.

<sup>9</sup>The case in which the  $L^{2p_k}$  norm of  $u_k$  is infinite can easily be treated by truncation. Namely, it suffices to consider the sequence  $\tilde{u}_k = \varphi(u_k/N_k)$ , where  $N_k$  is sufficiently large, and  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  is a function such that  $\varphi(r) = r$  for  $|r| \leq 1$  and  $\varphi(r) = \text{sgn}(r)$  for  $|r| > 1$ .

## 7.15 Functions in Sobolev spaces with a discrete essential range

Let  $(X, \mathcal{F})$  be a measurable spaces, let  $\mu$  be a positive measure on it, let  $Y$  be a Polish space endowed with its Borel  $\sigma$ -algebra, and let  $f : X \rightarrow Y$  be a measurable map. The *essential range* of  $f$  is defined as the set of points  $y \in Y$  such that  $\mu(f^{-1}(\mathcal{O})) > 0$  for any open set  $\mathcal{O} \subset Y$  containing  $y$ . In other words, the essential range of  $f$  is the support of the measure  $f_*(\mu)$ .

It is an obvious fact that that if  $D \subset \mathbb{R}^d$  is a connected open set, then a continuous function  $f : D \rightarrow \mathbb{R}$  with a discrete essential range must be constant. The following lemma shows that a similar result is true for functions in a Sobolev space of order  $r > 1/2$ . Its stronger versions can be found in [HKL90, BD95, BBM00]. For the reader's convenience, we shall give a proof of the lemma, following the argument in [BD95].

**Lemma 7.15.1.** *If  $u \in H^r(\mathbb{T}^2)$  with  $r > \frac{1}{2}$  and the essential range of  $u$  is a discrete subset of  $\mathbb{R}$ , then  $u = C$  almost everywhere.*

*Proof.* For  $i = 1, 2$ , denote by  $\mathbb{T}_i$  the one-dimensional torus  $\mathbb{R}/2\pi\mathbb{Z}$ ,  $\mathbb{T}_i = \{x_i\}$ . Then  $\mathbb{T}^2 = \mathbb{T}_1 \times \mathbb{T}_2$ . It follows immediately from the definition of the Sobolev norm that

$$\int_0^{2\pi} \|u(x_1, \cdot)\|_{H^r(\mathbb{T}_2)}^2 dx_1 \leq \|u\|_r^2.$$

Hence, there exists a full measure set  $A \subset \mathbb{T}_1$  such that  $u(x_1, \cdot) \in H^r(\mathbb{T}_2)$  for  $x_1 \in A$ . Since  $H^r(\mathbb{T}_2) \subset C(\mathbb{T}_2)$  and the essential range of  $u$  is a discrete subset, Fubini's theorem implies that, for almost every  $x_1 \in A$ , the essential range of  $u(x_1, \cdot)$  is a discrete subset. It follows that, for almost every  $x_1 \in \mathbb{T}_1$ , the function  $u(x_1, \cdot)$  is constant almost everywhere. Similarly, for almost every  $x_2 \in \mathbb{T}_2$ , the function  $u(\cdot, x_2)$  is constant almost everywhere.

For  $x = (x_1, x_2)$  and  $y = (y_1, y_2)$ , we now write

$$|u(x) - u(y)| \leq |u(x_1, x_2) - u(x_1, y_2)| + |u(x_1, y_2) - u(y_1, y_2)|.$$

Taking into account what has been said above and applying Fubini's theorem, we see that

$$\iint_{\mathbb{T}^2 \times \mathbb{T}^2} |u(x) - u(y)| dx dy = 0.$$

It follows that  $u(x) - u(y) = 0$  almost everywhere on  $\mathbb{T}^2 \times \mathbb{T}^2$ . Let  $B \subset \mathbb{T}^2$  be a subset of full measure such that, for  $x \in B$ , we have  $u(x) = u(y)$  almost everywhere in  $y$ . Integrating this relation in  $y$ , we conclude that  $u(x) = \int_{\mathbb{T}^2} u(y) dy$  for  $x \in B$ . This completes the proof of the lemma.  $\square$

## Chapter 8

# Solutions to some exercises

*Exercise 1.2.4.* (i) Let us assume that  $X$  is compact and prove that  $C_b(X)$  is separable. Using relation (1.19) and taking into account inequality (1.20), it is easy to approximate each continuous function  $f : X \rightarrow \mathbb{R}$  by bounded Lipschitz functions. Therefore, it suffices to construct countably many continuous functions whose finite linear combinations are dense in the unit ball of  $L_b(X)$  for the norm  $\|\cdot\|_\infty$ . Let us fix  $\varepsilon > 0$  and denote by  $\{x_1, \dots, x_n\} \subset X$  an  $\varepsilon$ -net in  $X$  and by  $\{\varphi_k\}_{k=1}^n \subset C_b(X)$  a partition of unity subordinate to the covering of  $X$  by the balls  $B_X(x_k, 2\varepsilon)$ . In this case, if  $f$  is an element of the unit ball in  $L_b(X)$ , then the norm of the function  $f - \sum_k f(x_k)\varphi_k$  does not exceed  $2\varepsilon$ .

Conversely, let us assume that  $X$  is not compact and prove that  $C_b(X)$  is not separable. Indeed, since  $X$  is not compact, there is  $\varepsilon > 0$  and a sequence  $\{x_k\} \subset X$  such that  $\text{dist}_X(x_k, x_m) \geq 2\varepsilon > 0$ . Let  $\varphi_k \in C_b(X)$  be such that  $\text{supp } \varphi_k \subset B_X(x_k, \varepsilon)$  and  $\varphi_k(x_k) = 1$ . Then the infinite linear combinations of the form  $\sum_k c_k \varphi_k$  with  $c_k = \pm 1$  form an uncountable family of functions such that the distance between any two functions is equal to 2.

(ii) Let us denote by  $\varphi_\alpha(x)$  a function equal to 0 for  $x < \alpha$  and to 1 for  $x \geq \alpha$  and define  $\psi_\alpha(x) = \int_0^x \varphi dy$ . Then the family  $\{\psi_\alpha, 0 < \alpha < 1\} \subset L_b(X)$  is uncountable and the distance in  $L_b(X)$  between any of its two elements is equal to 1.  $\square$

*Exercise 1.2.10.* Let us denote by  $\|\mu_1 - \mu_2\|'_{\text{var}}$  the right-hand side of (1.26). Since  $C_b(X) \subset L^\infty(X)$ , it follows from (1.13) that  $\|\mu_1 - \mu_2\|_{\text{var}} \leq \|\mu_1 - \mu_2\|'_{\text{var}}$ . To prove the converse inequality, we use a simple approximation argument. Namely, let us fix an arbitrary  $\varepsilon > 0$  and choose a function  $f \in L^\infty(X)$  with  $\|f\|_\infty \leq 1$  such that

$$\|\mu_1 - \mu_2\|'_{\text{var}} \leq \frac{1}{2}((f, \mu_1) - (f, \mu_2)) + \varepsilon.$$

We now choose a finitely many disjoint subsets  $\Gamma_k \in \mathcal{B}(X)$  and real constants  $c_k$  such that  $|c_k| \leq 1$  and  $|f(x) - \sum_k c_k \mathbb{1}_{\Gamma_k}(x)| \leq \varepsilon$  for all  $x \in X$ . Combining this with the

above inequality and using (1.22), we derive

$$\begin{aligned} \|\mu_1 - \mu_2\|'_{\text{var}} &\leq \frac{1}{2} \sum_k c_k (\mu_1(\Gamma_k) - \mu_2(\Gamma_k)) + 3\varepsilon \\ &\leq \frac{1}{2} \sum_{k \in \Lambda^+} (\mu_1(\Gamma_k) - \mu_2(\Gamma_k)) + \frac{1}{2} \sum_{k \in \Lambda^-} (\mu_2(\Gamma_k) - \mu_1(\Gamma_k)) + 3\varepsilon \\ &\leq \frac{1}{2} (\mu_1(\Gamma^+) - \mu_2(\Gamma^+)) + \frac{1}{2} (\mu_2(\Gamma^-) - \mu_1(\Gamma^-)) + 3\varepsilon \\ &\leq \|\mu_1 - \mu_2\|_{\text{var}} + 3\varepsilon, \end{aligned}$$

where  $\Lambda^\pm$  stands for the set of indices  $k$  such that  $\pm c_k > 0$  and  $\pm(\mu_1(\Gamma_k) - \mu_2(\Gamma_k)) > 0$ , and  $\Gamma^\pm$  is the union of  $\Gamma_k$  with  $k \in \Lambda^\pm$ . Since  $\varepsilon > 0$  was arbitrary, we arrive at the required inequality.  $\square$

*Exercise 1.2.12.* We first recall that if a sequence of functions  $\{f_n\}$  converges in the space  $L^1(X, m)$  to a limit  $f$ , then there is a subsequence  $\{f_{n_k}\}$  that converges to  $f$  almost surely. Thus, given a bounded measurable function  $f : X \rightarrow \mathbb{R}$  and a constant  $\varepsilon > 0$ , it suffices to construct  $f_\varepsilon$  such that  $\|f - f_\varepsilon\|_{L^1(X, m)} \leq \varepsilon$  and  $|f(x)| \leq \|f\|_\infty$  for all  $x \in X$ . A simple approximation argument shows that it suffices to consider the case when  $f$  is the indicator function of a Borel set  $\Gamma$ . Let  $K \subset \Gamma$  be a compact set and let  $G \supset \Gamma$  be an open set such that  $m(G \setminus K) \leq \varepsilon$ . Then the function  $f_\varepsilon(x) = \frac{\text{dist}_X(x, G)}{\text{dist}_X(x, K) + \text{dist}_X(x, G)}$  is continuous and satisfies the required inequality.  $\square$

*Exercise 1.2.13.* We confine ourselves to the proof of (i). For any  $\Gamma \in \mathcal{F}_2$ , we have

$$|f_*(\mu_1)(\Gamma) - f_*(\mu_2)(\Gamma)| = |\mu_1(f^{-1}(\Gamma)) - \mu_2(f^{-1}(\Gamma))| \leq \|\mu_1 - \mu_2\|_{\text{var}}.$$

Since  $\Gamma$  was arbitrary, we arrive at the required inequality.  $\square$

*Exercise 1.2.18. (i)* Let  $\{x_k\} \subset X$  be an arbitrary dense sequence. We claim that for any  $\mu \in \mathcal{P}(X)$  and  $\varepsilon > 0$  there is sequence  $\{c_k\} \subset \mathbb{R}$  with finitely many non-zero elements such that

$$\left\| \mu - \sum_k c_k \delta_{x_k} \right\|_L^* \leq \varepsilon, \quad (8.1)$$

where  $\delta_y$  stands for the Dirac measure concentrated at  $y$ . This property immediately implies the required result.

Let us fix  $\mu \in \mathcal{P}(X)$  and  $\varepsilon > 0$  and choose a compact subset  $\mathcal{C} \subset X$  such that  $\mu(X \setminus \mathcal{C}) \leq \varepsilon/2$ . We next cover  $\mathcal{C}$  by finitely many disjoint subsets  $B_j$  with nonempty interior and diameters  $\leq \varepsilon/2$ . Choosing arbitrary points  $x_{k_j} \in B_j$ , for any function  $f \in L_b(X)$  with norm  $\|f\|_L \leq 1$  we obtain

$$\begin{aligned} \left| (f, \mu) - \sum_j f(x_{k_j}) \mu(B_j \cap \mathcal{C}) \right| &\leq \sum_j \left| (f \mathbb{I}_{B_j \cap \mathcal{C}}, \mu) - f(x_{k_j}) \mu(B_j \cap \mathcal{C}) \right| + |(f \mathbb{I}_{X \setminus \mathcal{C}}, \mu)| \\ &\leq \sum_j \mu(B_j \cap \mathcal{C}) \sup_{x \in B_j \cap \mathcal{C}} |f(x) - f(x_{k_j})| + \varepsilon/2 \leq \varepsilon, \end{aligned}$$

where we used the fact  $f$  is Lipschitz continuous with a constant  $\leq 1$ . Since  $f$  was arbitrary, we arrive at inequality (8.1) in which  $c_{k_j} = \mu(B_j \cap \mathcal{C})$  and  $c_k = 0$  for the other indices.

*(ii)* It is straightforward to see  $\{\delta_y, y \in \mathbb{R}\}$  is an uncountable subset  $\mathcal{P}(\mathbb{R})$  such that the distance between any two elements is equal to 1. This implies that the space  $\mathcal{P}(\mathbb{R})$  endowed with the total variation distance is not separable.  $\square$

*Exercise 1.2.19.* We need to show that

$$\sup_{\|f\|_L \leq 1} |\mathbb{E}(f(\zeta_m) - f(\zeta))| \rightarrow 0 \quad \text{as } m \rightarrow \infty. \quad (8.2)$$

Let us take any function  $f \in L_b(X)$  with  $\|f\|_L \leq 1$  and write

$$\begin{aligned} |\mathbb{E}(f(\zeta_m) - f(\zeta))| &\leq |\mathbb{E}(f(\zeta_m) - f(\zeta_m^n))| + |\mathbb{E}(f(\zeta) - f(\zeta^n))| + |\mathbb{E}(f(\zeta_m^n) - f(\zeta^n))| \\ &\leq \sup_{m \geq 1} \mathbb{E}(\text{dist}_X(\zeta_m, \zeta_m^n) + \text{dist}_X(\zeta, \zeta^n)) + \|\mathcal{D}(\zeta_m^n) - \mathcal{D}(\zeta^n)\|_L^*. \end{aligned}$$

The first term on the left-hand side can be made arbitrarily small by choosing  $n \geq 1$  while the second goes to zero, for any fixed  $n \geq 1$ , as  $m \rightarrow \infty$ . Since  $f$  was arbitrary, we arrive at the required assertion.  $\square$

*Exercise 1.2.22.* Since  $\|\mu_1 - \mu_2\|_{\text{var}} = 1$ , by Corollary 1.2.11, there is a set  $A \in \mathcal{B}(X)$  for which  $\mu_1(A) = \mu_2(X \setminus A) = 1$ . It follows that if  $\xi_1$  and  $\xi_2$  are independent, then

$$\begin{aligned} \mathbb{P}\{\xi_1 \neq \xi_2\} &= \iint_{\{u_1 \neq u_2\}} \mu_1(du_1)\mu_2(du_2) = \int_A \mu_2(X \setminus \{u_1\})\mu(du_1) \\ &\geq \int_A \mu_2(X \setminus A)\mu(du_1) = 1. \end{aligned}$$

Furthermore, it is straightforward to see that the random variables  $\xi_1$  and  $\xi_2$  conditioned on the event  $N = \{\xi_1 \neq \xi_2\}$  are independent.

Conversely, if  $(\xi_1, \xi_2)$  is a maximal coupling for  $(\mu_1, \mu_2)$ , then  $\mathbb{P}(N) = 1$ , and for any  $\Gamma_1, \Gamma_2 \in \mathcal{B}(X)$  we have

$$\begin{aligned} \mathbb{P}\{\xi_1 \in \Gamma_1, \xi_2 \in \Gamma_2\} &= \mathbb{P}\{\xi_1 \in \Gamma_1, \xi_2 \in \Gamma_2 \mid N\} \\ &= \mathbb{P}\{\xi_1 \in \Gamma_1 \mid N\} \mathbb{P}\{\xi_2 \in \Gamma_2 \mid N\} \\ &= \mathbb{P}\{\xi_1 \in \Gamma_1\} \mathbb{P}\{\xi_2 \in \Gamma_2\}, \end{aligned}$$

whence we conclude that  $\xi_1$  and  $\xi_2$  are independent.  $\square$

*Exercise 1.2.23.* Let us set

$$a = \mathbb{P}(A_1 \mid E) = \mathbb{P}(A_2 \mid E), \quad b_1 = \mathbb{P}(A_1 \mid N), \quad b_2 = \mathbb{P}(A_2 \mid N),$$

where  $E = \{\xi_1 = \xi_2\}$  and  $A_i = \{\xi_i \in \Gamma\}$ . Noting that  $\mathbb{P}(E) + \mathbb{P}(N) = 1$ , we can write

$$\begin{aligned} \mathbb{P}\{\xi_1 \in \Gamma, \xi_2 \in \Gamma\} &= \mathbb{P}(A_1 A_2) = \mathbb{P}(A_1 A_2 \mid E) \mathbb{P}(E) + \mathbb{P}(A_1 A_2 \mid N) \mathbb{P}(N) \\ &= \mathbb{P}(A_1 \mid E) \mathbb{P}(E) + \mathbb{P}(A_1 \mid N) \mathbb{P}(A_2 \mid N) \mathbb{P}(N) = ax + b_1 b_2 (1 - x), \\ \mathbb{P}\{\xi_1 \in \Gamma\} \mathbb{P}\{\xi_2 \in \Gamma\} &= \mathbb{P}(A_1) \mathbb{P}(A_2) = (ax + b_1(1 - x))(ax + b_2(1 - x)), \end{aligned}$$

where  $x = \mathbb{P}(E)$ . We wish to show that the function

$$f(x) := ax + b_1 b_2 (1 - x) - (ax + b_1(1 - x))(ax + b_2(1 - x))$$

is non-negative for  $x \in [0, 1]$ . Note that  $f$  is a quadratic function such that  $f(0) = 0$  and  $f(1) = a - a^2 \geq 0$ . Therefore, the required property will be established if we show that the coefficient in front of the linear term is non-negative. Thus, it suffices to prove that

$$a + b_1 b_2 - a(b_1 + b_2) \geq 0 \quad \text{for } a, b_1, b_2 \in [0, 1].$$

This inequality can easily be checked.  $\square$

*Exercise 1.2.27.* Since convergence in the dual-Lipschitz metric is equivalent to the weak convergence of measures (see Theorem 1.2.15), the corresponding topology on the space  $\mathcal{P}(X)$  is generated by the open sets of the form  $\{\mu \in \mathcal{P}(X) : (f, \mu) < a\}$ , where  $f \in C_b(X)$  and  $a \in \mathbb{R}$ . It follows that a function  $z \mapsto \mu(z, \cdot)$  from  $Z$  to  $\mathcal{P}(X)$  is measurable if and only if, for any  $f \in C_b(X)$ , so is the function  $z \mapsto (f, \mu(z, \cdot))$  acting from  $Z$  to  $\mathbb{R}$ .

Now let  $\mu(z, \cdot)$  be a random probability measure and let  $f \in C_b(X)$ . Approximating  $f$  by finite linear combinations of indicator functions, we see that  $(f, \mu(z, \cdot))$  is a pointwise limit of measurable functions of the form  $\sum_k c_k \mu(z, \Gamma_k)$ . This observation immediately implies that  $(f, \mu(z, \cdot))$  is measurable. Conversely, if  $(f, \mu(z, \cdot))$  is measurable for any  $f \in C_b(X)$  and  $F \subset X$  is a closed subset, then approximating  $\mathbb{I}_F$  pointwise by continuous functions (see the proof of Corollary 7.1.3), we see that  $\mu(z, F) = (\mathbb{I}_F, \mu(z, \cdot))$  is measurable. The measurability of  $\mu(z, \Gamma)$  for an arbitrary  $\Gamma \in \mathcal{B}(X)$  follows by the monotone class technique.  $\square$

*Exercise 1.2.29.* The fact that  $\hat{\mu}_i(z, \cdot)$  and  $\mu(z, \cdot)$  are random probability measures follows from their definition and the measurability of  $\delta(z)$ . To prove that  $\delta$  is measurable, recall that (see Proposition 1.2.7)

$$\delta(z) = \|\mu_1(z, \cdot) - \mu_2(z, \cdot)\|_{\text{var}} = \frac{1}{2} \int_X |\rho_1(z, u) - \rho_2(z, u)| m_z(du),$$

where  $m_z = \frac{1}{2}(\mu_1(z, \cdot) + \mu_2(z, \cdot))$ , and  $\rho_i(z, u)$  is the density of  $\mu_i$  with respect to  $m_z$ . By a parameter version of the Radon–Nikodym theorem, the functions  $\rho_i$  can be assumed to be measurable in  $(z, u)$ . Thus, the required property will be established if we show that  $(f(z, \cdot), m_z)$  is measurable for any measurable function  $f : Z \times X \rightarrow \mathbb{R}$  such that  $f(z, \cdot) \in L^1(X, m_z)$  for any  $z \in Z$ . A standard approximation argument shows that it suffices to consider the case when  $f = \mathbb{I}_\Gamma$ , where  $\Gamma \in \mathcal{B}(Z \times X)$ . The claim is obvious for sets of the form  $\Gamma = \Gamma_1 \times \Gamma_2$  with  $\Gamma_1 \in \mathcal{B}(Z)$  and  $\Gamma_2 \in \mathcal{B}(X)$ . The general case follows by the monotone class technique.  $\square$

*Exercise 1.2.30.* Let us define measures  $\tilde{\mu}_i \in \mathcal{P}(X \times Y)$  by the relation

$$\tilde{\mu}_i(A \times B) = \mu_i(A \cap f^{-1}(B)) \quad \text{for } A \in \mathcal{B}(X), B \in \mathcal{B}(Y).$$

Then  $(P_Y)_* \tilde{\mu}_i = f_*(\mu_i)$ , where  $P_Y : X \times Y \rightarrow Y$  denotes the natural projection, and by the disintegration theorem, there are random probability measures  $\mu_i(y, dx)$  on  $X$  such that  $\tilde{\mu}_i(dx, dy) = f_*(\mu_i)(dy) \mu_i(y, dx)$ ,  $i = 1, 2$ . Now let  $(\eta_1, \eta_2)$  be a maximal coupling for  $(f_*(\mu_1), f_*(\mu_2))$  and let  $\zeta_i(y, \cdot)$ ,  $i = 1, 2$ , be some  $X$ -valued random variables independent of  $(\eta_1, \eta_2)$  such that  $\mathcal{D}(\zeta_i(y, \cdot)) = \mu_i(y, \cdot)$  for any  $y \in Y$ . Then the random variables  $\xi_i(\omega) = \zeta_i(\eta_i(\omega), \omega)$ ,  $i = 1, 2$ , satisfy the required property. This proves assertion (i). To establish (ii), it suffices to use in the above argument some parameter versions of the corresponding results.  $\square$

*Exercise 1.3.1.* In view of (1.48), for any  $f \in L^\infty(X)$  and any  $B \in \mathcal{F}_s$ , we have

$$\mathbb{E}_u(f(u_{t+s}) \mathbb{I}_B) = \mathbb{E}_u((\mathfrak{P}_t f)(u_s) \mathbb{I}_B).$$

Integrating this relation with respect to  $\mathbb{P}_\lambda(du)$ , we obtain

$$\mathbb{E}_\lambda(f(u_{t+s}) \mathbb{I}_B) = \mathbb{E}_\lambda((\mathfrak{P}_t f)(u_s) \mathbb{I}_B).$$

Since  $B \in \mathcal{F}_s$  was arbitrary, we arrive at (1.48).

When proving (1.49), we assume for simplicity that  $m = 2$ . It suffices to consider the case when  $f(u_1, u_2) = \mathbb{I}_{\Gamma_1}(u_1)\mathbb{I}_{\Gamma_2}(u_2)$ . Using (1.48), we write

$$\begin{aligned} \mathbb{E}_\lambda\{f(u_{t_1+s}, u_{t_2+s}) \mid \mathcal{F}_s\} &= \mathbb{E}_\lambda(\mathbb{I}_{\Gamma_1}(u_{t_1+s}) \mathbb{E}_\lambda\{\mathbb{I}_{\Gamma_2}(u_{t_2+s}) \mid \mathcal{F}_{t_1+s}\} \mid \mathcal{F}_s) \\ &= \mathbb{E}_\lambda(\mathbb{I}_{\Gamma_1}(u_{t_1+s}) P_{t_2-t_1}(u_{t_1+s}, \Gamma_2) \mid \mathcal{F}_s) \\ &= \mathfrak{P}_{t_1}(\mathbb{I}_{\Gamma_1} P_{t_2-t_1}(\cdot, \Gamma_2))(u_s). \end{aligned} \quad (8.3)$$

On the other hand, using again the Markov property, we derive

$$\begin{aligned} \mathbb{E}_v f(u_{t_1}, u_{t_2}) &= \mathbb{E}_v(\mathbb{I}_{\Gamma_1}(u_{t_1})\mathbb{I}_{\Gamma_2}(u_{t_2})) = \mathbb{E}_v(\mathbb{I}_{\Gamma_1}(u_{t_1})\mathbb{E}_v\{\mathbb{I}_{\Gamma_2}(u_{t_2}) \mid \mathcal{F}_{t_1}\}) \\ &= \mathbb{E}_v(\mathbb{I}_{\Gamma_1}(u_{t_1})P_{t_2-t_1}(u_{t_1}, \Gamma_2)) = \mathfrak{P}_{t_1}(\mathbb{I}_{\Gamma_1} P_{t_2-t_1}(\cdot, \Gamma_2))(v). \end{aligned}$$

Taking  $v = u_s$  in this relation and comparing it with (8.3), we obtain the required result.  $\square$

*Exercise 1.3.3.* It is clear that  $\mathfrak{P}_0$  is the identity operator in  $C_b(X)$ , and we shall prove the semigroup property. A simple approximation argument shows that it suffices to prove the relation

$$\mathfrak{P}_t(\mathfrak{P}_s \mathbb{I}_\Gamma) = \mathfrak{P}_{t+s} \mathbb{I}_\Gamma \quad \text{for any } \Gamma \in \mathcal{B}(X).$$

However, since  $\mathfrak{P}_r \mathbb{I}_\Gamma(u) = P_r(u, \Gamma)$ , the above equality coincides with the Kolmogorov–Chapman relation. The proof of the claim concerning  $\mathfrak{P}_t^*$  is similar.

We now prove the duality relation. It suffices to consider the case  $f = \mathbb{I}_\Gamma$ . In this situation, we have

$$(\mathfrak{P}_t \mathbb{I}_\Gamma, \mu) = \int_X P_t(u, \Gamma) \mu(du) = \mathfrak{P}_t^* \mu(\Gamma) = (\mathbb{I}_\Gamma, \mathfrak{P}_t^* \mu).$$

$\square$

*Exercise 1.3.4.* By definition, the  $\mathbb{P}_v$  law of  $u_t$  coincides with  $P_t(v, \cdot)$ . It follows that

$$\mathbb{P}_\lambda\{u_t \in \Gamma\} = \int_X \mathbb{P}_v\{u_t \in \Gamma\} \lambda(dv) = \int_X P_t(v, \Gamma) \lambda(dv) = \mathfrak{P}_t^* \lambda(\Gamma),$$

where  $\Gamma \in \mathcal{B}(X)$  is an arbitrary set. To prove (1.51), we write

$$\mathbb{E}_\lambda f(u_t) = \int_\Omega f(u_t) \mathbb{P}_\lambda(d\omega) = \int_X f(v) (\mathfrak{P}_t^* \lambda)(dv) = \int_X \mathfrak{P}_t f(v) \lambda(dv),$$

where we used the duality relation of Exercise 1.3.3.  $\square$

*Exercise 1.3.9.* We need to prove that  $\{\tau \leq t\} \cap \{\tau \leq s\}$  belongs to  $\mathcal{F}_t$  for any  $t, s \in \mathcal{T}_+$ . This fact follows immediately from the definition of a stopping time.  $\square$

*Exercise 1.3.12.* We confine our case to the case of a discrete time. It suffices to establish relation (1.59) for indicator functions  $f = \mathbb{I}_B$  with a Borel set  $B$  of the form

$$B = \{v = (v_j, j \in \mathbb{Z}_+) : v_j \in B_j \text{ for } j = 0, \dots, n\},$$

where  $B_j \in \mathcal{B}(X)$ . The case of an arbitrary Borel set  $B \subset X^{\mathbb{Z}_+}$  will follow by the monotone class technique, and an approximation argument will yield relation (1.59) for functions  $f$  satisfying the hypotheses of the exercise.

Thus, we wish to prove that

$$\mathbb{E}_u \left( \mathbb{I}_{\{\tau < \infty\}} \prod_{j=0}^m \mathbb{I}_{B_j}(u_{\tau+j}) \mid \mathcal{F}_\tau \right) = \mathbb{I}_{\{\tau < \infty\}} \left( \mathbb{E}_v \prod_{j=0}^m \mathbb{I}_{B_j}(u_{\tau+j}) \right) \Big|_{v=u_\tau}$$

with  $\mathbb{P}_u$ -probability 1. To this end, it suffices to show that

$$\mathbb{E}_u \left( \mathbb{I}_{\Gamma \cap \{\tau < \infty\}} \prod_{j=0}^m \mathbb{I}_{B_j}(u_{\tau+j}) \right) = \mathbb{E}_u \left( \mathbb{I}_{\Gamma \cap \{\tau < \infty\}} \mathbb{E}_{u_\tau} \prod_{j=0}^m \mathbb{I}_{B_j}(u_j) \right), \quad (8.4)$$

where  $\Gamma \in \mathcal{F}_\tau$  is an arbitrary subset. Using (1.49) and (1.56), we write

$$\begin{aligned} \mathbb{E}_u \left( \mathbb{I}_{\Gamma \cap \{\tau < \infty\}} \prod_{j=0}^m \mathbb{I}_{B_j}(u_{\tau+j}) \right) &= \sum_{n=0}^{\infty} \mathbb{E}_u \left( \mathbb{I}_{\Gamma \cap \{\tau=n\}} \prod_{j=0}^m \mathbb{I}_{B_j}(u_{n+j}) \right) \\ &= \sum_{n=0}^{\infty} \mathbb{E}_u \left( \mathbb{I}_{\Gamma \cap \{\tau=n\}} \mathbb{E}_u \left\{ \prod_{j=0}^m \mathbb{I}_{B_j}(u_{n+j}) \mid \mathcal{F}_n \right\} \right) \\ &= \sum_{n=0}^{\infty} \mathbb{E}_u \left( \mathbb{I}_{\Gamma \cap \{\tau=n\}} \mathbb{E}_{u_n} \prod_{j=0}^m \mathbb{I}_{B_j}(u_j) \right) \\ &= \mathbb{E}_u \left( \mathbb{I}_{\Gamma \cap \{\tau < \infty\}} \mathbb{E}_{u_\tau} \prod_{j=0}^m \mathbb{I}_{B_j}(u_j) \right). \end{aligned}$$

This completes the proof of (1.59) in the case of the discrete time.  $\square$

*Exercise 1.3.16.* Since  $\theta_s(\omega_j, j \in \mathbb{Z}) = (\omega_{j+s}, j \in \mathbb{Z})$ , the definition of  $\Phi$  implies that  $\varphi_t^{\theta_s \omega} u$  depends only on  $\omega_{s+1}, \dots, \omega_{s+t}$ . It follows that  $\mathcal{F}_{[p,q]}$  is contained in the  $\sigma$ -algebra generated by the cylindrical sets depending on  $\omega_{p+1}, \dots, \omega_q$ . On the other hand, for any integer  $k$  satisfying the inequality  $p+1 \leq k \leq q$  the  $\sigma$ -algebra generated by the random variables  $\varphi_1^{\theta_{k-1} \omega} u = S(u) + \omega_k$ ,  $u \in H$ , coincides with the cylindrical sets defined by  $\omega_k$ . What has been said implies that  $\mathcal{F}_{[p,q]}$  coincides with the  $\sigma$ -algebra of the sets of the form (1.65).  $\square$

*Exercise 1.3.17.* To be precise, we assume that  $p$  and  $q$  are finite. Since the family of those  $\Gamma \in \mathcal{F}[p, q]$  for which  $\theta_r^{-1}(\Gamma) \in \mathcal{F}_{[p+r, q+r]}$  is a  $\sigma$ -algebra, it suffices to prove this inclusion for any subset of the form

$$\Gamma = \{\omega \in \Omega : \varphi_t^{\theta_s \omega} u \in A\}, \quad A \in \mathcal{B}(X),$$

where  $p, q, t, s$ , and  $u$  satisfy conditions (1.64). To this end, note that, by the group property for  $\theta$ , we have

$$\theta_r^{-1}(\Gamma) = \{\omega \in \Omega : \varphi_t^{\theta_s(\theta_r \omega)} u \in A\} = \{\omega \in \Omega : \varphi_t^{\theta_{s+r} \omega} u \in A\},$$

whence we conclude that  $\theta_r^{-1}(\Gamma) \in \mathcal{F}_{[p+r, q+r]}$ .  $\square$

*Exercise 1.3.19.* It was proved in Exercise 1.3.16 that  $\mathcal{F}^-$  and  $\mathcal{F}^+$  coincide with the  $\sigma$ -algebras generated by the cylindrical sets depending on  $\omega_j$  with  $j \leq 0$  and  $j \geq 1$ , respectively. The required assertion follows now from the definition of the probability  $\mathbb{P}$ .  $\square$



*Exercise 1.3.20.* Let us note that, for any  $\Gamma^+ \in \mathcal{F}^+$  and  $\Gamma^- \in \mathcal{F}^-$ , we have

$$\begin{aligned} \mathbb{P}(\theta_t^{-1}(\Gamma^+) \cap \theta_t^{-1}(\Gamma^-)) &= \mathbb{P}(\theta_t^{-1}(\Gamma^+ \Gamma^-)) = \mathbb{P}(\Gamma^+ \Gamma^-) \\ &= \mathbb{P}(\Gamma^+) \mathbb{P}(\Gamma^-) = \mathbb{P}(\theta_t^{-1}(\Gamma^+)) \mathbb{P}(\theta_t^{-1}(\Gamma^-)). \end{aligned}$$

Combining this relation with Exercise 1.3.17, we arrive at the required result.  $\square$

*Exercise 1.3.23.* We need to show that if  $f \in C_b(X)$ , then  $\mathfrak{P}_t f(u)$  is continuous in  $u$ . Recall that

$$\mathfrak{P}_t f(u) = \int_{\Omega} f(\varphi_t^\omega u) \mathbb{P}(d\omega).$$

Since  $f$  is bounded and  $\varphi_t^\omega u$  is continuous in  $u$ , Lebesgue's theorem on dominated convergence implies the required result.  $\square$

*Exercise 1.3.24.* Recall that  $P_k(u, \cdot)$  is the law of the trajectory issued from  $u$ . Since the mapping  $\varphi_t$  is deterministic, we see that  $P_k(u, \cdot)$  is the Dirac measure concentrated at  $\psi^k(u)$ , where  $\psi^k$  is the  $k^{\text{th}}$  iteration of  $\psi$ . It follows that

$$\mathfrak{P}_k f(u) = f(\psi^k(u)), \quad \mathfrak{P}_k^* \mu = \mu(\psi_k^{-1}(\Gamma)).$$

$\square$

*Exercise 1.3.25.* Using the group and cocycle properties, we write

$$\Theta_t(\Theta_s(\omega, u)) = \Theta_t(\omega_s \omega, \varphi_s^\omega u) = (\theta_t(\theta_s \omega), \varphi_t^{\theta_s \omega}(\varphi_s^\omega u)) = (\theta_{t+s} \omega, \varphi_{t+s}^\omega u) = \Theta_{t+s}(\omega, u).$$

The fact that  $\Theta_0$  is the identity operator is obvious.  $\square$

*Exercise 1.3.29.* By Corollary 1.3.22, the law of  $\varphi_t u$  coincides with  $\mathfrak{P}_t^* \mu$ . Since  $\mu$  is a stationary measure, we see that  $\mathcal{D}(\varphi_t u) = \mu$  for all  $t \in \mathcal{T}_+$ .  $\square$

*Exercise 2.1.1.* Any (real-valued) function  $u \in H_\sigma^m$  can be written as a Fourier series

$$u(x) = \sum_{k \in \mathbb{Z}^2} u_k e^{ikx}, \quad (8.5)$$

where  $u_k \in \mathbb{C}^2$ ,  $\langle u_k, k \rangle = 0$ , and  $\bar{u}_k = u_{-k}$ . It is straightforward to see that the functions defined by the truncated series  $\sum_{|k| \leq N} u_k e^{ikx}$  belong to  $\mathcal{V}$  and approximate  $u$  in the  $H^m$  norm.  $\square$

*Exercise 2.1.14.* Let us note that if  $u$  is a solution of the Navier–Stokes system (2.19), then  $\dot{u} = -\nu Lu - B(u) + \Pi f$  belongs to  $L^2(0, T; V^*)$ . Therefore, using (2.10) and (2.11), we obtain

$$\frac{1}{2} (|u(t)|_2^2 - |u_0|_2^2) = \int_0^t \langle -\nu Lu - B(u) + \Pi f, u \rangle ds = \int_0^t (-\nu |\nabla u(s)|_2^2 + \langle f, u \rangle) ds.$$

This coincides with the required relation (2.39).  $\square$

*Exercise 2.1.17.* The fact that  $\psi$  is unique up to an additive constant is obvious. Let us fix a function  $u \in H$  and consider its Fourier series (8.5). It is easy to see that the coefficients  $u_k$  can be written as  $u_k = c_k k^\perp$ , where  $k^\perp = (-k_2, k_1)$ ,  $c_{-k} = -\bar{c}_k \in \mathbb{C}$ , and

$$\|u\|_m^2 = \sum_{k \in \mathbb{Z}_0^2} (1 + |k|^2)^m |k|^2 |c_k|^2.$$

It follows that the function

$$\psi(x) = -i \sum_{k \in \mathbb{Z}_0^2} c_k e^{ikx}$$

is real-valued, belongs to  $H^1(\mathbb{T}^2)$ , and satisfies the required properties.  $\square$

*Exercise 2.1.23.* Combining (2.39), (2.51), and the inequality  $2|\langle f, u \rangle| \leq \|f\|_1^2 + \|u\|_1^2$ , we derive

$$\begin{aligned} |u(0)|_2^2 &= |u(t)|_2^2 + 2\nu \int_0^t |\nabla u|_2^2 ds - 2 \int_0^t \langle f, u \rangle ds \\ &\leq e^{-\alpha_1 t} |u(0)|_2^2 + \int_0^t (C_1 \|u\|_1^2 + 2 \|f\|_{-1}^2) ds. \end{aligned}$$

This estimate readily implies the required inequality.  $\square$

*Exercise 2.1.26.* Let  $u \in \mathcal{H}$  be a solution of the Navier–Stokes system (2.19). Differentiating the function  $\varphi(t) = t |L^{1/4} u(t)|_2^2$  in time and using some standard estimates, we get

$$\begin{aligned} \varphi'(t) &= |L^{1/4} u|_2^2 + 2t \langle L^{1/2} u, \dot{u} \rangle = |L^{1/4} u|_2^2 - 2\nu t |L^{3/4} u|_2^2 + 2t \langle L^{1/2} u, \Pi f - B(u) \rangle \\ &\leq |L^{1/4} u|_2^2 - \frac{3\nu t}{2} |L^{3/4} u|_2^2 + C_1 \nu^{-1} t |f|_2^2 + C_2 t |L^{3/4} u|_2 |L^{1/4} u|_2 \|u\|_1 \\ &\leq |L^{1/4} u|_2^2 - \nu t |L^{3/4} u|_2^2 + C_1 \nu^{-1} t |f|_2^2 + C_2 \nu^{-1} t \|u\|_1^2 |L^{1/4} u|_2^2. \end{aligned}$$

It follows that

$$\partial_t (t |L^{1/4} u(t)|_2^2) + \nu t \|u\|_{\frac{3}{2}}^2 \leq \|u\|_{\frac{1}{2}}^2 + C_3 \nu^{-1} \|u\|_1^2 (t |L^{1/4} u(t)|_2^2) + C_3 \nu^{-1} t |f|_2^2.$$

A simple argument based on Gronwall's inequality and (2.24) now implies that

$$t \|u(t)\|_{\frac{1}{2}}^2 + \int_0^t s \|u(s)\|_{\frac{3}{2}}^2 ds \leq C_4(\nu) \exp\left(C_4(\nu) \int_0^t \|u\|_1^2 ds\right) \left(|u_0|_2^2 + \int_0^t |f(s)|_2^2 ds\right). \quad (8.6)$$

Now let  $u_1, u_2 \in \mathcal{H}$  be two solutions of (2.19). Their difference  $u = u_1 - u_2$  satisfies Eq. (2.25). Multiplying it by  $2tL^{1/2}u$  in  $H$  and carrying out some transformations, we derive

$$\begin{aligned} \partial_t (t |L^{1/4} u(t)|_2^2) + \nu t \|u\|_{\frac{3}{2}}^2 &\leq \|u\|_{\frac{1}{2}}^2 + C_1 \nu^{-1} t |f_1 - f_2|_2^2 \\ &\quad + C_5 t \|u\|_{\frac{3}{2}}^{4/3} |u|_2^{2/3} (|u_1|_2 \|u_1\|_{\frac{3}{2}}^2 + |u_2|_2 \|u_2\|_{\frac{3}{2}}^2)^{1/3} \end{aligned}$$

whence it follows that

$$\begin{aligned} t \|u(t)\|_{\frac{1}{2}}^2 &\leq \int_0^t (\|u\|_{\frac{1}{2}}^2 + C_1 \nu^{-1} s |f_1 - f_2|_2^2) ds \\ &\quad + C_6(\nu) \int_0^t s |u|_2^2 (|u_1|_2 \|u_1\|_{\frac{3}{2}}^2 + |u_2|_2 \|u_2\|_{\frac{3}{2}}^2) ds. \end{aligned}$$

The required inequality can now be derived by the same argument as in the proof of Proposition 2.1.25.  $\square$

*Exercise 2.1.27.* To prove the existence of a solution, we assume without loss of generality that  $g(0) = 0$  and seek a solution in the form  $u = g + v$ . Then  $v$  must be a solution of the problem

$$\dot{v} + \nu Lv + B(v + g) = h, \quad v(0) = u_0. \quad (8.7)$$

The existence of a solution for this problem can be established using the methods similar to those in Section 2.1.5.

To prove the Lipschitz continuity, we assume again that  $g_1(0) = g_2(0) = 0$ . Representing the solutions in the form  $u_i = g_i + v_i$ , we see that  $v_i$  must be a solution of problem (8.7) with  $h \equiv 0$  and  $u_0 = u_i(0)$ . It follows that the difference  $v = v_1 - v_2$  satisfies the equations

$$\dot{v} + \nu Lv + B(v + g, u_1) + B(u_2, v + g) = 0, \quad v(0) = u_1(0) - u_2(0),$$

where  $g = g_1 - g_2$ . Taking the scalar product in  $H$  of the first equation with  $2v$  and repeating essentially the scheme used in the proof of assertion (i) of Proposition 2.1.25, we arrive at the required result.  $\square$

*Exercise 2.2.3.* Let us set  $k_\varepsilon = [t_0/\varepsilon] + 1$ . It follows from inequality (2.74) and the independence of  $\eta_k$  that

$$\begin{aligned} \mathbb{P}\left\{\sup_{0 \leq t \leq t_0} |\zeta_\varepsilon(t) - \tilde{\zeta}_\varepsilon(t)|_2 \leq \delta\right\} &\geq \mathbb{P}\{\max\{|\eta_k|_2, 1 \leq k \leq k_\varepsilon\} \leq \frac{\delta}{\sqrt{\varepsilon}}\} \\ &\geq \prod_{k=1}^{k_\varepsilon} \left(1 - \mathbb{P}\{|\eta_k|_2 > \frac{\delta}{\sqrt{\varepsilon}}\}\right) \geq (1 - C_1(\delta)\varepsilon^{q/2})^{k_\varepsilon} \\ &\geq \exp(-C_2(\delta)k_\varepsilon\varepsilon^{q/2}). \end{aligned}$$

Since  $k_\varepsilon \sim 1/\varepsilon$  and  $q > 2$ , we see that the right-hand side of this inequality goes to 1 as  $\varepsilon \rightarrow 0$ .  $\square$

*Exercise 2.2.4.* Theorem 8.2 of [Bil99] implies that

$$\mathcal{D}(\tilde{\zeta}_\varepsilon^n) \rightarrow \mathcal{D}(\zeta^n) \quad \text{in } \mathcal{P}(C(0, t_0; H)) \text{ as } \varepsilon \rightarrow 0^+,$$

where the process  $\zeta^n$  is defined in the proof of Theorem 2.2.2 and  $\tilde{\zeta}_\varepsilon^n$  is a similar finite-dimensional approximation for  $\tilde{\zeta}_\varepsilon$ . Combining this convergence with inequality (2.73), its analogue for  $\tilde{\zeta}_\varepsilon$ , and Exercise 1.2.19, we arrive at the required result.  $\square$

*Exercise 2.4.4.* Let us denote by  $\{\hat{e}_j\}$  an orthonormal basis in  $H$  composed of the eigenfunction of the Stokes operator  $L$ . For any integer  $n \geq 1$ , we denote by  $P_n$  the orthogonal projection in  $H$  to the vector space spanned by  $\hat{e}_j$ ,  $j = 1, \dots, n$ . Let us consider the  $n$ -dimensional linear stochastic equation

$$\dot{z}_n + \nu L_n z_n = \partial_t P_n \zeta^n(t), \quad (8.8)$$

where  $L_n = P_n L$  and  $\zeta^n(t) = \sum_{j=1}^n b_j(t) \beta_j e_j$ . In view of the well-known results on existence and uniqueness of solutions for stochastic ODE's (e.g., see Chapter 5 in [Øks03]), Eq. (8.8) has a unique solution  $z_n$  satisfying the initial condition  $z_n(0) = 0$ .

It suffices to prove that there is a subsequence of  $\{z_n\}$  that converges to  $z(t)$  almost surely in the space  $\mathcal{X}_T = C(0, T; H) \cap L^2(0, T; V)$  for any  $T > 0$ .

The proof of this fact is divided into two steps: we first derive some a priori estimates for  $z_n$  with the help of Itô's formula and Doob's moment inequality and then, using a similar technique, we show that the difference  $z_n - z_m$  converges to zero in an appropriate space and that the limiting function satisfies (2.103).

*Step 1: A priori estimates.* We claim that

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} |z_n(t)|_2^2 + \int_0^T \|z_n(t)\|_1^2 dt \right) \leq C(\mathfrak{B}, T) \nu^{-1}, \quad (8.9)$$

where the constant  $C(\mathfrak{B}, T)$  does not depend on  $n$ . Indeed, by the finite-dimensional Itô formula (see Chapter 4 in [Øks03]), we have

$$|z_n(t)|_2^2 = \int_0^t \left( -2\nu |\nabla z_n(s)|_2^2 + \sum_{j=1}^n b_j^2 |P_n e_j|_2^2 \right) ds + 2 \int_0^t \langle z_n, d\zeta^n \rangle. \quad (8.10)$$

Taking the mean value and using Friedrichs' inequality, we obtain

$$\sup_{0 \leq t \leq T} \mathbb{E} |z_n(t)|_2^2 + 2\nu \mathbb{E} \int_0^T \|z_n\|_1^2 dt \leq C_1(\mathfrak{B}, T) \quad \text{for all } n \geq 1. \quad (8.11)$$

Furthermore, in view of Doob's moment inequality, we have

$$\begin{aligned} \mathbb{E} \sup_{0 \leq t \leq T} \left| \int_0^t \langle z_n, d\zeta^n \rangle \right|^2 &\leq \mathbb{E} \int_0^T \sum_{j=1}^n b_j^2 (z_n(t), e_j)^2 ds \\ &\leq C_2(\mathfrak{B}) \mathbb{E} \int_0^T |z_n(t)|_2^2 dt \leq C_3(\mathfrak{B}, T) \nu^{-1}, \end{aligned}$$

where we used (8.11) in the last inequality. Combining this with (8.10) and (8.11), we arrive at (8.9).

*Step 2: Convergence.* We now prove that a subsequence of  $\{z_n\}$  converges in  $\mathcal{X}_T$  almost surely and that the limit function coincides with  $z$ . Let us fix some integers  $m < n$ . The difference  $z_{mn} = z_n - z_m$  vanishes at  $t = 0$  and satisfies the equation

$$\dot{z}_{mn} + \nu P_n L z_{mn} = \partial_t \zeta_{mn}(t), \quad (8.12)$$

where we used the relation  $(P_n - P_m)Lz_m = 0$  and set

$$\zeta_{mn}(t) = \sum_{j=1}^m b_j \beta_j(t) (P_n - P_m) e_j + \sum_{j=m+1}^n b_j \beta_j(t) P_n e_j.$$

Applying Itô's formula, we derive

$$|z_{mn}(t)|_2^2 = \int_0^t \left( -2\nu |\nabla z_{mn}(s)|_2^2 + F_{mn} \right) ds + 2 \int_0^t \langle z_{mn}, d\zeta_{mn} \rangle, \quad (8.13)$$

where

$$F_{mn} = \sum_{j=1}^m b_j^2 |(P_n - P_m) e_j|_2^2 + \sum_{j=m+1}^n b_j^2 |P_n e_j|_2^2.$$

Taking the mean value in (8.13), we see that (2.107) remains true. The proof can now be completed by exactly the same argument as in the case of Proposition 2.4.2.  $\square$

*Exercise 2.4.17.* Let us set  $\Lambda_N = \{j \in \Lambda : j \leq N\}$ . Then

$$\tilde{\zeta}_N(t, x) = \sum_{j \in \Lambda_N} \omega_t^{(j)} e_j(x), \quad t \in \mathbb{R},$$

is an  $H$ -valued process with continuous trajectories whose range is contained in the vector span of  $e_j$ ,  $j \in \Lambda_N$ . Let us denote by  $\mathcal{F}_t$ ,  $t \in \mathbb{R}$ , the natural filtration on  $\Omega$  augmented<sup>1</sup> with respect to  $(\mathcal{F}, \mathbb{P})$ . The process  $\tilde{\zeta}_N$  is the sum of finitely many Brownian motions and therefore is a zero-mean martingale with respect to  $\mathcal{F}_t$ . By Doob's moment inequality, for any  $T > 0$  and  $M < N$  we have

$$\mathbb{E} \left( \sup_{|t| \leq T} |\tilde{\zeta}_N(t) - \tilde{\zeta}_M(t)|_2^2 \right) \leq 4 \mathbb{E} |\tilde{\zeta}_N(T) - \tilde{\zeta}_M(T)|_2^2 = \sum_{j \in \Lambda_N \setminus \Lambda_M} b_j^2.$$

Thus, the series (2.149) converges in  $L^2(\Omega, C(-T, T; H))$  for any  $T > 0$  and defines an  $H$ -valued process with continuous trajectories. Furthermore, since the laws of finite-dimensional approximation for  $\zeta$  and  $\tilde{\zeta}|_{\mathbb{R}_+}$  coincide, this convergence implies that the same is true for the limiting processes.  $\square$

*Exercise 2.4.18.* The fact that  $\theta$  is a group is obvious. Now note that the restriction of  $\theta_t$  to the space spanned by  $e_j$  with  $j \in \Lambda_N$  is the usual shift on the canonical probability space of a finite-dimensional Brownian motion and therefore is measure preserving. The monotone class technique now implies the required result.  $\square$

*Exercise 2.5.2.* Let us set  $\bar{\lambda}_t = t^{-1} \int_0^t P_s(u, \cdot) ds$ . We claim that the family  $\{\bar{\lambda}_t, t \geq 0\}$  is tight, so that the hypotheses of Theorem 2.5.1 are satisfied for the initial measure  $\delta_u$ . Let us fix  $\varepsilon > 0$  and choose  $m \geq 1$  so large that  $P_t(u, K_m^c) \leq \varepsilon/2$  for  $t \geq t_m$ . On the other hand, the time continuity of trajectories implies that the mapping  $t \mapsto P_t(u, \cdot)$  is continuous from  $\mathbb{R}_+$  to the space  $\mathcal{P}(X)$  endowed with the weak topology. Therefore the image of the interval  $[0, t_m]$  is compact in  $\mathcal{P}(X)$ , whence it follows that there is a compact set  $K_0 \subset X$  such that  $P_t(u, K_0^c) \leq \varepsilon/2$  for  $0 \leq t \leq t_m$ . Combining these two inequalities, we see that  $\bar{\lambda}_t(K_0 \cup K_m) \geq 1 - \varepsilon$  for all  $t \in \mathcal{T}_+$ .  $\square$

*Exercise 2.5.9.* Let us establish the uniqueness of a stationary measure. To this end, it suffices to prove that if the right-hand side is sufficiently small, then the stochastic Navier–Stokes system satisfies (1.75). To this end, we repeat the argument used in the proof of inequality (2.54). Namely, if  $u_0, v_0 \in H$  are two initial functions and  $u, v$  are the corresponding solutions, then their difference  $w = u - v$  satisfies the equation (cf. (2.166))

$$\dot{w} + \nu Lw + B(w, u) + B(v, w) = 0.$$

Taking the scalar product of this equation with  $2w$  in  $H$  and carrying out some transformations, we derive

$$\partial_t |w|_2^2 + (\alpha_1 \nu - C_1 \nu^{-1} \|u\|_1^2) |w|_2^2 \leq 0.$$

Applying Gronwall's inequality, we obtain

$$|w(t)|_2^2 \leq |u_0 - v_0|_2^2 \exp(-\alpha_1 \nu t / 2) \Psi(t, u_0), \quad t \geq 0,$$

<sup>1</sup>In other words, we take the  $\sigma$ -algebra generated by the random variables  $\omega_s^{(j)}$  with  $s \leq t$  and  $j \in \Lambda$  and add zero-measure subsets of  $\mathcal{F}$ .

where we set

$$\Psi(t, u_0) = \exp\left(-\frac{1}{2}\alpha_1\nu t + C_1\nu^{-1} \int_0^t \|u\|_1^2 ds\right).$$

It follows from (2.130) that if  $|h|_2 + \mathfrak{B}$  is sufficiently small, then

$$\mathbb{P}\left\{\sup_{t \geq 0} \Psi(t, u_0) \geq \rho\right\} \leq \exp(-\delta \ln \rho + \delta |u_0|_2^2),$$

where  $\delta > 0$  is a small constant. We thus obtain inequality (1.75) with  $\alpha(t) = e^{-\alpha_1\nu t/2}$  and  $\psi_{u_0, v_0} = \sup_{t \geq 0} \Psi(t, u_0)$ . Further analysis based on inequalities (1.76) and (2.157) enables one to prove exponential convergence to the unique stationary measure in the dual-Lipschitz distance.  $\square$

*Exercise 3.5.7.* The existence and uniqueness of a solution whose almost every trajectory belongs to  $C(\mathbb{R} \setminus \mathcal{T}_\omega; H) \cap L_{\text{loc}}^2(\mathbb{R}_+; V)$  is obvious, since we can treat the equation pathwise (cf. the case of a random kick force considered in Section 2.3). Furthermore, the existence of left- and right-hand limits at  $t_k$  follows from the properties of solution for deterministic Navier–Stokes system. To prove the Markov property, let us denote by  $\mathcal{F}_t$  the filtration generated by the random variables  $\zeta(s)$  with  $0 \leq s \leq t$ . It is well known that the process  $\zeta(t+s) - \zeta(t)$ ,  $s \geq 0$ , is independent of  $\mathcal{F}_t$ ; e.g. see Problem 3.2 of Chapter 1 in [KS91] for the case of a Poisson process. The Markov property can now be established by an argument similar to that used in the proof of Proposition 1.3.21.  $\square$

*Exercise 3.5.9.* Let us denote by  $\mathcal{F}_t$  the filtration associated with the problem in question, that is, the  $\sigma$ -algebra generated by the process

$$\int_0^t \eta(s) ds = \sum_{k=1}^{\infty} \eta_k(x) H(t - \tau_k),$$

where  $H(t) = 1$  for  $t \geq 0$  and  $H(t) = 0$  for  $t < 0$ . Since the indicator function  $\mathbb{I}_{[0, \tau_1]}(t)$  is  $\mathcal{F}_t$ -measurable, for any initial measure  $\lambda \in \mathcal{P}(H)$  and any  $s > 0$  we have

$$\begin{aligned} \mathbb{E}_\lambda \int_0^{\tau_1} f(u_{t+s}) dt &= \mathbb{E}_\lambda \int_0^{\infty} \mathbb{I}_{[0, \tau_1]}(t) f(u_{t+s}) dt \\ &= \mathbb{E}_\lambda \int_0^{\infty} \mathbb{I}_{[0, \tau_1]}(t) \mathbb{E}_\lambda \{f(u_{t+s}) \mid \mathcal{F}_t\} dt \\ &= \mathbb{E}_\lambda \int_0^{\infty} \mathbb{I}_{[0, \tau_1]}(t) (\mathfrak{P}_s f)(u_t) dt = \mathbb{E}_\lambda \int_0^{\tau_1} (\mathfrak{P}_s f)(u_t) dt. \end{aligned}$$

Therefore, setting  $c = (\mathbb{E}\tau_1)^{-1}$  and using the definition of  $\mu$ , for any  $f \in C_b(H)$  we obtain

$$\begin{aligned} (\mathfrak{P}_s f, \mu) &= c \mathbb{E}_{\bar{\mu}} \int_0^{\tau_1} (\mathfrak{P}_s f)(u_t) dt = c \mathbb{E}_{\bar{\mu}} \int_s^{s+\tau_1} f(u_t) dt \\ &= (f, \mu) + c \mathbb{E}_{\bar{\mu}} \int_{\tau_1}^{s+\tau_1} f(u_t) dt - c \mathbb{E}_{\bar{\mu}} \int_0^s f(u_t) dt. \end{aligned}$$

It remains to note that, due to the strong Markov property, the second and third terms on the right-hand side coincide and, hence,  $\mathfrak{P}_s^* \mu = \mu$  for any  $s > 0$ .  $\square$

# Bibliography

- [Ada75] R. A. Adams, *Sobolev Spaces*, Academic Press, New York–London, 1975.
- [AKSS07] A. Agrachev, S. Kuksin, A. Sarychev, and A. Shirikyan, *On finite-dimensional projections of distributions for solutions of randomly forced 2D Navier–Stokes equations*, Ann. Inst. H. Poincaré Probab. Statist. **43** (2007), no. 4, 399–415.
- [Ale43] A. D. Alexandroff, *Additive set-functions in abstract spaces*, Rec. Math. [Mat. Sbornik] N.S. **13(55)** (1943), 169–238.
- [App04] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge University Press, Cambridge, 2004.
- [Arn98] L. Arnold, *Random Dynamical Systems*, Springer-Verlag, Berlin, 1998.
- [AS05] A. A. Agrachev and A. V. Sarychev, *Navier–Stokes equations: controllability by means of low modes forcing*, J. Math. Fluid Mech. **7** (2005), no. 1, 108–152.
- [AS06] ———, *Controllability of 2D Euler and Navier–Stokes equations by degenerate forcing*, Comm. Math. Phys. **265** (2006), no. 3, 673–697.
- [Bat69] G. K. Batchelor, *Computation of the energy spectrum in homogeneous two-dimensional turbulence*, Phys. Fluids Suppl. **11** (1969), 233–239.
- [Bat82] ———, *The Theory of Homogeneous Turbulence*, Cambridge University Press, Cambridge, 1982.
- [Bax89] P. Baxendale, *Lyapunov exponents and relative entropy for a stochastic flow of diffeomorphisms*, Probab. Theory Related Fields **81** (1989), no. 4, 521–554.
- [BBM00] J. Bourgain, H. Brezis, and P. Mironescu, *Lifting in Sobolev spaces*, J. Anal. Math. **80** (2000), 37–86.

- [BD95] F. Bethuel and F. Demengel, *Extensions for Sobolev mappings between manifolds*, Calc. Var. Partial Differential Equations **3** (1995), no. 4, 475–491.
- [BD07] V. Barbu and G. Da Prato, *Existence and ergodicity for the two-dimensional stochastic magneto-hydrodynamics equations*, Appl. Math. Optim. **56** (2007), no. 2, 145–168.
- [Ber00] D. Bernard, *Influence of friction on the direct cascade of the 2d forced turbulence*, Europhys. Lett. **50** (2000), 333–339.
- [BF09] Z. Brzeźniak and B. Ferrario, *2D Navier–Stokes equation in Besov spaces of negative order*, Nonlinear Anal. **70** (2009), no. 11, 3902–3916.
- [Bil99] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1999.
- [BIN79] O. V. Besov, V. P. Il'in, and S. M. Nikol'skiĭ, *Integral Representations of Functions and Imbedding Theorems. Vol. I, II*, V. H. Winston & Sons, Washington, D.C., 1979.
- [Bis81] J.-M. Bismut, *Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions*, Z. Wahrsch. Verw. Gebiete **56** (1981), no. 4, 469–505.
- [BK07] J. Bec and K. Khanin, *Burgers turbulence*, Phys. Rep. **447** (2007), no. 1-2, 1–66.
- [BKL00] J. Bricmont, A. Kupiainen, and R. Lefevere, *Probabilistic estimates for the two-dimensional stochastic Navier–Stokes equations*, J. Statist. Phys. **100** (2000), no. 3-4, 743–756.
- [BKL01] ———, *Ergodicity of the 2D Navier–Stokes equations with random forcing*, Comm. Math. Phys. **224** (2001), no. 1, 65–81.
- [BKL02] ———, *Exponential mixing of the 2D stochastic Navier–Stokes dynamics*, Comm. Math. Phys. **230** (2002), no. 1, 87–132.
- [BMS11] W. Bołt, A. A. Majewski, and T. Szarek, *An invariance principle for the law of the iterated logarithm for some Markov chains*, Preprint (2011).
- [Bog98] V. I. Bogachev, *Gaussian Measures*, Mathematical Surveys and Monographs, vol. 62, American Mathematical Society, Providence, RI, 1998.
- [Bog07] ———, *Measure Theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.



- [Bol82] E. Bolthausen, *The berry–esseén theorem for strongly mixing Harris recurrent Markov chains*, Z. Wahrsch. Verw. Gebiete **60** (1982), no. 3, 283–289.
- [Bor12] A. Boritchev, *Sharp estimates for turbulence in white-forced generalised Burgers equation*, Preprint (2012), arXiv:1201.5567.
- [Bow75] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Mathematics, Vol. 470, Springer-Verlag, Berlin, 1975.
- [BPR06] P. Berti, L. Pratelli, and P. Rigo, *Almost sure weak convergence of random probability measures*, Stochastics **78** (2006), no. 2, 91–97.
- [Bri02] J. Bricmont, *Ergodicity and mixing for stochastic partial differential equations*, Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002) (Beijing), Higher Ed. Press, 2002, pp. 567–585.
- [Bro71] B. M. Brown, *Martingale central limit theorems*, Ann. Math. Statist. **42** (1971), 59–66.
- [BT73] A. Bensoussan and R. Temam, *Équations stochastiques du type Navier–Stokes*, J. Functional Analysis **13** (1973), 195–222.
- [BV92] A. V. Babin and M. I. Vishik, *Attractors of Evolution Equations*, North-Holland Publishing, Amsterdam, 1992.
- [BV11] F. Bouchet and A. Venaille, *Statistical mechanics of two-dimensional and geophysical flows*, Physics Reports (2011), to appear.
- [CC92] M. Capiński and N. Cutland, *A simple proof of existence of weak and statistical solutions of Navier–Stokes equations*, Proc. Roy. Soc. London Ser. A **436** (1992), no. 1896, 1–11.
- [CDF97] H. Crauel, A. Debussche, and F. Flandoli, *Random attractors*, J. Dynam. Differential Equations **9** (1997), no. 2, 307–341.
- [CF88] P. Constantin and C. Foiaş, *Navier–Stokes equations*, University of Chicago Press, Chicago, IL, 1988.
- [CF94] H. Crauel and F. Flandoli, *Attractors for random dynamical systems*, Probab. Theory Related Fields **100** (1994), no. 3, 365–393.
- [CFNT89] P. Constantin, C. Foiaş, B. Nicolaenko, and R. Temam, *Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations*, Springer-Verlag, New York, 1989.

- [CG94] M. Capiński and D. Gatarek, *Stochastic equations in Hilbert space with application to Navier-Stokes equations in any dimension*, J. Funct. Anal. **126** (1994), no. 1, 26–35.
- [Che98] J.-Y. Chemin, *Perfect Incompressible Fluids*, Oxford Lecture Series in Mathematics and its Applications, vol. 14, The Clarendon Press Oxford University Press, New York, 1998.
- [CK08a] I. Chueshov and S. Kuksin, *Random kick-forced 3D Navier–Stokes equations in a thin domain*, Arch. Rational Mech. Anal. **188** (2008), 117–153.
- [CK08b] ———, *Stochastic 3D Navier–Stokes equations in a thin domain and its  $\alpha$ -approximation*, Phys. D **237** (2008), no. 10–12, 1352–1367.
- [Cra91] H. Crauel, *Markov measures for random dynamical systems*, Stochastics Stochastics Rep. **37** (1991), no. 3, 153–173.
- [Cra01] ———, *Random point attractors versus random set attractors*, J. London Math. Soc. (2) **63** (2001), no. 2, 413–427.
- [Cra02] ———, *Random Probability Measures on Polish Spaces*, Taylor & Francis, London, 2002.
- [CV77] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions*, Springer, Berlin, 1977.
- [CV02] V. V. Chepyzhov and M. I. Vishik, *Attractors for Equations of Mathematical Physics*, AMS Coll. Publ., vol. 49, AMS, Providence, 2002.
- [DD02] G. Da Prato and A. Debussche, *Two-dimensional Navier–Stokes equations driven by a space-time white noise*, J. Funct. Anal. **196** (2002), no. 1, 180–210.
- [DD03] ———, *Ergodicity for the 3D stochastic Navier–Stokes equations*, J. Math. Pures Appl. **82** (2003), 877–947.
- [DDL<sup>+</sup>07] J. Dedecker, P. Doukhan, G. Lang, J. R. León, S. Louhichi, and C. Priour, *Weak Dependence: with Examples and Applications*, Springer, New York, 2007.
- [DDT05] G. Da Prato, A. Debussche, and L. Tubaro, *Coupling for some partial differential equations driven by white noise*, Stochastic Process. Appl. **115** (2005), no. 8, 1384–1407.
- [Deb97] A. Debussche, *On the finite dimensionality of random attractors*, Stochastic Anal. Appl. **15** (1997), no. 4, 473–491.
- [Deb98] ———, *Hausdorff dimension of a random invariant set*, J. Math. Pures Appl. (9) **77** (1998), no. 10, 967–988.

- [Den04] D. Denisov, Personal communication (2004).
- [DG95] C. R. Doering and J. D. Gibbon, *Applied Analysis of the Navier–Stokes Equations*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 1995.
- [DM78] C. Dellacherie and P.-A. Meyer, *Probabilities and Potential*, North-Holland, Amsterdam, 1978.
- [DM10] J. Dedecker and F. Merlevède, *On the almost sure invariance principle for stationary sequences of Hilbert-valued random variables*, Dependence in Probability, Analysis and Number Theory, Kendrick Press, Heber City, UT, 2010, pp. 157–175.
- [DO05] A. Debussche and C. Odasso, *Ergodicity for a weakly damped stochastic non-linear Schrödinger equation*, J. Evol. Equ. **5** (2005), no. 3, 317–356.
- [DO06] ———, *Markov solutions for the 3D stochastic Navier–Stokes equations with state dependent noise*, J. Evol. Equ. **6** (2006), no. 2, 305–324.
- [Dob68] R. L. Dobrušin, *Description of a random field by means of conditional probabilities and conditions for its regularity*, Teor. Veroyatnost. i Primenen **13** (1968), 201–229.
- [Dob74] ———, *Conditions for the absence of phase transitions in one-dimensional classical systems*, Math. USSR-Sb. **22** (1974), no. 1, 28–48.
- [Doe38] W. Doeblin, *Exposé de la théorie des chaînes simples constantes de Markov à un nombre fini d'états*, Rev. Math. Union Interbalkan **2** (1938), 77–105.
- [Doe40] ———, *Éléments d'une théorie générale des chaînes simples constantes de Markoff*, Ann. Sci. École Norm. Sup. (3) **57** (1940), 61–111.
- [Doo48] J. L. Doob, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. **63** (1948), 393–421.
- [DRRW06] G. Da Prato, M. Röckner, B. L. Rozovskii, and F.-Y. Wang, *Strong solutions of stochastic generalized porous media equations: existence, uniqueness, and ergodicity*, Comm. Partial Differential Equations **31** (2006), no. 1-3, 277–291.
- [Dud02] R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2002.
- [DZ92] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, 1992.

- [DZ96] ———, *Ergodicity for Infinite Dimensional Systems*, Cambridge University Press, Cambridge, 1996.
- [EH01] J.-P. Eckmann and M. Hairer, *Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise*, Comm. Math. Phys. **219** (2001), no. 3, 523–565.
- [Elw92] K. D. Elworthy, *Stochastic flows on Riemannian manifolds*, Diffusion processes and related problems in analysis, Vol. II (Charlotte, NC, 1990), Birkhäuser, Boston, MA, 1992, pp. 37–72.
- [EMS01] W. E, J. C. Mattingly, and Ya. Sinai, *Gibbsian dynamics and ergodicity for the stochastically forced Navier–Stokes equation*, Comm. Math. Phys. **224** (2001), no. 1, 83–106.
- [ES00] W. E and Ya. G. Sinaĭ, *New results in mathematical and statistical hydrodynamics*, Russian Math. Surveys **55** (2000), no. 4(334), 635–666.
- [Fel71] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol. II.*, John Wiley & Sons, New York, 1971.
- [Fer97] B. Ferrario, *Ergodic results for stochastic Navier–Stokes equation*, Stochastics Stochastics Rep. **60** (1997), no. 3-4, 271–288.
- [Fer99] ———, *Stochastic Navier–Stokes equations: analysis of the noise to have a unique invariant measure*, Ann. Mat. Pura Appl. (4) **177** (1999), 331–347.
- [Fer03] ———, *Uniqueness result for the 2D Navier–Stokes equation with additive noise*, Stoch. Stoch. Rep. **75** (2003), no. 6, 435–442.
- [FG95] F. Flandoli and D. Gałarek, *Martingale and stationary solutions for stochastic Navier–Stokes equations*, Probab. Theory Related Fields **102** (1995), no. 3, 367–391.
- [Fla94] F. Flandoli, *Dissipativity and invariant measures for stochastic Navier–Stokes equations*, NoDEA Nonlinear Differential Equations Appl. **1** (1994), no. 4, 403–423.
- [Fla97] ———, *Irreducibility of the 3-D stochastic Navier–Stokes equation*, J. Funct. Anal. **149** (1997), no. 1, 160–177.
- [FM95] F. Flandoli and B. Maslowski, *Ergodicity of the 2D Navier–Stokes equation under random perturbations*, Comm. Math. Phys. **172** (1995), no. 1, 119–141.
- [FMRT01] C. Foias, O. Manley, R. Rosa, and R. Temam, *Navier–Stokes Equations and Turbulence*, Encyclopedia of Mathematics and its Applications, vol. 83, Cambridge University Press, Cambridge, 2001.

- [FP67] C. Foias and G. Prodi, *Sur le comportement global des solutions non-stationnaires des équations de Navier–Stokes en dimension 2*, Rend. Sem. Mat. Univ. Padova **39** (1967), 1–34.
- [FR08] F. Flandoli and M. Romito, *Markov selections for the 3D stochastic Navier–Stokes equations*, Probab. Theory Related Fields **172** (2008), 407–458.
- [Fri95] U. Frisch, *Turbulence. The Legacy of A. N. Kolmogorov*, Cambridge University Press, Cambridge, 1995.
- [FT89] C. Foias and R. Temam, *Gevrey class regularity for the solutions of the Navier–Stokes equations*, J. Funct. Anal. **87** (1989), no. 2, 359–369.
- [Gal02] G. Gallavotti, *Foundations of Fluid Dynamics*, Springer-Verlag, Berlin, 2002.
- [GK03] I. Gyöngy and N. Krylov, *On the splitting-up method and stochastic partial differential equations*, Ann. Probab. **31** (2003), no. 2, 564–591.
- [GM05] B. Goldys and B. Maslowski, *Exponential ergodicity for stochastic Burgers and 2D Navier–Stokes equations*, J. Funct. Anal. **226** (2005), no. 1, 230–255.
- [Gor69] M. I. Gordin, *The central limit theorem for stationary processes*, Dokl. Akad. Nauk SSSR **188** (1969), 739–741.
- [GS80] Ī. Ī. Gihman and A. V. Skorohod, *The Theory of Stochastic Processes. I*, Springer-Verlag, Berlin, 1980.
- [Hai02a] M. Hairer, *Exponential mixing for a stochastic partial differential equation driven by degenerate noise*, Nonlinearity **15** (2002), no. 2, 271–279.
- [Hai02b] ———, *Exponential mixing properties of stochastic PDEs through asymptotic coupling*, Probab. Theory Related Fields **124** (2002), no. 3, 345–380.
- [Hai05] ———, *Coupling stochastic PDEs*, XIVth International Congress on Mathematical Physics, World Sci. Publ., Hackensack, NJ, 2005, pp. 281–289.
- [Har55] T. E. Harris, *On chains of infinite order*, Pacific J. Math. **5** (1955), 707–724.
- [Har64] P. Hartman, *Ordinary Differential Equations*, John Wiley & Sons, New York, 1964.

- [Has80] R. Z. Has'minskii, *Stochastic Stability of Differential Equations*, Sijthoff & Noordhoff, Alphen aan den Rijn, 1980.
- [Hen81] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Springer-Verlag, Berlin, 1981.
- [Hey73] C. C. Heyde, *An iterated logarithm result for martingales and its application in estimation theory for autoregressive processes*, J. Appl. Probability **10** (1973), 146–157.
- [HH80] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Application*, Academic Press, New York, 1980.
- [HKL90] R. Hardt, D. Kinderlehrer, and F. H. Lin, *The variety of configurations of static liquid crystals*, Variational methods (Paris, 1988), Progr. Nonlinear Differential Equations Appl., vol. 4, Birkhäuser Boston, Boston, MA, 1990, pp. 115–131.
- [HM06] M. Hairer and J. C. Mattingly, *Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing*, Ann. of Math. (2) **164** (2006), no. 3, 993–1032.
- [HM08] ———, *Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations*, Ann. Probab. **36** (2008), no. 6, 2050–2091.
- [HM11] ———, *A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs*, Electron. J. Probab. **16** (2011), no. 23, 658–738.
- [Hör97] L. Hörmander, *Lectures on Nonlinear Hyperbolic Differential Equations*, Springer-Verlag, Berlin, 1997.
- [JS87] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin, 1987.
- [Jud63] V. I. Judovič, *Non-stationary flows of an ideal incompressible fluid*, Ž. Vychisl. Mat. i Mat. Fiz. **3** (1963), 1032–1066.
- [KA82] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, 1982.
- [KB37] N. Kryloff and N. Bogoliouboff, *La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire*, Ann. of Math. (2) **38** (1937), no. 1, 65–113.
- [KF75] A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis*, Dover Publications Inc., New York, 1975.

- [KH95] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, 1995.
- [Kif86] Y. Kifer, *Ergodic Theory of Random Transformations*, Birkhäuser, Boston, MA, 1986.
- [KP05] S. Kuksin and O. Penrose, *A family of balance relations for the two-dimensional Navier–Stokes equations with random forcing*, J. Statist. Phys. **118** (2005), no. 3-4, 437–449.
- [KP08] S. Kuksin and A. Piatnitski, *Khasminskii-Whitham averaging for randomly perturbed KdV equation*, J. Math. Pures Appl. (9) **89** (2008), no. 4, 400–428.
- [KPS02] S. Kuksin, A. Piatnitski, and A. Shirikyan, *A coupling approach to randomly forced nonlinear PDEs. II*, Comm. Math. Phys. **230** (2002), no. 1, 81–85.
- [Kra67] R. H. Kraichnan, *Inertial ranges in two-dimensional turbulence*, Physics of Fluids **10** (1967), 1417–1423.
- [Kry73] N. V. Krylov, *The selection of a Markov process from a Markov system of processes, and the construction of quasidiffusion processes*, Izv. Akad. Nauk SSSR Ser. Mat. **37** (1973), 691–708.
- [Kry74] ———, *Some estimates for the density of the distribution of a stochastic integral*, Izv. Akad. Nauk SSSR Ser. Mat. **38** (1974), 228–248.
- [Kry80] ———, *Controlled Diffusion Processes*, Applications of Mathematics, vol. 14, Springer-Verlag, New York, 1980.
- [Kry86] ———, *Estimates of the maximum of the solution of a parabolic equation and estimates of the distribution of a semimartingale*, Mat. Sb. (N.S.) **130(172)** (1986), no. 2, 207–221, 284.
- [Kry87] ———, *Nonlinear Elliptic and Parabolic Equations of the Second Order*, D. Reidel Publishing Co., Dordrecht, 1987.
- [Kry96] ———, *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, American Mathematical Society, Providence, RI, 1996.
- [Kry02] ———, *Introduction to the Theory of Random Processes*, Graduate Studies in Mathematics, vol. 43, American Mathematical Society, Providence, RI, 2002.
- [KS91] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1991.
- [KS00] S. Kuksin and A. Shirikyan, *Stochastic dissipative PDEs and Gibbs measures*, Comm. Math. Phys. **213** (2000), no. 2, 291–330.

- [KS01a] ———, *A coupling approach to randomly forced nonlinear PDE's. I*, Comm. Math. Phys. **221** (2001), no. 2, 351–366.
- [KS01b] ———, *Ergodicity for the randomly forced 2D Navier–Stokes equations*, Math. Phys. Anal. Geom. **4** (2001), no. 2, 147–195.
- [KS02a] ———, *Coupling approach to white-forced nonlinear PDEs*, J. Math. Pures Appl. (9) **81** (2002), no. 6, 567–602.
- [KS02b] ———, *On dissipative systems perturbed by bounded random kick-forces*, Ergodic Theory Dynam. Systems **22** (2002), no. 5, 1487–1495.
- [KS03] ———, *Some limiting properties of randomly forced two-dimensional Navier–Stokes equations*, Proc. Roy. Soc. Edinburgh Sect. A **133** (2003), no. 4, 875–891.
- [KS04a] ———, *On random attractors for systems of mixing type*, Funktsional. Anal. i Prilozhen. **38** (2004), no. 1, 34–46, 95.
- [KS04b] ———, *Randomly forced CGL equation: stationary measures and the inviscid limit*, J. Phys. A **37** (2004), no. 12, 3805–3822.
- [Kuk02a] S. Kuksin, *Ergodic theorems for 2D statistical hydrodynamics*, Rev. Math. Phys. **14** (2002), no. 6, 585–600.
- [Kuk02b] ———, *On exponential convergence to a stationary measure for nonlinear PDEs perturbed by random kick-forces, and the turbulence limit*, Partial differential equations, Amer. Math. Soc. Transl. Ser. 2, vol. 206, Amer. Math. Soc., Providence, RI, 2002, pp. 161–176.
- [Kuk04] ———, *The Eulerian limit for 2D statistical hydrodynamics*, J. Statist. Phys. **115** (2004), no. 1-2, 469–492.
- [Kuk06a] ———, *Randomly Forced Nonlinear PDEs and Statistical Hydrodynamics in 2 Space Dimensions*, European Mathematical Society (EMS), Zürich, 2006.
- [Kuk06b] ———, *Remarks on the balance relations for the two-dimensional Navier–Stokes equation with random forcing*, J. Statist. Phys. **122** (2006), no. 1, 101–114.
- [Kuk07] ———, *Eulerian limit for 2D Navier–Stokes equation and damped/driven KdV equation as its model*, Tr. Mat. Inst. Steklova **259** (2007), Anal. i Osob. Ch. 2, 134–142.
- [Kuk08] ———, *On distribution of energy and vorticity for solutions of 2D Navier–Stokes equation with small viscosity*, Comm. Math. Phys. **284** (2008), no. 2, 407–424.



- [Kuk10a] ———, *Damped-driven KdV and effective equations for long-time behaviour of its solutions*, *Geom. Funct. Anal.* **20** (2010), no. 6, 1431–1463.
- [Kuk10b] ———, *Dissipative perturbations of KdV*, Proceedings of the 16th International Congress on Mathematical Physics (Prague 2009) (P. Exner, ed.), World Scientific, 2010, pp. 323–327.
- [Kuk12] ———, *Weakly nonlinear stochastic CGL equations*, *Ann. Inst. H. Poincaré Probab. Statist.* (2012), to appear.
- [Kup10] A. Kupiainen, *Ergodicity of two-dimensional turbulence*, Preprint (2010), arXiv:1005.0587.
- [Lad59] O. A. Ladyženskaja, *Solution “in the large” of the nonstationary boundary value problem for the Navier–Stokes system with two space variables*, *Comm. Pure Appl. Math.* **12** (1959), 427–433.
- [Lad63] O. A. Ladyzhenskaya, *The Mathematical Theory of Viscous Incompressible Flow*, Gordon and Breach, New York, 1963.
- [Le 87] Y. Le Jan, *Équilibre statistique pour les produits de difféomorphismes aléatoires indépendants*, *Ann. Inst. H. Poincaré Probab. Statist.* **23** (1987), no. 1, 111–120.
- [Led86] F. Ledrappier, *Positivity of the exponent for stationary sequences of matrices*, *Lyapunov Exponents* (Bremen, 1984), Springer, Berlin, 1986, pp. 56–73.
- [Ler34] J. Leray, *Sur le mouvement d’un liquide visqueux emplissant l’espace*, *Acta Math.* **63** (1934), no. 1, 193–248.
- [Lio69] J.-L. Lions, *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires*, Dunod, 1969.
- [LM72] J.-L. Lions and E. Magenes, *Non-Nomogeneous Boundary Value Problems and Applications. Vol. I*, Springer-Verlag, New York, 1972.
- [LP59] J.-L. Lions and G. Prodi, *Un théorème d’existence et unicité dans les équations de Navier–Stokes en dimension 2*, *C. R. Acad. Sci. Paris* **248** (1959), 3519–3521.
- [LR06] J. A. Langa and J. C. Robinson, *Fractal dimension of a random invariant set*, *J. Math. Pures Appl.* (9) **85** (2006), no. 2, 269–294.
- [LS06] A. Lasota and T. Szarek, *Lower bound technique in the theory of a stochastic differential equation*, *J. Differential Equations* **231** (2006), no. 2, 513–533.

- [LY94] A. Lasota and J. A. Yorke, *Lower bound technique for Markov operators and iterated function systems*, Random Comput. Dynam. **2** (1994), no. 1, 41–77.
- [Mat99] J. C. Mattingly, *Ergodicity of 2D Navier–Stokes equations with random forcing and large viscosity*, Comm. Math. Phys. **206** (1999), no. 2, 273–288.
- [Mat02a] ———, *The dissipative scale of the stochastic Navier–Stokes equation: regularization and analyticity*, J. Statist. Phys. **108** (2002), no. 5-6, 1157–1179.
- [Mat02b] ———, *Exponential convergence for the stochastically forced Navier–Stokes equations and other partially dissipative dynamics*, Comm. Math. Phys. **230** (2002), no. 3, 421–462.
- [Mat03] ———, *On recent progress for the stochastic Navier–Stokes equations*, Journées “Équations aux Dérivées Partielles”, Univ. Nantes, Nantes, 2003, pp. Exp. No. XI, 52.
- [Maz85] V. G. Maz’ja, *Sobolev Spaces*, Springer-Verlag, Berlin, 1985.
- [McK69] H. P. McKean, *Stochastic Integrals*, Academic Press, New York, 1969.
- [Mey66] P.-A. Meyer, *Probability and Potentials*, Blaisdell Publishing Co. Ginn and Co., Waltham, Mass.-Toronto, Ont.-London, 1966.
- [MP06] J. C. Mattingly and É. Pardoux, *Malliavin calculus for the stochastic 2D Navier–Stokes equation*, Comm. Pure Appl. Math. **59** (2006), no. 12, 1742–1790.
- [MT93] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag London, London, 1993.
- [MY02] N. Masmoudi and L.-S. Young, *Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs*, Comm. Math. Phys. **227** (2002), no. 3, 461–481.
- [Ner08] V. Nersesyan, *Polynomial mixing for the complex Ginzburg–Landau equation perturbed by a random force at random times*, J. Evol. Equ. **8** (2008), no. 1, 1–29.
- [Nov05] D. Novikov, *Hahn decomposition and Radon–Nikodym theorem with a parameter*, arXiv:math/0501215 (2005).
- [Oda07] C. Odasso, *Exponential mixing for the 3D stochastic Navier–Stokes equations*, Comm. Math. Phys. **270** (2007), no. 1, 109–139.
- [Oda08] ———, *Exponential mixing for stochastic PDEs: the non-additive case*, Probab. Theory Related Fields **140** (2008), no. 1-2, 41–82.

- [Øks03] B. Øksendal, *Stochastic Differential Equations*, Springer-Verlag, Berlin, 2003.
- [Ons49] L. Onsager, *Statistical hydrodynamics*, *Nuovo Cimento* (9) **6** (1949), no. 2, 279–287.
- [Pit74] J. W. Pitman, *Uniform rates of convergence for Markov chain transition probabilities*, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **29** (1974), 193–227.
- [PSS89] N. I. Portenko, A. V. Skorokhod, and V. M. Shurenkov, *Markov processes*, Current problems in mathematics. Fundamental directions, Vol. 46 (Russian), Akad. Nauk SSSR Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1989, pp. 5–245.
- [Rev84] D. Revuz, *Markov Chains*, North-Holland, Amsterdam, 1984.
- [Rio00] E. Rio, *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*, Springer-Verlag, Berlin, 2000.
- [Rom08] M. Romito, *Analysis of equilibrium states of Markov solutions to the 3D Navier–Stokes equations driven by additive noise*, *J. Statist. Phys.* **131** (2008), no. 3, 415–444.
- [Roz90] B. L. Rozovskii, *Stochastic Evolution Systems. Linear Theory and Applications to Non-Linear Filtering*, Kluwer, Dordrecht, 1990.
- [RS93] G. Raugel and G. R. Sell, *Navier–Stokes equations on thin 3D domains. I. Global attractors and global regularity of solutions*, *J. Amer. Math. Soc.* **6** (1993), no. 3, 503–568.
- [Rue68] D. Ruelle, *Statistical mechanics of a one-dimensional lattice gas*, *Comm. Math. Phys.* **9** (1968), 267–278.
- [RY99] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, vol. 293, Springer-Verlag, Berlin, 1999.
- [Sei97] J. Seidler, *Ergodic behaviour of stochastic parabolic equations*, *Czechoslovak Math. J.* **47(122)** (1997), no. 2, 277–316.
- [Shi02] A. Shirikyan, *Analyticity of solutions of randomly perturbed two-dimensional Navier–Stokes equations*, *Russian Math. Surveys* **57** (2002), no. 4, 785–799.
- [Shi04] ———, *Exponential mixing for 2D Navier–Stokes equations perturbed by an unbounded noise*, *J. Math. Fluid Mech.* **6** (2004), no. 2, 169–193.
- [Shi05a] ———, *Ergodicity for a class of Markov processes and applications to randomly forced PDE’s. I*, *Russ. J. Math. Phys.* **12** (2005), no. 1, 81–96.

- [Shi05b] ———, *Some mathematical problems of statistical hydrodynamics*, XIVth International Congress on Mathematical Physics, World Sci. Publ., Hackensack, NJ, 2005, pp. 304–311.
- [Shi06a] ———, *Approximate controllability of three-dimensional Navier-Stokes equations*, *Comm. Math. Phys.* **266** (2006), no. 1, 123–151.
- [Shi06b] ———, *Ergodicity for a class of Markov processes and applications to randomly forced PDE's. II*, *Discrete Contin. Dyn. Syst. Ser. B* **6** (2006), no. 4, 911–926 (electronic).
- [Shi06c] ———, *Law of large numbers and central limit theorem for randomly forced PDE's*, *Probab. Theory Related Fields* **134** (2006), no. 2, 215–247.
- [Shi07a] ———, *Contrôlabilité exacte en projections pour les équations de Navier-Stokes tridimensionnelles*, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **24** (2007), no. 4, 521–537.
- [Shi07b] ———, *Qualitative properties of stationary measures for three-dimensional Navier–Stokes equations*, *J. Funct. Anal.* **249** (2007), 284–306.
- [Shi08] ———, *Exponential mixing for randomly forced partial differential equations: method of coupling*, *Instability in models connected with fluid flows. II*, *Int. Math. Ser. (N. Y.)*, vol. 7, Springer, New York, 2008, pp. 155–188.
- [Shi11a] ———, *Control and mixing for 2D Navier–Stokes equations with space-time localised noise*, Preprint (2011), arXiv:1110.0596.
- [Shi11b] ———, *Local times for solutions of the complex Ginzburg–Landau equation and the inviscid limit*, *J. Math. Anal. Appl.* **384** (2011), 130–137.
- [Sin91] Ya. G. Sinaï, *Two results concerning asymptotic behavior of solutions of the Burgers equation with force*, *J. Statist. Phys.* **64** (1991), no. 1-2, 1–12.
- [Soh01] H. Sohr, *The Navier–Stokes Equations*, Birkhäuser, Basel, 2001.
- [Ste70] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, N.J., 1970.
- [Ste94] L. Stettner, *Remarks on ergodic conditions for Markov processes on Polish spaces*, *Bull. Polish Acad. Sci. Math.* **42** (1994), no. 2, 103–114.
- [Str93] D. Stroock, *Probability. An Analytic Viewpoint*, Cambridge University Press, Cambridge, 1993.

- [SV79] D. Stroock and S. R. S. Varadhan, *Multidimensional Diffusion Processes*, Springer, Berlin, 1979.
- [Sza97] T. Szarek, *Markov operators acting on Polish spaces*, Ann. Polon. Math. **67** (1997), no. 3, 247–257.
- [Tay97] M. E. Taylor, *Partial Differential Equations. I–III*, Springer-Verlag, New York, 1996–97.
- [Tem68] R. Temam, *Une méthode d’approximation de la solution des équations de Navier-Stokes*, Bull. Soc. Math. France **96** (1968), 115–152.
- [Tem79] ———, *Navier–Stokes Equations*, North-Holland, Amsterdam, 1979.
- [Tem88] ———, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag, New York, 1988.
- [TZ96] R. Temam and M. Ziane, *Navier–Stokes equations in three-dimensional thin domains with various boundary conditions*, Adv. Differential Equations **1** (1996), no. 4, 499–546.
- [VF88] M. I. Vishik and A. V. Fursikov, *Mathematical Problems in Statistical Hydromechanics*, Kluwer, Dordrecht, 1988.
- [Vio75] M. Viot, *Équations aux dérivées partielles stochastiques: formulation faible*, Séminaire sur les Équations aux Dérivées Partielles (1974–1975), III, Exp. No. 1, Collège de France, Paris, 1975, p. 16.
- [Vio76] ———, *Solutions faibles d’équations aux dérivées partielles stochastiques non linéaires*, Thèses de Doctorat (1976), Paris–VI.
- [VKF79] M. I. Vishik, A. I. Komech, and A. V. Fursikov, *Some mathematical problems of statistical hydromechanics*, Uspekhi Mat. Nauk **34** (1979), no. 5(209), 135–210.
- [Yos95] K. Yosida, *Functional Analysis*, Springer-Verlag, Berlin, 1995.

## Notation and conventions

$\mathbb{Z}$ ,  $\mathbb{R}$ ,  $\mathbb{C}$  denote the sets of integer, real, and complex numbers, respectively.

$\mathbb{R}^n$  stands for the  $n$ -dimensional Euclidean space.

$\mathbb{Z}^n$  is the set of integer vectors  $(s_1, \dots, s_n)$  and  $\mathbb{Z}_0^n$  is the set of non-zero vectors in  $\mathbb{Z}^n$ .

$\mathcal{T}$  stands for  $\mathbb{Z}$  or  $\mathbb{R}$ ,  $\mathcal{T}_+ = \{t \in \mathcal{T} : t \geq 0\}$ , and  $\mathcal{T}_- = \{t \in \mathcal{T} : t \leq 0\}$ .

If  $Y \subset X$  is a subset, then  $Y^c$  stands for the complement of  $Y$  in  $X$ .

The infimum over an empty subset of  $\mathbb{R}$  is  $+\infty$ . Given two real numbers  $a$  and  $b$ , we denote by  $a \vee b$  and  $a \wedge b$  their maximum and minimum, respectively.

We often deal with random forces decomposed in an orthonormal basis  $\{e_j\}$  of the space  $L^2(Q; \mathbb{R}^2)$ . We denote by  $b_j$  the corresponding coefficients and always assume that

$$\mathfrak{B} := \sum_{j=1}^{\infty} b_j^2 < \infty.$$

When  $\{e_j\}$  is a basis formed of the eigenfunctions of the Stokes operator, we set

$$\mathfrak{B}_k = \sum_{j=1}^{\infty} \alpha_j^k b_j^2,$$

where  $\alpha_j$  stands for the eigenvalue associated with  $e_j$ . Note that  $\mathfrak{B}_0 = \mathfrak{B}$ . In case of the standard torus  $\mathbb{T}^2$ , the eigenfunctions (as well as the eigenvalues and coefficients) are indexed by  $s \in \mathbb{Z}_0^2$ , and we obtain

$$\mathfrak{B}_k = \sum_{s \in \mathbb{Z}_0^2} |s|^{2k} b_s^2.$$

## Abstract spaces and functions

For an arbitrary set  $Y$ , we denote by  $\text{Id}_Y$  the identity operator on  $Y$  (which take an element  $y \in Y$  to itself). If  $\Gamma \subset Y$ , then we write  $\mathbb{I}_\Gamma$  the indicator function of  $\Gamma$  (equal to 1 on  $\Gamma$  and to zero otherwise).

If  $\Omega$  is an arbitrary set and  $\mathcal{F}$  is a  $\sigma$ -algebra on  $\Omega$ , then  $(\Omega, \mathcal{F})$  is called a *measurable space*.

Let  $(\Omega_i, \mathcal{F}_i)$ ,  $i = 1, 2$ , be measurable spaces and  $f : \Omega_1 \rightarrow \Omega_2$  be a mapping. We say that  $f$  is *measurable* if  $f^{-1}(\Gamma) \in \mathcal{F}_1$  for any  $\Gamma \in \mathcal{F}_2$ . For any probability  $\mu$  on  $(\Omega_1, \mathcal{F}_1)$ , we denote by  $f_*(\mu)$  or  $f_*\mu$  the image of  $\mu$  under  $f$ , that is, a measure on  $(\Omega_2, \mathcal{F}_2)$  defined by the relation  $f_*(\mu)(\Gamma) = \mu(f^{-1}(\Gamma))$  for  $\Gamma \in \mathcal{F}_2$ .

$X$  denotes either a Polish space (that is, a separable complete metric space) or a separable Banach space. We write  $\text{dist}_X$  and  $\|\cdot\|_X$  for the corresponding metric and norm. All Polish spaces are endowed with the Borel  $\sigma$ -algebra  $\mathcal{B} = \mathcal{B}(X)$

and are considered as measurable spaces. We denote by  $L^\infty(X)$  the space of real-valued bounded measurable functions on  $X$  and by  $\mathcal{P}(X)$  the set of probability measures on  $(X, \mathcal{B}(X))$ .

$C_b(X)$  denotes the space of bounded continuous functions  $f : X \rightarrow \mathbb{R}$  endowed with the norm

$$\|f\|_\infty = \sup_{u \in X} |f(u)|.$$

$L_b(X)$  stands for the space of functions  $f \in C_b(X)$  such that

$$\text{Lip}(f) := \sup_{u_1, u_2 \in X} \frac{|f(u_1) - f(u_2)|}{\text{dist}_X(u_1, u_2)} < \infty.$$

This is a Banach space with respect to the norm  $\|f\|_L = \|f\|_\infty + \text{Lip}(f)$ .

$C_b^\gamma(X)$  denotes the space of functions  $f \in C_b(X)$  such that

$$|f|_\gamma := \|f\|_\infty + \sup_{0 < \text{dist}_X(u, v) \leq 1} \frac{|f(u) - f(v)|}{\text{dist}_X(u, v)^\gamma} < \infty.$$

The *canonical space* on an interval  $J \subset \mathbb{R}$  is defined as the space  $\Omega = \Omega_X$  of continuous functions  $\omega : J \rightarrow X$  endowed with the metric of uniform convergence on bounded subintervals. We consider only two cases:  $J = \mathbb{R}$  or  $J = \mathbb{R}_+$ . The *shift operator*  $\theta_s : \Omega \rightarrow \Omega$  is given by  $(\theta_s \omega)(t) = \omega(s + t)$ ,  $t \in J$ , where  $s \in \mathbb{R}$  in the first case and  $s \geq 0$  in the second. The *canonical process* on  $\Omega$  is defined by  $x_t(\omega) = \omega(t)$  for  $t \in J$ .

Given a random process  $\{\xi_t, t \geq 0\}$  valued in a Banach space, we denote by  $\tau(B) \leq \infty$  its *first hitting time* of a closed set  $B$ ; that is,  $\tau(B) = \inf\{t \geq 0 : \xi_t \in B\}$ .

$H$  stands for a separable Hilbert space with a scalar product  $(\cdot, \cdot)_H$ .

$(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space. We often assume that it is complete, that is, if  $A \in \mathcal{F}$  and  $\mathbb{P}(A) = 0$ , then any subset of  $A$  belongs to  $\mathcal{F}$ . If, in addition,  $\{\mathcal{F}_t, t \geq 0\}$  is a filtration on  $(\Omega, \mathcal{F}, \mathbb{P})$ , then unless otherwise stated, we assume that  $\mathcal{F}_t$  is right-continuous, and  $\mathcal{F}_0$  contains all negligible sets of  $\mathcal{F}$ .

$B_X(u, r)$  denotes the closed ball in  $X$  of radius  $r$  centred at  $u$ . If  $X$  is a Banach space and  $u = 0$ , then we write  $B_X(r)$ .

If  $X$  is a Polish space, then  $\mathbf{X}$  stands for the direct product  $X \times X$  endowed with the natural metric, and we write  $\mathbf{u} = (u, u')$  to denote elements of  $\mathbf{X}$ . Similarly, if  $B \subset X$  is a subset, then  $\mathbf{B}$  stands for  $B \times B$ . Finally, if  $(\Omega, \mathcal{F})$  is a measurable space, then we denote by  $(\mathbf{\Omega}, \mathbf{\mathcal{F}})$  the direct product of its two copies, that is,  $\mathbf{\Omega} = \Omega \times \Omega$  and  $\mathbf{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}$ .

## Functional spaces

Let  $Q$  be a domain in  $\mathbb{R}^d$  or a manifold, let  $J \subset \mathbb{R}$  be a closed interval, let  $1 \leq p \leq \infty$ , and let  $X$  be a Banach space. We use the following functional spaces.

$C_b^\infty(Q; \mathbb{R}^n)$  is the Fréchet space of infinitely smooth functions  $f : Q \rightarrow \mathbb{R}^n$  that are bounded together with all their derivatives.

$L^p(Q; \mathbb{R}^n)$  is the space of measurable functions  $f : Q \rightarrow \mathbb{R}^n$  such that

$$\|f\|_p := \left( \int_Q |f(x)|^p \right)^{1/p} < \infty.$$

In the case  $p = \infty$ , the above norm should be replaced by

$$\|f\|_\infty := \operatorname{ess\,sup}_{x \in Q} |f(x)|.$$

$H^s(Q; \mathbb{R}^n)$  is the Sobolev space of order  $s \in \mathbb{R}$  endowed with its standard norm  $\|\cdot\|_s$ .

$L_\sigma^2(\mathbb{T}^2; \mathbb{R}^2)$  denotes the space of functions  $u \in L^2(\mathbb{T}^2; \mathbb{R}^2)$  such that  $\operatorname{div} u = 0$  in  $\mathbb{T}^2$  and  $H$  is the space of those  $u \in L_\sigma^2(\mathbb{T}^2; \mathbb{R}^2)$  for which  $\langle u \rangle := \int_{\mathbb{T}^2} u \, dx = 0$ .

If  $Q$  is a bounded domain with smooth boundary, then  $H = L_\sigma^2(Q; \mathbb{R}^2)$  stands for the space of functions  $u \in L^2(Q; \mathbb{R}^2)$  such that  $\operatorname{div} u = 0$  in  $Q$  and  $\langle u, \mathbf{n} \rangle = 0$  on  $\partial Q$ , where  $\mathbf{n}$  is the outward unit normal to  $\partial Q$ .

In the case of a torus, we set  $V^k = H^k(\mathbb{T}^2; \mathbb{R}^2) \cap H$  and write  $V = V^1$ . In the case of a bounded domain, we define  $V = H_0^1(Q; \mathbb{R}^2) \cap L_\sigma^2(Q; \mathbb{R}^2)$ . In both cases, we denote by  $V^*$  the dual space of  $V$  with respect to the scalar product in  $L^2$ .

$L^p(J; X)$  denotes the space of Borel-measurable functions  $u : J \rightarrow X$  for which  $\|u\|_{L^p(J; X)} < \infty$ ; see page 4 for the definition of this norm.

$C(J; X)$  is the space of continuous functions on  $J$  endowed with the supremum norm  $\|u\|_{C(J; X)}$  defined on page 4.

## Measures and applications

A mapping  $F : X \rightarrow Y$  between two Banach spaces is said to be *locally Lipschitz* if for any  $R > 0$  there is  $C_R > 0$  such that

$$\|F(u_1) - F(u_2)\|_Y \leq C_R \|u_1 - u_2\|_X \quad \text{for any } u_1, u_2 \in B_X(R).$$

$\mathcal{L}(X, Y)$  denotes the space of continuous linear applications from  $X$  to  $Y$ . In the case  $X = Y$ , we shall write  $\mathcal{L}(X)$ .

Given a measure  $\mu$  on a Banach space  $X$ , we denote by  $\mathfrak{m}_k(\mu)$  its  $k^{\text{th}}$  moment:

$$\mathfrak{m}_k(\mu) = \int_X \|u\|_X^k \mu(du).$$

For any measurable space  $(\Omega, \mathcal{F})$  and a point  $a \in \Omega$ , we denote by  $\delta_a$  the *Dirac measure* concentrated at  $a$ , that is,  $\delta_a(\Gamma) = 1$  if  $\Gamma \in \mathcal{F}$  contains  $a$  and  $\delta_a(\Gamma) = 0$  otherwise.



## Subject Index

- $\pi$ -system, 259
- $\sigma$ -algebra, 5
  
- absorbing set, 32
- adapted random process, 6, 60
- admissible martingale solution, 251
- Alexandrov theorem, 265
- augmented filtration, 268
  
- balance relations, 207
- Bogolyubov–Krylov argument, 83
- Borel  $\sigma$ -algebra, 6
- Burgers equation, 105
  
- Calderón–Zygmund theorem, 209
- canonical process, 132, 315
- canonical space, 315
- central limit theorem, viii, 173
- characteristic scales, 55
- closed random set, 175
- co-area form of balance relations, 208
- compact random set, 175
- complete  $\sigma$ -algebra, 6
- complete probability space, 6
- completion of a  $\sigma$ -algebra, 6
- compound Poisson process, 57
- conditional law, 18
- countably generated  $\sigma$ -algebra, 261
- coupling, 16
- coupling operators, 113
  
- Da Prato–Debussche–Odasso theorem, 247
- Dirac measure, 316
- disintegration, 30
- distribution function, 15, 19, 262
- distribution of a random variable, 6
- Dobrushin lemma, 16
- Doob convergence theorem, 181, 281
- Doob moment inequality, 57, 59, 69, 280, 296, 297
- Doob optional sampling theorem, 73, 75, 129, 281
  
- Doob–Kolmogorov inequality, 80, 269, 271, 280
- dual-Lipschitz distance, 8
  
- energy, 55
- energy balance, 45, 73
- energy inequality, 252
- energy spectrum of turbulence, 233
- enstrophy, 55
- exponential mixing, 98, 101, 106, 140
- exponential supermartingale, 280
- extension of a Markov process, 103
  
- Fatou lemma, 7
- Feller property, 24
- filtered probability space, 6
- first hitting time, 315
- Flandoli–Romito theorem, 251
- Foias–Prodi inequality, 52
- function of bounded variation, 119
  
- Gelfand triple, 46, 266, 267
  
- image of a measure, 5
- interpolation, 3
- Itô formula, 268
  
- Kantorovich density, 20
- Kantorovich distance, 9
- Kantorovich functional, 20
- Kantorovich–Rubinstein theorem, 21
- Khas’minskii relation, 146
- Kolmogorov inequality, 81
- Kolmogorov–Chapman relation, 22, 29
- Krylov estimate, 222, 225, 276
  
- Ladyzhenskaya inequality, 4
- law of a random variable, 6
- law of iterated logarithm, 170
- law of large numbers, 165
- Lebesgue spaces, 1
- Lebesgue theorem on dominated convergence, 7

- Ledrappier–Le Jan–Crauel theorem, 179  
 Leray decomposition, 36  
 Leray projection, 37  
 Lindeberg condition, 282  
 Lipschitz domain, 1  
 local time, 274, 276  
 Lyapunov–Schmidt reduction, 149
- Markov chain, 22  
 Markov invariant measures, 179  
 Markov process, 21  
 Markov property, 22, 28  
 Markov RDS, 28  
 Markov selection, 252  
 Markov semigroups, 22  
 martingale, 279  
 martingale approximation, 166, 283  
 martingale solution, 249  
 maximal coupling, 16, 98  
 measurable map, 5, 6  
 measurable Polish space, 6  
 measurable space, 5  
 measure-preserving group, 26  
 metric of uniform convergence on bounded intervals, 5  
 minimum of two measures, 18  
 mixing, 103  
 mixing extension, 106  
 modified Skorokhod metric, 57  
 monotone class, 259  
 monotone convergence theorem, 7  
 mutually singular measures, 12
- Navier–Stokes process, 50, 51
- Polish space, 4, 6  
 portmanteau theorem, 14  
 probability space, 6  
 product of measurable spaces, 5  
 progressively measurable random process, 6  
 projection theorem, 261  
 Prokhorov theorem, 13, 14, 84, 105, 195, 210, 212, 228, 266
- quadratic variation, 279
- random dynamical system, 26  
 random kick force, 56  
 random point attractor, 175  
 random probability measure, 18, 30  
 random process, 6  
 random variable, 6  
 rate of dissipation of energy, 55, 86  
 rate of dissipation of enstrophy, 55, 88
- Reynolds number, 55  
 rigged Hilbert space, 267  
 Ruelle–Perron–Frobenius theorem, 150
- smoothing property, 47  
 Sobolev embedding, 3  
 Sobolev space, 1, 2  
 spatially regular white noise, 56  
 squeezing property, 112  
 stationary measure, 23, 31, 61, 73  
 stationary solution, 23, 61, 73  
 Stokes operator, 38, 54  
 stopping time, 23  
 stream function, 47  
 strong Markov property, 23–25  
 submartingale, 280  
 supermartingale, 280
- tight family of measures, 13  
 Tikhonov topology, 5  
 total variation distance, 8, 10  
 transition function, 21, 29  
 trigonometric basis, 43
- Ulam theorem, 6, 12, 85, 224  
 uniformly mixing Markov process, 283  
 universality of white noise, 57, 79, 199  
 universally measurable subset, 261  
 usual hypotheses, 6, 67
- velocity correlation tensors, 143  
 vorticity, 55
- weak convergence of measures, 13  
 weight functions, 122
- Yosida approximation, 266

**Author Index**

- Adams, 33  
 Agrachev, 142, 256  
 Alexandrov, 265  
 Arnold, 33, 201  
  
 Barbu, 161  
 Batchelor, ix, 160, 234  
 Bec, vi  
 Bensoussan, 95  
 Berti, 265  
 Bismut, 140  
 Bogachev, 33  
 Bogoliouboff, 83, 95  
 Bowen, 162  
 Bricmont, 95, 140, 161, 162  
 Brown, 282  
 Brzeźniak, 95  
  
 Chueshov, 202, 241  
 Crauel, 179, 201, 202  
  
 Da Prato, ix, 95, 161, 245  
 Debussche, ix, 95, 161, 163, 201, 245  
 Dobrushin, 18, 33  
 Doebelin, 33, 97  
 Donsker, 57  
 Dudley, 33  
  
 Eckmann, 161  
 Elworthy, 140  
  
 Ferrario, 95, 140, 161  
 Flandoli, ix, 95, 138, 161, 201, 247,  
     251, 253, 258  
 Foaş, 35, 95  
 Fursikov, 95  
  
 Goldys, 140, 161  
 Gordin, 166, 201, 283  
 Gyöngy, 95  
  
 Hairer, 142, 162, 163, 202  
 Harris, 33  
 Heyde, 281  
  
 Kantorovich, 16, 21  
 Karatzas, 33  
 Khanin, vi  
 Kifer, 33  
 Kolmogorov, vi  
 Komech, 95  
 Kraichnan, ix  
 Kryloff, 83, 95  
 Krylov, 95, 245, 251, 276  
 Kuksin, 95, 162, 201, 202, 235, 241,  
     256  
 Kupiainen, 95, 140, 142, 161, 162  
  
 Ladyzhenskaya, 94  
 Langa, 202  
 Le Jan, 179, 202  
 Ledrappier, 179, 202  
 Lefevere, 95, 140, 161, 162  
 Leray, 36, 94  
 Lions, 33, 94  
  
 Magenes, 33  
 Maslowski, 138, 140, 161  
 Masmoudi, 163  
 Mattingly, 95, 142, 162, 163, 202  
 Maz'ja, 33  
  
 Nersesyan, 146  
  
 Odasso, ix, 163, 245  
  
 Pardoux, 142  
 Penrose, 235  
 Pitman, 33  
 Pratelli, 265  
 Prodi, 35, 95  
  
 Röckner, 161  
 Raugel, 238  
 Revuz, 33  
 Rigo, 265  
 Robinson, 202  
 Romito, ix, 247, 251, 253  
 Rozovskii, 161

- Sarychev, 142, 256  
Sell, 238  
Shirikyan, 95, 162–164, 201, 202, 256,  
258  
Shreve, 33  
Sinai, 95, 162  
Stroock, 245
- Taylor, 33  
Temam, 94, 95, 238  
Tubaro, 161
- Varadhan, 245  
Vishik, 95
- Wang, 161  
Weinan E, 95, 162
- Young, 163  
Yudovich, 216
- Ziane, 238

## Citation Index

- [BK07], vi, viii  
 [VF88], 125  
 [Bat82], vii  
 [Bor12], vi, viii  
 [Dud02], 57  
 [Kry96], 244  
 [AKSS07], 142, 164, 256  
 [AS05], 142, 256  
 [AS06], 142, 256  
 [BBM00], 286  
 [BD95], 286  
 [BD07], 161  
 [BF09], 95  
 [BIN79], 4  
 [BK07], 106  
 [BKL00], 88, 95  
 [BKL01], 140, 161  
 [BKL02], 140, 162  
 [BMS11], 201  
 [BPR06], 265  
 [BT73], 95, 248  
 [BV11], 235  
 [BV92], 95, 224  
 [CC92], 248  
 [CDF97], 201  
 [CF94], 201  
 [CF88], 95, 139  
 [CFNT89], 95  
 [CG94], 257  
 [CK08a], 202, 239, 241, 242  
 [CK08b], 242  
 [CV77], 261  
 [CV02], 201  
 [DD02], 95  
 [DD03], 245  
 [DDL<sup>+</sup>07], 201  
 [DDT05], 161  
 [DG95], 50  
 [DM10], 201  
 [DM78], 218  
 [DO05], 163  
 [DO06], 245, 250  
 [DRRW06], 161  
 [DZ92], 68, 72, 269, 270, 280  
 [DZ96], 105, 139, 280  
 [EH01], 162  
 [EMS01], 95, 162, 163  
 [ES00], 163  
 [FG95], 248, 257  
 [FM95], 138, 139, 161  
 [FMRT01], v  
 [FP67], 95  
 [FR08], 247, 252, 253  
 [FT89], 50  
 [GK03], 95  
 [GM05], 140, 161  
 [GS80], 265  
 [HH80], 173  
 [HKL90], 286  
 [HM06], 142, 164  
 [HM08], 142, 164, 202  
 [HM11], 142, 164  
 [JS87], 201  
 [KA82], 16, 21  
 [KB37], 83, 95  
 [KF75], 264  
 [KH95], 95  
 [KP08], 231  
 [KP05], 235  
 [KPS02], 163  
 [KS00], 149, 151, 162  
 [KS01a], 163  
 [KS02b], 162  
 [KS04a], 202  
 [KS02a], 95, 140, 163  
 [KS04b], 232, 235  
 [KS01b], 64, 95, 162  
 [KS03], 88, 95, 202  
 [KS91], 21, 24, 33, 244, 245, 275,  
     281, 298  
 [LM72], 33, 38, 212  
 [LP59], 94  
 [LR06], 190, 202  
 [LS06], 162  
 [LY94], 162  
 [MP06], 142, 163

- [MT93], 201  
 [MY02], 163  
 [PSS89], 100, 106  
 [RS93], 238  
 [RY99], 24, 29, 281  
 [SV79], 245, 247  
 [TZ96], 238  
 [VF88], vi, 74, 95, 248, 257  
 [VKF79], 95  
 [Ada75], 33  
 [Ale43], 265  
 [App04], 57  
 [Arn98], 33, 185, 201  
 [Bat69], 234  
 [Bat82], vii, 160, 200, 234  
 [Bax89], 179  
 [Ber00], 235  
 [Bil99], 57–59, 295  
 [Bis81], 140  
 [Bog98], 262  
 [Bog07], 33  
 [Bol82], 173  
 [Bor12], 106  
 [Bow75], 162  
 [Bri02], 163  
 [Bro71], 282  
 [Che98], 216  
 [Cra91], 202  
 [Cra01], 179  
 [Cra02], 261  
 [Deb97], 190, 202  
 [Deb98], 190, 202  
 [Den04], 201  
 [Dob68], 33  
 [Dob74], 18, 33  
 [Doe38], 33, 97  
 [Doe40], 33, 97  
 [Doo48], 139, 161  
 [Dud02], 6, 13, 14, 21, 29, 31, 33,  
     57, 81, 155, 214, 261, 263,  
     265, 281  
 [Elw92], 140  
 [Fel71], 100, 106, 281  
 [Fer97], 161  
 [Fer99], 140, 161  
 [Fer03], 95  
 [Fla94], 95  
 [Fla97], 258  
 [Fri95], vii, 200  
 [Gal02], vii  
 [Gor69], 201, 283  
 [Hai02b], 147, 163  
 [Hai05], 163  
 [Hai02a], 162  
 [Har55], 33  
 [Har64], 217  
 [Has80], 201  
 [Hen81], 68  
 [Hey73], 281  
 [Hör97], 105  
 [Jud63], 216  
 [Kif86], 33  
 [Kra07], 234  
 [Kry73], 245, 247, 251  
 [Kry74], 276  
 [Kry86], 276  
 [Kry80], 276  
 [Kry87], 277  
 [Kry02], 72, 76, 212, 270, 281  
 [Kuk12], 231, 232  
 [Kuk02b], 163  
 [Kuk08], 235  
 [Kuk10a], 231  
 [Kuk10b], 235  
 [Kuk04], 235  
 [Kuk06b], 208, 235  
 [Kuk02a], 163, 173, 201  
 [Kuk07], 235  
 [Kuk06a], x  
 [Kup10], 142  
 [Lad59], 94  
 [Lad63], 94  
 [Led86], 202  
 [Le 87], 202  
 [Ler34], 94  
 [Lio69], 4, 33, 38, 94, 212  
 [Mat99], 95, 162  
 [Mat02b], 20, 95, 140, 163  
 [Mat03], 163  
 [Maz85], 33  
 [McK69], 68, 76  
 [Mey66], 281

[Ner08], 146, 163  
[Nov05], 20  
[Oda07], 245, 250  
[Oda08], 20, 140, 141, 147, 163  
[Øks03], 72, 274, 278, 279, 295, 296  
[Ons49], 233  
[Pit74], 33  
[Rev84], 33  
[Rio00], 56, 201  
[Rom08], 247, 252  
[Roz90], 270  
[Rue68], 162  
[Sei97], 139  
[Shi07a], 257  
[Shi11a], 164  
[Shi08], 141, 147, 163  
[Shi06a], 257  
[Shi06b], 163  
[Shi05b], 163  
[Shi07b], 252, 258  
[Shi11b], 232, 235  
[Shi04], 95, 163  
[Shi06c], 173, 201  
[Shi05a], 163  
[Shi02], 79, 88, 95  
[Sin91], 162  
[Soh01], 95, 252  
[Ste70], 4, 112, 209  
[Ste94], 139  
[Str93], 26, 132  
[Sza97], 162  
[Tay97], 4, 33, 37, 50, 217  
[Tem68], 94  
[Tem79], 37, 42, 46, 95  
[Tem88], 95  
[Vio75], 95  
[Vio76], 95  
[Yos95], 4, 33