

# Mixing in deterministic and random PDE: lecture notes for the Santa Barbara Kinetic summer school

Lenya Ryzhik\*

June 13, 2016

## 1 Introduction

We will discuss evolution problems of the form

$$\frac{\partial \phi}{\partial t} = iL\phi - \varepsilon\Gamma\phi. \quad (1.1)$$

Here,  $L$  and  $\Gamma$  are two operators, and we think of (1.1) as a perturbation of the background dynamics

$$\frac{\partial \phi}{\partial t} = iL\phi. \quad (1.2)$$

The parameter  $\varepsilon \ll 1$  is small. The operator  $L$  is self-adjoint, so that (1.2) is non-dissipative. The basic question is how a combination of the mixing or dissipative properties of  $L$  and  $\Gamma$  leads to a non-trivial behavior of the solutions of (1.1) on long time scales that is very different both from the dynamics of (1.2) and that of

$$\frac{\partial \phi}{\partial t} = -\Gamma\phi. \quad (1.3)$$

The perturbation  $\Gamma$  may be either random or deterministic.

We will focus on two complementary situations. The first regime is when  $L$  is very mixing but non-dissipative, while  $\Gamma$  is dissipative. Their interaction leads to a dramatically increased dissipation, that can not be achieved by either of them alone. This phenomenon is known as relaxation enhancement. The second set of problems involves random perturbations  $\Gamma$  of a simple deterministic background operator  $L$ . Here, the long time dynamics leads to diffusive limits that are a result of the interaction of the deterministic dynamics generated by  $L$  and the random dynamics coming from  $\Gamma$ . We will look at several examples of such diffusive behavior, in the order of increasing difficulty.

In this introductory section, we will consider two simple examples of each of these two classes of problems, just to illustrate what one may expect.

---

\*Department of Mathematics, Stanford University, Stanford CA, 94305, USA; ryzhik@math.stanford.edu

# An example of relaxation enhancement: the Dirichlet eigenvalues

## The eigenvalues of the Laplacian

Before we explain how relaxation enhancement comes about, let us first recall some very basic facts about the principal Dirichlet eigenvalues for the Laplacian on a bounded domain [14]. For any smooth bounded domain  $\Omega$  there exists an eigenvalue  $\lambda_1$  (called the principal eigenvalue) that corresponds to a positive eigenfunction  $\phi_1 > 0$  in  $\Omega$ :

$$\begin{aligned} -\Delta\phi_1 &= \lambda_1\phi_1, & x \in \Omega, \\ \phi_1 &= 0 \text{ on } \partial\Omega. \end{aligned} \tag{1.4}$$

Moreover,  $\lambda_1$  is the smallest of all eigenvalues of the Dirichlet Laplacian on  $\Omega$ ,  $\lambda_1$  is a simple eigenvalue and all other eigenfunctions of the Laplacian change sign in  $\Omega$ . For example, if  $\Omega$  is an interval  $(0, 1)$ , the eigenvalues of the operator  $Lu = -u''$  with the Dirichlet boundary conditions  $u(0) = u(1) = 0$  are  $\lambda_n = n^2\pi^2$ , and the corresponding eigenfunctions are

$$u_n(x) = \sin(n\pi x).$$

In this case, the principal eigenvalue is  $\lambda_1 = \pi^2$ .

In general, the principal eigenvalue of the Laplacian is given by the variational formula:

$$\lambda_1 = \inf_{\substack{\psi \in H_0^1(\Omega) \\ \|\psi\|_2=1}} \int_{\Omega} |\nabla\psi|^2 dx. \tag{1.5}$$

The principal eigenvalue determines the long time decay of solutions of the parabolic initial value problem in the following way. Consider the initial value problem

$$\begin{aligned} \psi_t &= \Delta\psi, & t > 0, x \in \Omega, \\ \psi(t, x) &= 0 \text{ on } \partial\Omega, \\ \psi(0, x) &= g(x). \end{aligned} \tag{1.6}$$

As  $\phi_1(x) > 0$  in  $\Omega$ , and, as follows from the Hopf lemma,  $\partial\phi_1/\partial\nu < 0$  on  $\partial\Omega$ , we can find a constant  $C > 0$  so that  $|\psi(t = 1, x)| \leq C\phi_1(x)$  – we can not quite have such estimate at  $t = 0$  since the initial condition  $g(x)$  may not satisfy the Dirichlet boundary conditions. The maximum principle implies that

$$\psi(t, x) \leq Ce^{-\lambda_1(t-1)}\phi_1(x), \tag{1.7}$$

for  $t > 1$ , and, similarly,

$$-\psi(t, x) \leq Ce^{-\lambda_1(t-1)}\phi_1(x), \tag{1.8}$$

so that

$$|\psi(t, x)| \leq Ce^{-\lambda_1(t-1)}\phi_1(x), \quad t \geq 1. \tag{1.9}$$

Therefore, all solutions of the Cauchy problem decay at the exponential rate determined by  $\lambda_1$  as  $t \rightarrow +\infty$ .

## The Dirichlet eigenvalues with a drift

Let us now consider the Dirichlet principal eigenvalue problem in a smooth bounded domain  $\Omega$ , for a diffusion with a strong incompressible flow:

$$\begin{aligned} -\Delta\phi + \frac{1}{\varepsilon}u \cdot \nabla\phi &= \lambda_1(\varepsilon)\phi, & \phi(x) > 0 \text{ in } \Omega, \\ \phi &= 0 \text{ on } \partial\Omega. \end{aligned} \tag{1.10}$$

This is an example of a problem like (1.1), with  $\Gamma = -\Delta$ , and  $iL = u \cdot \nabla$ . We assume that  $u$  is an incompressible flow:  $\nabla \cdot u = 0$ , and that it does not penetrate the boundary:

$$u \cdot \nu = 0 \text{ on } \partial\Omega. \tag{1.11}$$

The effect of the positive parameter  $\varepsilon > 0$  will become clear very soon.

The operator in (1.10) is not self-adjoint (so that its eigenvalues are not necessarily real), and its eigenvalues do not obey an integral variational principle such as (1.5). Nevertheless, the Krein-Rutman theory for positive operators (see Chapter VIII of [11]) implies that it has a unique eigenvalue  $\lambda_1(\varepsilon)$  that corresponds to a positive eigenfunction  $\phi_1(x)$ . This eigenvalue is real and simple, has the smallest real part of all eigenvalues, and is called the principal eigenvalue. As for the Laplacian, the maximum principle implies that the principal eigenvalue determines the long time decay of the solutions of the corresponding Cauchy problem:

$$\begin{aligned} \psi_t + \frac{1}{\varepsilon}u \cdot \nabla\psi &= \Delta\psi, & t > 0, x \in \Omega, \\ \psi(t, x) &= 0 \text{ on } \partial\Omega, \\ \psi(0, x) &= g(x), \end{aligned} \tag{1.12}$$

that is,

$$\psi(t, x) \sim e^{-\lambda_1(\varepsilon)t}\phi_1(x), \quad \text{as } t \rightarrow +\infty. \tag{1.13}$$

Note that when  $u = 0$  (or, in our general terminology,  $L = 0$ ) the exponential rate of decay for the solutions of (1.12) is simply the principal eigenvalue of the Laplacian. On the other hand, solutions of the Laplacian-less problem

$$\psi_t + \frac{1}{\varepsilon}u \cdot \nabla\psi = 0 \tag{1.14}$$

do not decay at all – their  $L^2$  norm is preserved, as are all  $L^p$ -norms for  $p \geq 1$ . This is because the flow  $u$  is incompressible and parallel to  $\partial\Omega$  on the boundary.

Let us now understand whether it is possible that solutions of the "combined" Cauchy problem (1.12) decay much faster in time than when  $u = 0$  despite the fact that solutions of (1.14) have no decay whatsoever. To quantify this questions, let us ask if it is possible that

$$\lambda_1(\varepsilon) \rightarrow +\infty \text{ as } \varepsilon \rightarrow 0. \tag{1.15}$$

The above considerations make it clear that such phenomenon may only come from an interaction of the drift and the Laplacian.

Let us recall the probabilistic interpretation of the solutions of the Cauchy problem (1.12). Consider the stochastic differential equation

$$dX_t = -\frac{1}{\varepsilon}u(X_t)dt + \sqrt{2}dW_t, \quad X_0 = x, \quad (1.16)$$

starting at a point  $x \in \Omega$ , and let  $\tau$  be the first time that the process  $X_t$  hits the boundary  $\partial\Omega$ . Then solution of the Cauchy problem (1.12) can be expressed in terms of the diffusion  $X_t$  as

$$\psi(t, x) = \mathbb{E}_x[g(X_{\min(t, \tau)})], \quad (1.17)$$

with the convention that

$$g(X_\tau) = 0. \quad (1.18)$$

When would we expect  $\psi(t, x)$  to be small as  $\varepsilon \rightarrow 0$ ? As one sees from (1.18), this would be true if, with a high probability we have  $\tau < t$  – the particle hits the boundary before a given time  $t$ . Intuitively, if the trajectories of the incompressible flow are “sufficiently mixing”, then, for any starting point  $x_0$  in the interior of  $\Omega$ , the trajectory of (1.16) that starts at  $x_0$  eventually comes close to the boundary  $\partial\Omega$ . Therefore, such flow, when sufficiently fast, will force solutions of (1.17) very quickly to pass very close to  $\partial\Omega$ , and at that time diffusion term in (1.16) will force  $X_t$  to exit  $\Omega$  with a very high probability. Hence, when  $\varepsilon > 0$  is sufficiently small, the exit time  $\tau$  of the solutions of (1.16) should be smaller than a given time  $t > 0$  with a high probability. As we have mentioned, this makes  $\psi(t, x)$  given by (1.17) very small because of (1.18). Physically, this means that a sufficiently mixing flow, together with diffusion, should dramatically increase the cooling of the interior by the boundary. A natural question is what “mixing” means in this context, and how one can quantify such property. Usually, the mixing properties of a flow are defined in terms of the dynamic properties of the ODE

$$\dot{X} = u(X),$$

behave. Here, we are asking a PDE question – hence, the first problem is to define what “mixing” means for us. This is quantified by the following beautiful result due to Berestycki, Hamel and Nadirashvili [4]. We denote by  $\mathcal{I}_0$  the set of all first integrals of  $u$ , solutions of

$$u \cdot \nabla \phi = 0 \text{ a.e. in } \Omega, \quad (1.19)$$

in the space  $H_0^1(\Omega)$ .

**Theorem 1.1** *The principal eigenvalue  $\lambda_1(\varepsilon)$  of (1.10) tends to  $+\infty$  as  $\varepsilon \rightarrow 0$  if and only if the flow  $u$  has no first integral in  $H_0^1(\Omega)$ . Moreover, if  $u$  has a first integral in  $H_0^1(\Omega)$ , then*

$$\lambda_1(\varepsilon) \rightarrow \bar{\lambda} := \min_{w \in \mathcal{I}_0} \frac{\int_{\Omega} |\nabla w|^2 dx}{\int_{\Omega} |w|^2 dx} \text{ as } \varepsilon \rightarrow 0, \quad (1.20)$$

*and the minimum in the right side is achieved.*

A couple of comments are in order. First, notice that the only information about the Laplacian operator in (1.10) that survives in the statement of the theorem is in the condition that the first integral lies in  $H_0^1(\Omega)$ . This regularity requirement comes exactly from the presence of the

Laplacian in (1.10), as irregular first integrals do not prevent strong decay of the solutions of the Cauchy problem. Second, the strong flow essentially forces the eigenfunction to be close to a first integral, and then the variational principle (1.6) for the Laplacian operator is replaced by essentially the same expression (1.20) except that the set of allowed test functions is restricted to the first integrals.

### Proof of Theorem 1.1

The proof of this Theorem is nicely short. First, we claim that if  $u$  has a non-zero first integral  $w$  in  $H_0^1(\Omega)$ , normalized so that

$$\|w\|_{L^2} = 1,$$

then we have

$$0 \leq \lambda_1(\varepsilon) \leq \int_{\Omega} |\nabla w(x)|^2 dx, \quad (1.21)$$

for any  $\varepsilon \in \mathbb{R}$ . In order to show that (1.21) holds, we take any  $w \in \mathcal{I}_0$ , and multiply (1.10) by  $w^2/(\phi + \delta)$  with  $\delta > 0$  fixed:

$$- \int_{\Omega} \frac{w^2 \Delta \phi}{\phi + \delta} dx + \int_{\Omega} \frac{w^2}{\phi + \delta} (u \cdot \nabla \phi) dx = \lambda_1(\varepsilon) \int_{\Omega} \frac{w^2}{\phi + \delta} \phi dx. \quad (1.22)$$

Integrating by parts in the first term gives

$$\begin{aligned} - \int_{\Omega} \frac{w^2 \Delta \phi}{\phi + \delta} dx &= \int_{\Omega} \nabla \phi \cdot \nabla \left( \frac{w^2}{\phi + \delta} \right) dx = \int_{\Omega} \frac{2w(\phi + \delta) \nabla \phi \cdot \nabla w - w^2 |\nabla \phi|^2}{(\phi + \delta)^2} dx \\ &\leq \int_{\Omega} |\nabla w|^2 dx. \end{aligned}$$

The second term in the left side of (1.22) vanishes because  $\nabla \cdot u = 0$  and  $w$  is a first integral:

$$\int_{\Omega} \frac{w^2}{\phi + \delta} (u \cdot \nabla \phi) dx = \int_{\Omega} w^2 (u \cdot \nabla (\log \phi + \delta)) dx = - \int_{\Omega} 2w \log(\phi + \delta) (u \cdot \nabla w) dx = 0.$$

The boundary terms above vanish since  $w \in H_0^1(\Omega)$  (it vanishes on the boundary). We conclude that

$$\lambda_1(\varepsilon) \int_{\Omega} \frac{w^2}{\phi + \delta} \phi dx \leq \int_{\Omega} |\nabla w|^2 dx, \quad (1.23)$$

for any  $w \in \mathcal{I}_0$ . Passing to the limit  $\delta \rightarrow 0$  gives (1.21). Thus, existence of a first integral implies that  $\lambda_1(\varepsilon)$  are uniformly bounded for all  $\varepsilon \in \mathbb{R}$ .

On the other hand, if there exists a sequence  $\varepsilon_n \rightarrow 0$  such that  $\lambda_1(\varepsilon_n)$  are bounded, then

$$\int_{\Omega} |\nabla \phi_n(x)|^2 dx = \lambda_1(\varepsilon_n) \int_{\Omega} |\phi_n(x)|^2 dx = \lambda_1(\varepsilon_n). \quad (1.24)$$

Here,  $\phi_n(x)$  are the associated positive eigenfunctions  $\phi_n(x)$  normalized so that  $\|\phi_n\|_{L^2(\Omega)} = 1$ . Then, there exists a subsequence  $n_k$  so that the sequence  $\phi_{n_k}$  converges weakly in  $H_0^1(\Omega)$  and

strongly in  $L^2(\Omega)$  to a function  $\bar{w}(x) \in H_0^1(\Omega)$ . Moreover, multiplying (1.10) by  $\varepsilon_{n_k}$  and passing to the limit  $k \rightarrow +\infty$  gives

$$u \cdot \nabla \bar{w} = 0, \quad \text{weakly in } H_0^1(\Omega),$$

and

$$\|\bar{w}\|_{L^2(\Omega)} = 1. \quad (1.25)$$

Hence,  $\bar{w}$  is a first integral of  $u$  in  $H_0^1(\Omega)$ . Thus, the non-existence of the first integral in  $H_0^1(\Omega)$  implies that

$$\lim_{\varepsilon \rightarrow 0} \lambda_1(\varepsilon) = +\infty. \quad (1.26)$$

Finally, to show that (1.20) holds, let us assume, once again, that there exists a sequence  $\varepsilon_n \rightarrow 0$  such that  $\lambda_1(\varepsilon_n)$  are bounded. As the convergence of the subsequence  $\phi_{n_k}$  to the first integral  $\bar{w}$  is strong in  $L^2(\Omega)$  and weak in  $H_0^1(\Omega)$ , it follows from (1.24), (1.25) and Fatou's lemma that

$$\liminf_{n \rightarrow +\infty} \lambda_1(\varepsilon_n) \geq \int_{\Omega} |\nabla \bar{w}(x)|^2 dx. \quad (1.27)$$

It remains to notice that (1.27) and (1.21) together imply the Rayleigh quotient formula (1.20), finishing the proof of Theorem 1.1.

### When do things happen in a weakly random medium?

Before turning to an example of the second set of problems of the type (1.1), with a random perturbation  $\Gamma$ , let us illustrate when one can expect a weak random fluctuation to have a non-trivial effect. Probably, the simplest such situation is evolution of a particle in a random time-dependent velocity field:

$$\frac{dX(t)}{dt} = \varepsilon V(t), \quad X(0) = 0, \quad (1.28)$$

that is,

$$X(t) = \varepsilon \int_0^t V(s) ds. \quad (1.29)$$

One can think of (1.28) as (1.1) with the background  $L = 0$ .

We need to make some assumptions on  $V(t)$ : we assume that it is a statistically homogeneous in time random field. Intuitively, it means that the statistics of  $V(t)$  is “the same at all times” – which is a reasonable model for “unknown complex environments”. On a more formal level, this condition holds if given any collection of times  $t_1, t_2, \dots, t_N$ , and a shift  $h$ , the joint law of the random variables  $V(t_1+h), V(t_2+h), \dots, V(t_N+h)$  does not depend on  $h$ . This means, in particular, that the expected value  $\bar{V} = \langle V(t) \rangle$  does not depend on  $t$ , and that the two-point correlation matrix  $R_{ij}(t, s) = \mathbb{E}[V_i(t)V_j(s)]$  depends only on the difference  $t - s$ . Accordingly, we define

$$R_{ij}(t) = \mathbb{E}[V_i(0)V_j(t)],$$

and the power-spectrum matrix as the Fourier transform of the two-point correlation matrix

$$\hat{R}_{ij}(\omega) = \int e^{-it\omega} R_{ij}(t) dt.$$

The stationarity condition can be relaxed to local stationarity – so that the random medium characteristics can vary on a macroscopic or mesoscopic scale but we will not discuss this direction here.

Going back to the particle trajectory (1.29), we see that its average position is

$$\bar{X}(t) = \mathbb{E}[X(t)] = \varepsilon \bar{V}t,$$

where  $\bar{V} = \langle V(0) \rangle$  is the mean velocity. Therefore, if  $\bar{V} \neq 0$ , then the particle moves by a distance  $O(1)$  after a time  $t \sim \varepsilon^{-1}$ , which is by no means a surprising result. If  $\bar{V} = 0$ , then  $\bar{X}(t) = 0$  for all  $t > 0$ , and the way to find out if the particle performs a non-trivial motion is to look at its variance:

$$\begin{aligned} \langle X_i(t)X_j(t) \rangle &= \varepsilon^2 \int_0^t ds_1 \int_0^t ds_2 \mathbb{E}(V_i(s_1)V_j(s_2)) = \varepsilon^2 \int_0^t ds_1 \int_0^t ds_2 R_{ij}(s_1 - s_2) \\ &= \varepsilon^2 \int_0^t ds_1 \int_0^{s_1} ds_2 R_{ij}(s_1 - s_2) + \varepsilon^2 \int_0^t ds_1 \int_{s_1}^t ds_2 R_{ij}(s_1 - s_2) \\ &= \varepsilon^2 \int_0^t ds_1 \int_0^{s_1} ds_2 R_{ij}(s_2) + \varepsilon^2 \int_0^t ds_1 \int_0^{t-s_1} ds_2 R_{ij}(-s_2) \\ &= \varepsilon^2 \int_0^t (t - s_2)[R_{ij}(s_2) + R_{ij}(-s_2)]ds_2 = \varepsilon^2[D_{ij}t + O(1)], \quad \text{as } t \rightarrow +\infty. \end{aligned} \tag{1.30}$$

with the diffusivity matrix

$$D_{ij} = \int_{-\infty}^{\infty} R_{ij}(s)ds = \hat{R}_{ij}(0). \tag{1.31}$$

Expression (1.30) tells us (at least) two things: first, we should expect a non-trivial behavior for the particle at times of the order  $t \sim \varepsilon^{-2}$ , and, second, that the particle behavior at this time scale should be a Brownian motion  $B_D(t)$  with the correlation matrix  $D_{ij}$ . Strictly speaking, we have only computed that its variance agrees with that of  $B_D(t)$  but it is not difficult to make this rigorous.

**Theorem 1.2** *Let  $V(t)$  be a stationary in time random process with mean zero and correlation function*

$$\mathbb{E}(V(s)V(t)) = R(t - s).$$

*Assume that the function  $R(t)$  is of the Schwartz class. Then the process*

$$Y_\varepsilon(t) = \varepsilon \int_0^{t/\varepsilon^2} V(s)ds$$

*converges in law to  $Y(t) = DB(t)$ . Here,  $B(t)$  is the standard Brownian motion, and the diffusion coefficient*

$$D = \int_{-\infty}^{\infty} R(s)ds. \tag{1.32}$$

That is, we have the following result: if  $X(t)$  solves (1.28) with a mean-zero statistically time homogeneous random field  $V(t)$  then the process  $X_\varepsilon(t) = X(t/\varepsilon^2)$  converges, as  $t \rightarrow +\infty$ , to a Brownian motion with the covariance matrix  $D_{ij}$ . The main observation here is that

“mean-zero randomness of size  $\varepsilon$  has a non-trivial effect on the time scales of the order  $\varepsilon^{-2}$ ” – something that any probabilist knows very well from the classical central limit theorem, going at least as far back as de Moivre and 1733.

It is instructive to observe that the diffusivity matrix  $D_{ij}$  is positive-definite (otherwise, the above claim would make no sense). This is a consequence of Bochner’s theorem that asserts that for any statistically time homogeneous process  $V(t) \in \mathbb{R}^n$  the power-spectrum matrix  $\hat{R}_{ij}(\omega)$  is nonnegative-definite for each  $\omega \in \mathbb{R}$ .

Of course, in order for the above discussion to make sense, the diffusivity matrix  $D_{ij}$  needs to be finite – otherwise, obviously, the conclusion can not hold. This imposes a decay condition on the two-point correlation matrix  $D_{ij}$ . What happens if it is violated, that is, if the matrix  $D_{ij}$  is infinite? This tells us that by the times of the order  $t \sim \varepsilon^{-2}$  the particle is “already at infinity”, hence something non-trivial happens before the “classical” times scale  $t \sim \varepsilon^{-2}$  – this has very interesting implications beyond the scope of these notes.

## A toy example: advection equation with a random potential

Let us now give an extremely simple example of dynamics (1.1) with a random operator  $\Gamma$  and  $L \neq 0$ , where everything can be computed quite explicitly. One of the “real” examples is the Schrödinger equation

$$i\phi_t + \Delta\phi - \varepsilon V(x)\phi = 0 \tag{1.33}$$

with a random potential  $iV(x)$ . This is an instance of (1.1) with  $L = \Delta$ , and  $\Gamma$  the multiplication operator by the random function  $V(x)$ . Instead, as a toy model, let us consider a linear advection equation with the same perturbation:

$$\frac{\partial\phi}{\partial t} + \bar{u} \cdot \nabla\phi + i\varepsilon V(x)\phi = 0, \quad \phi(0, x) = \phi_0(x). \tag{1.34}$$

Here,  $\bar{u} \neq 0$  is a constant drift, and  $V(x)$  is a spatially homogeneous random field. For the moment, we will not make the assumptions on  $V(x)$  very precise, except that its mean is zero:

$$\mathbb{E}V(x) \equiv 0. \tag{1.35}$$

The two-point correlation function of  $V(x)$  depends only on the displacement between  $x$  and  $y$ :

$$\mathbb{E}(V(x)V(y)) = R(x - y). \tag{1.36}$$

This follows from spatial homogeneity of  $V(x)$ . We assume that the function  $R(x)$  is of the Schwartz class for now. Equation (1.34) is easy to solve explicitly:

$$\phi(t, x) = \phi_0(x - \bar{u}t) \exp\left(-i\varepsilon \int_0^t V(x - s\bar{u}) ds\right). \tag{1.37}$$

In order for the integral in the exponential to have a non-trivial effect, we need to wait until times of the order  $t \sim O(\varepsilon^{-2})$ . Hence, we define

$$\phi_\varepsilon(t, x) = \phi\left(\frac{t}{\varepsilon^2}, x\right), \tag{1.38}$$



so that

$$\phi_\varepsilon(t, x) = \phi_0\left(x - \frac{t}{\varepsilon^2}\bar{u}\right) \exp\left(-i\varepsilon \int_0^{t/\varepsilon^2} V(x - s\bar{u})ds\right). \quad (1.39)$$

The first term is simply the solution of the "fast" homogeneous problem (1.34) with  $V = 0$ , and the second factor

$$\zeta_\varepsilon(t, x) = \exp\left(-i\varepsilon \int_0^{t/\varepsilon^2} V(x - s\bar{u})ds\right) \quad (1.40)$$

comes from the "slow" random dynamics. Its limit can be computed from Theorem 1.2, which implies that  $\zeta_\varepsilon(t, x)$  converges in law to

$$\bar{\zeta}(t, x) = \exp(-i\sqrt{D}B(t)). \quad (1.41)$$

The diffusion coefficient  $D$  is

$$D = \int_{-\infty}^{\infty} R(\bar{u}s)ds. \quad (1.42)$$

Let us formulate this result as a theorem.

**Theorem 1.3** *Let  $V(x)$  be a spatially homogeneous mean-zero random field with a correlation function  $R(t) \in \mathcal{S}(\mathbb{R})$ . Let  $\phi(t, x)$  be the solution of (1.34),  $\phi_\varepsilon(t, x) = \phi(t/\varepsilon^2, x)$ , and  $\bar{\phi}(t, x)$  the solution of (1.34) with  $V = 0$ . Then  $\phi_\varepsilon(t, x)$  can be decomposed as*

$$\phi_\varepsilon(t, x) = \bar{\phi}\left(\frac{t}{\varepsilon^2}, x\right)\zeta_\varepsilon(t, x). \quad (1.43)$$

*The function  $\zeta_\varepsilon(t, x)$  converges in law, as  $\varepsilon \rightarrow 0$  to  $\bar{\zeta}(t, x)$  given by (1.41).*

This example is very simple but it has the main features that are much harder to prove in even slightly more complicated situations. In particular, the dynamics can be decomposed into a fast deterministic part that does not have a limit but is quite explicit, and the "slow" component that converges in law to a stochastic limit. As we have noticed before, the background dynamics is crucial for the limit theorem here: if  $\bar{u} = 0$ , the conclusion of Theorem 1.3 fails. Thus, the long time dynamics of the solutions of (1.34) involves a non-trivial interaction of the background dynamics and the random fluctuations.

## Organization of the notes

The notes will be organized one day.

## 2 Relaxation enhancement

### Relaxation enhancing flows

As we have discussed, one interpretation of the eigenvalue enhancement estimate in Theorem 1.1 is in terms of the long time decay rate of the solution of the Cauchy problem

$$\begin{aligned} \psi_t + \frac{1}{\varepsilon}u \cdot \nabla\psi &= \Delta\psi, \quad t > 0, x \in \Omega, \\ \psi(t, x) &= 0 \text{ on } \partial\Omega, \\ \psi(0, x) &= g(x), \end{aligned} \quad (2.1)$$

in  $\Omega$  with the Dirichlet boundary condition. Its solution has the long time asymptotics

$$\psi(t, x) \sim e^{-\lambda_1(\varepsilon)t} \phi(x) \quad (2.2)$$

for  $t \gg 1$ . Here,  $\phi(x)$  is the principal eigenfunction of the operator

$$-\Delta\phi + \frac{1}{\varepsilon}u \cdot \nabla\phi = \lambda_1(\varepsilon)\phi, \quad (2.3)$$

with the Dirichlet boundary conditions. We have seen in Theorem 1.1 that the principal eigenvalue, or the exponential rate of decay in (2.2), satisfies

$$\lambda_1(\varepsilon) \rightarrow +\infty \text{ as } \varepsilon \rightarrow 0 \quad (2.4)$$

if and only if the flow  $u$  has no first integrals in  $H_0^1(\Omega)$ .

Here, we focus on similar questions in the case of a compact manifold without boundary or the Neumann boundary conditions. Then, the principal eigenvalue is simply zero and corresponds to the constant eigenfunction. One may instead study the second eigenvalue but that is not simple since we do not even know a priori that the second eigenvalue is real, and finding estimates for the real part of a complex eigenvalue that corresponds to an eigenfunction that also need not be real would not be an easy task. Moreover, even if the spectral gap estimate were available, generally it only provides a long time dynamical information, and how fast the long time limit is achieved may depend on  $\varepsilon$ , since the operator in the left side of (2.3) is neither self-adjoint nor normal: it does not commute with its formal adjoint operator

$$L^*\phi = -\Delta\phi - \frac{1}{\varepsilon}\nabla \cdot (u\phi).$$

On the other hand, our general interest is in the speed of convergence of the solution to its average, the relaxation speed, and there are other ways to measure than in terms of the spectrum. Therefore, rather than try to address the spectral behavior, we will reformulate our questions purely in terms of the Cauchy problem.

Let  $\Omega$  be a smooth compact  $n$ -dimensional Riemannian manifold. We consider solutions of the passive scalar equation

$$\phi_t^\varepsilon + \frac{1}{\varepsilon}u(x) \cdot \nabla\phi^\varepsilon - \Delta\phi^\varepsilon = 0, \quad \phi^\varepsilon(0, x) = \phi_0(x). \quad (2.5)$$

Here,  $\Delta$  is the Laplace-Beltrami operator on  $\Omega$ ,  $u$  is a divergence free vector field,  $\nabla$  is the covariant derivative, and  $\varepsilon > 0$  is a parameter regulating the strength of the flow. For the sake of concreteness we will assume that  $\Omega$  is a torus, and (2.5) is supplemented by periodic boundary conditions but the results of this section apply verbatim to advection-diffusion equations on a compact manifold without boundary, or with the Neumann boundary conditions on a compact domain  $\Omega \subset \mathbb{R}^n$ , and with very slight modifications to the Robin boundary conditions on such domains – see [8] for the full cornucopia.

As time tends to infinity, the solution  $\phi^\varepsilon(t, x)$  tends to its average,

$$\bar{\phi}(t) \equiv \frac{1}{|\Omega|} \int_{\Omega} \phi^\varepsilon(t, x) d\mu = \frac{1}{|\Omega|} \int_{\Omega} \phi_0(x) dx. \quad (2.6)$$

Here  $|\Omega|$  is the volume of  $\Omega$ . To see that, first, integrating (2.5) over  $M$  and using incompressibility of  $u$  gives

$$\frac{d}{dt} \int_{\Omega} \phi(t, x) dx = 0,$$

hence  $\bar{\phi}(t) = \bar{\phi}(0)$  is preserved in time. Next, multiplying (2.5) by  $\phi^\varepsilon(t, x) - \bar{\phi}$ , and using incompressibility of  $u(x)$ , we have

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} |\phi(t, x) - \bar{\phi}|^2 dx = - \int_{\Omega} |\nabla \phi^\varepsilon(t, x)|^2 dx. \quad (2.7)$$

The Poincaré inequality implies that

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} |\phi(t, x) - \bar{\phi}|^2 dx \leq -C_p \int_{\Omega} |\phi^\varepsilon(t, x) - \bar{\phi}|^2 dx, \quad (2.8)$$

whence

$$\|\phi(t, \cdot) - \bar{\phi}\|_{L^2(\Omega)} \leq e^{-C_p t} \|\phi_0 - \bar{\phi}\|_{L^2(\Omega)}. \quad (2.9)$$

**Exercise 2.1** Strengthen this result to show that

$$\|\phi(t, \cdot) - \bar{\phi}\|_{L^\infty(\Omega)} \rightarrow 0 \text{ as } t \rightarrow +\infty. \quad (2.10)$$

We would like to understand how the speed of convergence to the average in (2.10) depends on the properties of the flow and determine which flows are particularly efficient in enhancing the relaxation process. We will use the following "fixed time" (no long time limit!) definition as a measure of the flow efficiency in improving the relaxation of the solution to a uniform state.

**Definition 2.2** *An incompressible flow  $u$  is relaxation enhancing if for all  $\tau > 0$  and  $\delta > 0$ , there exists  $\varepsilon(\tau, \delta)$  such that for any  $\varepsilon < \varepsilon(\tau, \delta)$  and any  $\phi_0 \in L^2(\Omega)$ , with  $\|\phi_0\|_{L^2(\Omega)} = 1$ , we have*

$$\|\phi^\varepsilon(\tau, \cdot) - \bar{\phi}\|_{L^2(\Omega)} < \delta, \quad (2.11)$$

where  $\phi^\varepsilon(t, x)$  is the solution of (2.5) and  $\bar{\phi}$  the average of  $\phi_0$ .

**Exercise 2.3** Show that the choice of the  $L^2$  norm in the definition is not essential and can be replaced by any  $L^p$ -norm with  $1 \leq p \leq \infty$ , without changing the class of relaxation enhancing flows.

Let us mention that there are various results on Gaussian and other estimates on the heat kernel corresponding to the incompressible drift and diffusion on manifolds such as in the work of Norris [41] and Franke [20], but these estimates lead to upper bounds on the convergence rate to the equilibrium which essentially do not improve as  $\varepsilon \rightarrow 0$ , and thus do not quite address the effect of a strong flow. Such general estimates often deteriorate as the flow gets stronger, which is exactly the opposite of what interests us.

The original motivation for this definition came from the work of Fannjiang, Nonnemacher and Wolowski [15, 16, 17], where relaxation enhancement was studied in the discrete setting (see also [33] for related earlier references). In these papers, a unitary evolution step (a

certain measure preserving map on the torus) alternates with a dissipation step, which, for example, acts simply by multiplying the Fourier coefficients by damping factors. The absence of sufficiently regular eigenfunctions appears as a key for the enhanced relaxation in this particular class of dynamical systems. In [15, 16, 17], the authors also provide finer estimates of the dissipation time for particular classes of toral automorphisms – they estimate how many steps are needed to reduce the  $L^2$  norm of the solution by a factor of two if the dissipation strength is  $\varepsilon$ .

To understand why and when we expect relaxation enhancement, let us first look at the time-splitting approximation for (2.5), in the spirit of [15, 16, 17]. Assume that  $\psi(t, x)$  solves the advection equation

$$\psi_t + \frac{2}{\varepsilon} u \cdot \nabla \psi = 0, \quad n\tau \leq t \leq (n + 1/2)\tau, \quad (2.12)$$

followed by the heat equation

$$\psi_t = 2\Delta\psi, \quad (n + 1/2)\tau \leq n\tau, \quad (2.13)$$

and then again (2.12) followed by (2.13), and so on. As the time step  $\tau \rightarrow 0$ , the solution of this time-splitting scheme converges to the solution of (2.12). However, the smallness of  $\tau$  that is required to make the error small depends on  $\varepsilon$  in a way that is very difficult to control efficiently. If we, in a cavalier fashion, instead fix the size of  $\tau$  that is independent of  $\varepsilon$ , then solution of the very first step is

$$\psi(\tau/2, x) = \phi_0(X(\tau/\varepsilon, x)), \quad (2.14)$$

where  $X(t, x)$  is the trajectory

$$\dot{X}(t) = -u(X), \quad X(0) = x. \quad (2.15)$$

If the flow of (2.15) is sufficiently complex and  $\varepsilon$  is sufficiently small, the points  $X(\tau/\varepsilon, x)$  and  $X(\tau/\varepsilon, x')$  may be very far apart, even if  $x$  and  $x'$  are very close. This would make the difference  $\psi(\tau/2, x) - \psi(\tau/2, x')$  large, so that the function  $\psi(\tau/2, x)$  given by (2.14) would have a large gradient. This means that the initial condition for the second step in the time-splitting scheme

$$\psi_t = 2\Delta\psi, \quad \tau/2 \leq \tau, \quad (2.16)$$

has a very large gradient. On the other hand, the dissipation identity for (2.16)

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} |\psi - \bar{\psi}|^2 = -2 \int_{\Omega} |\nabla\psi|^2 dx \quad (2.17)$$

tells us that solutions with a large gradient and zero average decay very fast. Therefore, we would deduce that for "sufficiently mixing" flows  $u(x)$  solutions of this time splitting scheme converge to their average very fast if  $\varepsilon$  is small. The problem with making this argument rigorous is that, as we have mentioned, for the convergence of the time-splitting scheme to the true solution we would need to take  $\tau$  not fixed but  $\tau \ll \varepsilon$ , making the interaction of advection and diffusion non-trivial and very difficult to account for carefully. Nevertheless, this intuition is correct. Here is the main result of this section.

**Theorem 2.4** ([8]) *A Lipschitz continuous incompressible flow  $u \in Lip(\Omega)$  is relaxation enhancing if and only if the operator  $u \cdot \nabla$  has no eigenfunctions in  $H^1(\Omega)$ , other than the constant function.*

The "sufficiently mixing" property of  $u$  is encoded in this theorem in the requirement that it does not have an eigenfunction in  $H^1(\Omega)$ . The reason for that condition can also be seen from the time-splitting scheme (2.14)-(2.16). The operator  $u \cdot \nabla$  is skew-symmetric when  $u$  is divergence free:

$$\int_{\Omega} (u \cdot \nabla \eta(x)) \eta(x) dx = 0, \quad (2.18)$$

for all  $\eta \in H^1(\Omega)$ . Therefore, all its eigenvalues  $\lambda = i\omega$  are purely imaginary, and if  $w \in H^1(\Omega)$  is an eigenfunction, then solution of (2.14)

$$\psi_t + \frac{2}{\varepsilon} u \cdot \nabla \psi = 0, \quad 0 \leq t \leq \tau/2 \quad (2.19)$$

with the initial data  $\psi(0, x) = w(x)$  satisfies

$$\psi(t, x) = e^{2i\omega t/\varepsilon} w(x).$$

Therefore,  $\|\psi(\tau/2, x)\|_{H^1(\Omega)} = \|w\|_{H^1(\Omega)}$  does not increase, hence the advection step does not prepare an irregular initial data for the heat equation in the second step, and there is no intuitive reason to expect relaxation enhancement when  $\varepsilon \rightarrow 0$ .

Furthermore, as in the eigenvalue enhancement in Theorem 1.1 the way the Laplacian operator (that is responsible for dissipation) enters the statement of Theorem 2.4 is in the requirement that the eigenfunction lies in the space  $H^1(\Omega)$  – rough eigenfunctions (outside of  $H^1(\Omega)$ ) do not prevent a flow  $u(x)$  from being relaxation enhancing.

The discrepancy between Theorems 1.1 and 2.4 may seem surprising – after all, on the physical level the conditions for the relaxation enhancement and eigenvalue enhancement need not be very different but the eigenvalue enhancement (with the Dirichlet boundary conditions) requires that the operator  $u \cdot \nabla$  does not have first integrals while relaxation enhancement (with the periodic or Neumann boundary conditions) requires that this operator does not have eigenfunctions in  $H^1(\Omega)$  with any eigenvalue (the first integral corresponds to a zero eigenvalue). This issue is resolved by the following

**Proposition 2.5** *Let  $u \in Lip(\Omega)$ . If  $\phi \in H^1(\Omega)$  is an eigenfunction of the operator  $u \cdot \nabla$  corresponding to an eigenvalue  $i\omega$ ,  $\omega \in \mathbb{R}$ , then  $|\phi| \in H^1(\Omega)$  and it is the first integral of  $u$ , that is,  $u \cdot \nabla |\phi| = 0$ .*

**Proof.** The fact that  $|\phi| \in H^1$  follows from the well-known properties of Sobolev functions (see, for example, [14]). If  $\phi(x)$  satisfies

$$u \cdot \nabla \phi = i\omega \phi$$

then

$$u \cdot \nabla |\phi|^2 = u \cdot \nabla (\phi \bar{\phi}) = \phi(u \cdot \nabla \bar{\phi}) + \bar{\phi}(u \cdot \nabla \phi) = -i\omega \phi \bar{\phi} + i\omega \phi \bar{\phi} = 0,$$

hence  $u \cdot \nabla |\phi| = 0$ .  $\square$

Therefore, in the case of the Dirichlet boundary conditions, if  $\phi \in H_0^1(\Omega)$  is an eigenfunction of the operator  $u \cdot \nabla$  then  $|\phi|$  is its first integral. Naturally,  $|\phi|$  can not be equal identically to a constant since  $\phi$  satisfies the Dirichlet boundary conditions, as it lies in  $H_0^1(\Omega)$ , and  $\phi \neq 0$ . Moreover, if  $\phi \in H_0^1(\Omega)$  is a first integral:  $u \cdot \nabla \phi = 0$  then it is an eigenfunction corresponding to eigenvalue  $\lambda = 0$ . Hence, for the Dirichlet boundary conditions the requirement that  $u \cdot \nabla$  does not have a first integral in  $H_0^1(\Omega)$  is equivalent to the condition that it does not have eigenfunctions in  $H_0^1(\Omega)$ .

On the other hand, existence of mean zero  $H^1(\Omega)$  eigenfunctions, without imposing the Dirichlet boundary condition, need not guarantee the existence of a mean zero first integral, as can be seen from the following well-known example. Let  $\alpha \in \mathbb{R}^n$  be a constant vector generating an irrational rotation on the  $n$ -dimensional torus  $\Omega$ , in the sense that the components of  $\alpha$  are independent over the rationals. The operator  $\alpha \cdot \nabla$  has eigenvalues  $2\pi i(\alpha \cdot k)$ , with any  $k \in \mathbb{Z}^n$ . The corresponding eigenfunctions are

$$w_k(x) = e^{2\pi i k \cdot x}.$$

Their absolute value is 1, which is a first integral of  $\alpha \cdot \nabla$  but there are no other first integrals since  $\alpha$  is irrational. Indeed, if there exists a function  $\psi \in L^1(\Omega)$  such that

$$\psi(x + \alpha t) = \psi(x), \quad \text{for all } x \in \Omega \text{ and all } t \in \mathbb{R},$$

then the Fourier coefficients of the function  $\psi$ , defined by

$$\psi(x) = \sum_{k \in \mathbb{Z}^n} e^{2\pi i k \cdot x} \hat{\psi}_k, \quad \hat{\psi}_k = \int_{\Omega} e^{-2\pi i k \cdot y} \psi(y) dy,$$

should satisfy

$$\hat{\psi}_k = e^{2\pi i k \cdot \alpha t} \hat{\psi}_k, \quad \text{for all } k \in \mathbb{Z}^n, \text{ and all } t \in \mathbb{R}.$$

Therefore, either all  $\hat{\psi}_k = 0$  for  $k \neq 0$ , or there exists  $k \neq 0$  such that

$$k \cdot \alpha = 0.$$

The latter, however, is impossible since  $\alpha$  is irrational. Hence,  $\hat{\psi}_k = 0$  for all  $k \neq 0$ , and the only first integrals of  $\alpha \cdot \nabla$  for an irrational  $\alpha$  are constant functions. Thus, this flow is not relaxation enhancing, since it has eigenfunctions in  $H^1(\Omega)$ , even though it has no first integrals other than a constant function.

## Mixing and weakly mixing flows

An important class of relaxation enhancing flows is given by mixing and weakly mixing flows. Let us recall how they are defined. An incompressible flow  $u \in \text{Lip}(\Omega)$  generates a unitary evolution group  $U^t$  on  $L^2(\Omega)$ , defined by

$$U^t f(x) = f(X(t; x)).$$

Here,  $X(t; x)$  is the measure preserving map associated with the flow, defined by (2.15):

$$\frac{dX}{dt} = -u(X), \quad X(0; x) = x. \quad (2.20)$$

The group  $U^t$  is a convenient tool to set up everything in  $L^2(\Omega)$  rather than at the level of trajectories. For example, a function  $f$  is a first integral of  $u$  if

$$U^t f(x) = f(x), \text{ for all } x \in \Omega \text{ and all } t \in \mathbb{R}. \quad (2.21)$$

We say that a flow  $u$  is ergodic if its only first integrals are constant functions.

A flow is mixing if the following condition holds: for any two functions  $f, h \in L^2(\Omega)$  we have

$$\lim_{t \rightarrow +\infty} \int_{\Omega} f(X(t; x))h(x)dx = \int_{\Omega} f(x)dx \int_{\Omega} h(x)dx. \quad (2.22)$$

The mixing condition (2.22) can be interpreted as follows. Let us start (2.20) at a random point  $x$ , equally distributed over the set  $\Omega$ . The Lebesgue measure on  $\Omega$  is invariant under the dynamics (2.20) since  $u$  is incompressible: for any measurable set  $A$  we have

$$\mathbb{P}(X(t) \in A) = \int_{\Omega} \chi_A(X(t; x))dx = \int_{\Omega} \chi_A(x)dx = |A|.$$

Consider two measurable sets  $A \subset \Omega$  and  $B \subset \Omega$ , and the corresponding characteristic functions  $h(x) = \chi_A(x)$  and  $f(x) = \chi_B(x)$ . Then (2.22) says that

$$P(X(t) \in B \text{ and } X(0) \in A) - |A| \cdot |B| \rightarrow 0, \text{ as } t \rightarrow +\infty, \quad (2.23)$$

that is, the events  $\{X(0) \in A\}$  and  $\{X(t) \in B\}$  become asymptotically (as  $t \rightarrow +\infty$ ) independent – the fact that you end up in  $B$  does not depend on where you start.

Mixing implies ergodicity: if  $U^t f(x) = f(x)$  for all  $t \in \mathbb{R}$  then

$$\int_{\Omega} f(X(t; x))h(x)dx = \int_{\Omega} f(x)h(x)dx, \text{ for all } t > 0, \quad (2.24)$$

for all  $h \in L^2(\Omega)$  which is incompatible with mixing unless  $f$  is a constant function.

A function  $f \in L^2(\Omega)$  is an eigenfunction of the flow  $u$  if for any  $t \in \mathbb{R}$  there exists  $c(t)$  so that

$$U^t f(x) = c(t)f(x). \quad (2.25)$$

This definition is equivalent to the condition that

$$u \cdot \nabla f = \lambda f. \quad (2.26)$$

Indeed, the function  $g(t, x) = U^t f(x)$  satisfies the advection equation

$$g_t + u \cdot \nabla g = 0, \quad g(0, x) = f(x). \quad (2.27)$$

therefore, if (2.26) holds then

$$f(X(t, x)) = e^{\lambda t} f(x).$$

On the other hand, if (2.25) holds then the solution of (2.27) has the form

$$g(t, x) = c(t)f(x).$$

Inserting this expression into (2.27) gives

$$\dot{c}(t)f(x) + c(t)u(x) \cdot \nabla f(x) = 0.$$

Separation of variables now implies that there exists  $\lambda \in \mathbb{C}$  such that (note that  $c(0) = 1$  automatically)

$$c(t) = e^{-\lambda t},$$

and

$$u(x) \cdot \nabla f(x) = \lambda f(x).$$

Moreover, as the map  $x \rightarrow X(t; x)$  is measure preserving for all  $t \in \mathbb{R}$ ,  $|c(t)| = 1$  for all  $t$ , whence  $\lambda$  is purely imaginary:  $\lambda = i\omega$  with a real number  $\omega$ .

An incompressible flow  $u$  is called weakly mixing if the corresponding operators  $U^t$  have only continuous spectrum, that is, the only eigenfunctions of  $U^t$  are constants. An equivalent definition is that (2.22) holds on average, that is:

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \left| \int_{\Omega} f(X(t; x))h(x)dx - \int_{\Omega} f(x)dx \int_{\Omega} h(x)dx \right| dt = 0, \quad (2.28)$$

and the convergence in (2.23) holds for a set of times of density one.

Weakly mixing flows are ergodic: first integrals are eigenfunctions with eigenvalue zero but weakly mixing flows are not necessarily mixing (see, for instance, [9]). On the other hand, mixing flows are weakly mixing: essentially for the same reason that mixing flows are ergodic – if

$$U^t f = c(t)f, \text{ for all } t \in \mathbb{R},$$

then

$$\int_{\Omega} f(X(t; x))h(x)dx = c(t) \int_{\Omega} f(x)h(x)dx, \text{ for all } t > 0, \quad (2.29)$$

for all  $h \in L^2(\Omega)$  which is also incompatible with mixing unless  $f$  is a constant function.

A direct consequence of Theorem 2.4 is the following Corollary.

**Corollary 2.6** *Any weakly mixing incompressible flow  $u \in Lip(\Omega)$  is relaxation enhancing.*

There exist fairly explicit examples of weakly mixing flows [2, 18, 19, 35, 44, 40], some of which we describe below but delving into the detailed constructions would take us too far outside of the PDE realm.

## Examples of relaxation enhancing flows

Before embarking on the proof of Theorem 2.4 we present in this section some examples of relaxation enhancing flows on a torus so as to assure the reader that this class is not empty. We first describe flows that have no eigenfunctions – they are weakly mixing, and then flows with very rough eigenfunctions none of which lie in  $H^1(\Omega)$ .



## Weakly mixing incompressible flows on a torus

According to Theorem 2.4, a flow  $u \in \text{Lip}(\Omega)$  is relaxation enhancing if and only if it has no eigenfunctions in  $H^1(\Omega)$ . As we have mentioned, a natural class satisfying this condition is weakly mixing flows – which have no eigenfunctions in  $L^2(\Omega)$  at all. Examples of weakly mixing flows on  $\mathbb{T}^2$  go back to von Neumann [40] and Kolmogorov [35]. The flow in von Neumann’s example is continuous, in the construction suggested by Kolmogorov the flow is smooth. The technical details of Kolmogorov’s construction have been carried out in [44], a good review of these results is [27]. More recently, Fayad [18] generalized this example to show that weakly mixing flows are generic. To describe the result of [18] in more detail, let us recall that a vector  $\alpha \in \mathbb{R}^n$  is called  $\beta$ -Diophantine if there exists a constant  $C$  such that for each  $k \in \mathbb{Z}^n \setminus \{0\}$  we have

$$\inf_{p \in \mathbb{Z}} |\langle \alpha, k \rangle + p| \geq \frac{C}{|k|^{n+\beta}}.$$

The vector  $\alpha$  is Liouvillean if it is not Diophantine for any  $\beta > 0$ . The Liouvillean numbers (and vectors) are the ones which can be very well approximated by rationals.

In order to construct a weakly mixing flow on a torus  $\mathbb{T}^{n+1}$  we start with a very simple flow that is a local time change of a linear translation flow:

$$\frac{dX}{dt} = \frac{\alpha}{F(X, Y)}, \quad \frac{dY}{dt} = \frac{1}{F(X, Y)}, \quad X(0) = x, \quad Y(0) = y, \quad (2.30)$$

with a smooth positive function  $F(x, y)$ ,  $x \in \mathbb{T}^n$ ,  $y \in \mathbb{T}$ . Such flows have a unique invariant measure

$$d\mu = F(x, y) dx dy.$$

Indeed, for any smooth function  $f(x, y)$  set

$$g(t, x, y) = U^t f(x, y) = f(X(t, x, y), Y(t, x, y)),$$

so that

$$\int_{\mathbb{T}^{n+1}} U^t f d\mu = \int_{\mathbb{T}^{n+1}} f(X(t, x, y), Y(t, x, y)) d\mu = \int_{\mathbb{T}^{n+1}} g(t, x, y) d\mu.$$

This function satisfies the first order advection equation

$$\frac{\partial g}{\partial t} - \frac{\alpha}{F(x, y)} \cdot \nabla_x g - \frac{1}{F(x, y)} \frac{\partial g}{\partial y} = 0, \quad g(0, x, y) = f(x, y). \quad (2.31)$$

Using this equation, we compute:

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{T}^{n+1}} U^t f d\mu &= \frac{d}{dt} \int_{\mathbb{T}^{n+1}} g(t, x, y) d\mu = \int_{\mathbb{T}^{n+1}} \left[ \frac{\alpha}{F(x, y)} \cdot \nabla_x g(x, y) + \frac{1}{F(x, y)} \frac{\partial g(x, y)}{\partial y} \right] d\mu \\ &= \int_{\mathbb{T}^{n+1}} \left[ \alpha \cdot \nabla_x g(t, x, y) + \frac{\partial g(t, x, y)}{\partial y} \right] dx dy = 0. \end{aligned} \quad (2.32)$$

Therefore,  $\mu$  is, indeed, an invariant measure for the flow (2.30). Let us denote by  $C_+^r(\mathbb{T}^d)$  the set of positive  $C^r$  functions on the torus. Fayad’s result is

**Proposition 2.7** ([18]) *Assume that the irrational vector  $\alpha \in \mathbb{R}^d$  is not  $\beta$ -Diophantine, for some  $\beta > 0$ . Then, for a dense  $G_\delta$  set of functions  $F$  in  $C_+^{\beta+n}(\mathbb{T}^{n+1})$  the flow (2.30) is weakly mixing (for the unique invariant measure  $F(x, y)dxdy$ ).*

The flows given by this proposition have an invariant measure  $F(x, y)dxdy$  and not the Lebesgue measure  $dxdy$  – they are not quite divergence free, rather they satisfy

$$\nabla \cdot (Fu) = 0.$$

To obtain examples of relaxation enhancing flows, we need to modify these flows so that the resulting flow is divergence-free but the weakly mixing property is preserved. Obviously, if the flow (2.30) is weakly mixing for some  $F$  then it is also weakly mixing for all positive multiples of  $F$  – multiplication of  $F$  by a constant amounts to a simple rescaling of time by the same constant. Thus, we may assume that  $F$  given by Proposition 2.7 is normalized so that

$$\int_{\mathbb{T}^{n+1}} F(x, y)dxdy = 1. \quad (2.33)$$

As  $\mathbb{T}^{n+1}$  is the unit torus, a theorem of Moser [39] says that then there exists a measure preserving invertible transformation

$$Z : \mathbb{T}^{n+1}(F(x, y)dxdy) \rightarrow \mathbb{T}^{n+1}(dpdq),$$

that is as smooth as the function  $F$ . If we denote by

$$w(x, y) = (\alpha/F(x, y), 1/F(x, y))$$

the vector field in (2.30), then the vector field

$$u(p, q) = Z \circ w \circ Z^{-1}$$

is going to be incompressible (with respect to the standard Lebesgue measure  $dpdq$ ). Moreover, the unitary evolutions generated by  $w$  and  $u$  in  $L^2(F(x, y)dxdy)$  and  $L^2(dp dq)$ , respectively, are unitary equivalent and so have the same spectra. Therefore, the flow  $u(p, q)$  is weakly mixing and thus relaxation enhancing.

## Flows with rough eigenfunctions

We now describe an example of a different class of flows to which Theorem 2.4 applies. Namely, we will sketch a construction of a smooth incompressible flow  $u(x, y)$ ,  $\nabla \cdot u = 0$ , on a torus  $\mathbb{T}^2$  that has a purely discrete spectrum but none of the eigenfunctions are in  $H^1(\mathbb{T}^2)$ . The idea of the construction goes back to Kolmogorov [35]. We only sketch the construction, without presenting the full technical details [2, 27].

As before, we denote by  $U^t$  the flow on  $L^2(\mathbb{T}^2)$  generated by  $u$ :

$$U^t f(x) = f(X(t; x)),$$

and by  $X(t, x)$  the trajectory of

$$\frac{dX}{dt} = -u(X), \quad X(0; x) = x.$$

**Proposition 2.8** *There exists a smooth incompressible (with respect to the Lebesgue measure) flow  $u(x, y)$  on a two-dimensional torus  $\mathbb{T}^2$  so that the corresponding unitary evolution  $U^t$  has a discrete spectrum on  $L^2(\mathbb{T}^2)$  but none of the eigenfunctions of  $U^t$  are in  $H^1(\mathbb{T}^2)$ .*

**Proof.** The example will be given by a flow of the type (2.30):

$$\frac{dX}{dt} = \frac{\alpha}{F(X, Y)}, \quad \frac{dY}{dt} = \frac{1}{F(X, Y)}, \quad X(0) = x, \quad Y(0) = y, \quad (2.34)$$

with  $d = 2$  and appropriately chosen  $\alpha \in \mathbb{R}$  and  $F(x, y)$ . The idea of the construction is to find a flow of this form which can be mapped to a constant flow  $(\alpha, 1)$  by a measure preserving map  $S$  with very low regularity properties. Since the eigenfunctions of the constant flow are explicitly computable, we can compute the eigenfunctions of the original flow. Due to the roughness of  $S$ , these will be highly irregular. To obtain an incompressible flow, rather than a flow that preserves the measure  $F(x, y)dxdy$ , we will then proceed as in the previous example.

It would be very convenient to take  $F(x, y)$  in the form

$$F(x, y) = Q(x - \alpha y), \quad (2.35)$$

so that  $F(x, y)$  would be constant on each trajectory of the flow (2.34). However, for  $F(x, y)$  as in (2.35) to be 1-periodic both in  $x$  and  $y$ , the function  $Q(x)$  has to be both 1-periodic and  $\alpha$ -periodic. If  $\alpha$  is irrational, this is impossible unless  $Q(x) \equiv 1$ .

Thus, instead of trying (2.35), we take a 1-periodic function  $Q(x) > 0$ , and choose  $m > 0$  so that  $m < \min Q(s)$ . We also take a smooth cut-off function  $\psi(y) \geq 0$  such that

$$\int_0^1 \psi(y)dy = 1, \quad (2.36)$$

and

$$\psi(y) = 0 \text{ for } 0 \leq y \leq y_0 \text{ and } y_1 \leq y \leq 1 \text{ with } y_0 \text{ close to zero and } y_1 \text{ close to one.} \quad (2.37)$$

The choice of  $m$  ensures that the function

$$F(x, y) = m + \psi(y)(Q(x - \alpha y) - m), \quad 0 \leq x, y \leq 1 \quad (2.38)$$

is positive. Next, we extend  $F(x, y)$  periodically in both variables to the whole plane  $\mathbb{R}^2$ . The periodicity of  $Q(x)$  and (2.37) imply that the extension is smooth. In addition, because of (2.42), it has total mass over the torus equal to one:

$$\int_0^1 \int_0^1 F(x, y)dxdy = 1.$$

In order to map the flow (2.34) to a constant speed flow  $(\alpha, 1)$  moving along the same straight lines, it is natural to attempt to define the transformation  $S : (x, y) \rightarrow (X, Y)$  as

$$X(x, y) = x + \alpha(Y(x, y) - y), \quad Y(x, y) = T(x - \alpha y, y), \quad (2.39)$$

with  $T(x, y)$  defined by

$$T(x, y) = \int_0^y F(x + \alpha z, z)dz.$$

The transformation (2.39) satisfies  $x - \alpha y = X - \alpha Y$ , thus it preserves the flow trajectories. In addition,  $T(x, y)$  is simply the time it takes for the particle to move from the point  $(x - \alpha y, 0)$  to the point  $(x, y)$ . Hence, the particle would move with a constant speed in the new variables, and the speed in the  $y$ -direction would equal to one. However, the transformation (2.39) is not well-defined:  $X(x, y)$  and  $Y(x, y)$  are not 1-periodic (even modulo 1) in  $x$  and  $y$ , so this is not a map of the torus  $\mathbb{T}^2$  to itself. Thus, we modify (2.39) as [35, 45]

$$X(x, y) = x + \alpha(Y(x, y) - y), \quad Y(x, y) = T(x - \alpha y, y) + R(x - \alpha y), \quad (2.40)$$

adding a shift  $R(x - \alpha y)$  that is constant on each trajectory.

We claim that if we choose the 1-periodic function  $R(x)$  that satisfies the homology equation

$$R(\xi + \alpha) - R(\xi) = Q(\xi) - 1, \quad \xi \in \mathbb{S}^1, \quad (2.41)$$

then the map (2.40) is well-defined on  $\mathbb{T}^2$ . Note that for (2.41) to have a measurable solution the function  $Q(\xi)$  should satisfy the normalization [2]

$$\int_0^1 Q(\xi) d\xi = 1. \quad (2.42)$$

The shift in  $x$  is simple: the function  $T(x, y)$  is clearly 1-periodic in  $x$  since  $F(x, y)$  is periodic in  $x$ , thus

$$Y(x + 1, y) = Y(x, y), \quad (2.43)$$

while

$$X(x + 1, y) = 1 + X(x, y). \quad (2.44)$$

To verify what happens under the shift  $y \rightarrow y + 1$ , we first make some preliminary observations. The normalization (2.36) implies that the functions  $F$  and  $Q$  are related by

$$\int_0^1 F(\xi + \alpha z, z) dz = \int_0^1 [m + \psi(z)(Q(\xi) - m)] dz = Q(\xi), \quad 0 \leq \xi \leq 1, \quad (2.45)$$

while for  $\xi \geq 1$  we have

$$\int_0^1 F(\xi + \alpha z, z) dz = Q(\{\xi\}), \quad (2.46)$$

where  $\{\xi\} = \xi - [\xi]$  is the fractional part of  $\xi$ , and  $[\xi]$  is the integer part of  $\xi$ . Furthermore, we have

$$\int_n^{n+1} F(\xi + \alpha z, z) dz = \int_0^1 F(\xi + n\alpha + \alpha z, z) dz = Q(\{\xi + n\alpha\}). \quad (2.47)$$

Now, it follows that

$$\begin{aligned} T(x, y + 1) &= \int_0^{1+y} F(x + \alpha z, z) dz = \int_0^1 F(x + \alpha z, z) dz + \int_1^{y+1} F(x + \alpha z, z) dz \\ &= Q(x) + \int_0^y F(x + \alpha + \alpha z, z) dz = T(x + \alpha, y) + Q(x) \end{aligned} \quad (2.48)$$

because of (2.45). Using this identity, and the homology equation (2.41) for the function  $R$  gives

$$\begin{aligned} Y(x, y + 1) &= T(x - \alpha y - \alpha, y + 1) + R(x - \alpha y - \alpha) \\ &= T(x - \alpha y, y) + Q(x - \alpha y - \alpha) + R(x - \alpha y) - Q(x - \alpha y - \alpha) + 1 \\ &= T(x - \alpha y, y) + R(x - \alpha y) + 1 = Y(x, y) + 1. \end{aligned} \tag{2.49}$$

Finally, for  $X(x, y)$  we have

$$X(x, y + 1) = x + \alpha(Y(x, y + 1) - y - 1) = x + \alpha(Y(x, y) + 1 - y - 1) = x + \alpha(Y(x, y) - y) = X(x, y). \tag{2.50}$$

We conclude from (2.43), (2.44), (2.49) and (2.50) that  $S$  is a well-defined mapping of  $\mathbb{T}^2$  to itself.

A key observation is that solutions  $R(x)$  of the homology equation (2.41) can be very rough even if the function  $Q \in C^\infty(\mathbb{S}^1)$  is smooth. To see that, let us go back to the homology equation

$$R(\xi + \alpha) - R(\xi) = Q(\xi) - 1, \quad \xi \in \mathbb{S}^1. \tag{2.51}$$

Note that it can be solved explicitly using the Fourier transform:

$$R(\xi) = \sum_{n \in \mathbb{Z}} \hat{R}_n e^{2\pi i n \xi}, \tag{2.52}$$

with the Fourier coefficients

$$\hat{R}_n = \frac{\hat{Q}_n}{\exp(2\pi i \alpha n) - 1}. \tag{2.53}$$

The denominators in (2.53) can be dangerously small if  $\alpha n$  can be very close to an integer, that is, if  $\alpha$  is a Liouvillean irrational number. The following Proposition is a particular case of Theorem 4.5 of [27].

**Proposition 2.9** *Let  $\alpha$  be a Liouvillean irrational number. There exists a  $C^\infty(\mathbb{S}^1)$  function  $Q(\xi)$  so that the homology equation (2.41) has a unique (up to an additive constant) measurable solution  $R(\xi) : \mathbb{S}^1 \rightarrow \mathbb{R}$  such that for any  $\lambda \in \mathbb{R} \setminus \{0\}$ , the function  $R_\lambda(\xi) = e^{i\lambda R(\xi)}$  is discontinuous everywhere.*

Without loss of generality we may assume that  $Q(\xi)$  given by Proposition 2.9 is positive – otherwise we choose  $M$  so that  $Q(\xi) + M > 1$  and consider a rescaled function

$$Q_M(\xi) = (M + Q(\xi))/(M + 1).$$

Then, the function

$$R_M(\xi) = \frac{R(\xi)}{M + 1}$$

is the solution of (2.41) with  $Q_M$  in the right side and, of course,  $R_M(\xi)$  has the same set discontinuities as  $R(\xi)$ .

Let us see what happens to the flow (2.34) under the map (2.40):

$$\frac{dx}{dt} = \frac{\alpha}{F(x, y)}, \quad \frac{dy}{dt} = \frac{1}{F(x, y)}, \quad x(0) = x_0, \quad y(0) = y_0. \tag{2.54}$$

Note that

$$x(t) - y(t) = x_0 - \alpha y_0,$$

hence  $Y(t)$  is given by

$$Y(t) = T(x(t) - \alpha y(t), y(t)) + R(x(t) - \alpha y(t)) = T(x_0 - \alpha y_0, y(t)) + R(x_0 - \alpha y_0), \quad (2.55)$$

so that

$$\begin{aligned} \frac{dY}{dt} &= \frac{\partial T(x_0 - \alpha y_0, y(t))}{\partial y} \dot{y}(t) = F(x_0 - \alpha y_0 + \alpha y(t), y(t)) \frac{1}{F(x(t), y(t))} \\ &= F(x(t) - \alpha y(t) + \alpha y(t), y(t)) \frac{1}{F(x(t), y(t))} = 1. \end{aligned} \quad (2.56)$$

On the other hand, for  $X(t)$  we have

$$\frac{dX}{dt} = \dot{x}(t) + \alpha(\dot{Y}(t) - \dot{y}(t)) = \frac{\alpha}{F(x(t), y(t))} + \alpha - \frac{\alpha}{F(x(t), y(t))} = \alpha. \quad (2.57)$$

Therefore, the image of the flow (2.34) under  $S$  is simply the uniform flow:

$$\frac{dX}{dt} = \alpha, \quad \frac{dY}{dt} = 1, \quad (2.58)$$

as we desired. We will denote  $w_{unif} = (\alpha, 1)$ .

Note that  $S$  is invertible with a measurable inverse. Indeed, we have

$$X - \alpha Y = x - \alpha y, \quad (2.59)$$

so that

$$Y = T(X - \alpha Y, y) + R(X - \alpha Y). \quad (2.60)$$

As the function  $F$  is positive, the function  $T(x, y)$  is strictly increasing in  $y$  so that (2.60) has a unique solution  $y(X, Y)$ , and then (2.59) defines  $x(X, Y)$  uniquely.

In addition,  $S$  is measure preserving in the following sense:

$$\int [S^* f](x, y) F(x, y) dx dy = \int f(S(x, y)) F(x, y) dx dy = \int f(X, Y) dX dY \quad (2.61)$$

for any function  $f \in C(\mathbb{T}^2)$ . In order to see that, let us introduce intermediate changes of variables:  $S = S_3 \circ S_2 \circ S_1$ , with  $S_1 : (x, y) \rightarrow (z, y_1)$  with

$$z = x - \alpha y, \quad y_1 = y,$$

followed by  $S_2 : (z, y_1) \rightarrow (Z, y_2)$

$$Z = z, \quad y_2 = T(z, y_1) + R(z),$$

and finally  $S_3 : (Z, y_2) \rightarrow (X, Y)$ , with

$$X = Z + \alpha y_2, \quad Y = y_2.$$

The corresponding Jacobians are:

$$J_1 = J_3 = 1, \quad J_2 = \frac{\partial T}{\partial y_1}(z, y_1) = F(z + \alpha y_1, y_1) = F(x, y).$$

Therefore, the Jacobian of  $S$  is, indeed,

$$J = J_1 J_2 J_3 = F(x, y),$$

hence (2.61) holds and  $S$  is measure-preserving.

Hence,  $S^*$  may be extended as an operator  $L^2(dxdy) \rightarrow L^2(d\mu)$  with the preservation of the corresponding norms. It follows that the unitary evolutions  $U_w^t$  and  $U_{unif}^t$  generated by the flow  $w$  given by (2.54) and the uniform flow  $w_{unif}$ , respectively, are conjugated by means of the unitary transformation

$$S^* : L^2(\mathbb{T}^2, dXdY) \rightarrow L^2(\mathbb{T}^2, d\mu),$$

that is, we have

$$U_{unif}^t = [S^*]^{-1} U_w^t S^*.$$

Therefore,  $U_w^t$  and  $U_{unif}^t$  have the same spectrum:

$$\lambda_{nl} = 2\pi in\alpha + 2\pi il, \quad l, n \in \mathbb{Z}.$$

It also follows that the eigenfunctions of the operator  $U_w$  may be written as

$$\begin{aligned} \psi_{nl}^w(x, y) &= e^{2\pi inX(x,y) + 2\pi ilY(x,y)} = e^{2\pi in(x-\alpha y + \alpha Y(x,y)) + 2\pi ilY(x,y)} \\ &= e^{2\pi in(x-\alpha y)} e^{(2\pi in\alpha + 2\pi il)(T(x-\alpha y, y) + R(x-\alpha y))} = \zeta(x, y) e^{(2\pi in\alpha + 2\pi il)R(x-\alpha y)} \end{aligned} \quad (2.62)$$

with a smooth function  $\zeta(x, y) \in C^\infty([0, 1]^2)$ . Note that the function

$$\zeta(x, y) = e^{2\pi in(x-\alpha y)} e^{(2\pi in\alpha + 2\pi il)T(x-\alpha y, y)}$$

is not periodic in  $y$  even though the function  $\psi_{nl}^w(x, y)$  is periodic. In order to verify that  $\psi_{nl}^w$  are not in  $H^1(\mathbb{T}^2)$  it suffices to check that the function

$$\Theta_\lambda(x, y) = e^{i\lambda R(x-\alpha y)} = R_\lambda(x - \alpha y)$$

is not in  $H^1([0, 1]^2)$  for any real  $\lambda \neq 0$ . Here,  $R_\lambda(s)$  is as defined in Proposition 2.9. Since the function  $\Theta_\lambda(x, y)$  is constant on the lines

$$x - \alpha y = \text{const},$$

if it were in  $H^1([0, 1]^2)$ , it would force the function  $R_\lambda(s)$  to be in  $H^1(\mathbb{S}^1)$  and hence continuous. However,  $R_\lambda$  is discontinuous everywhere according to Proposition 2.9. Therefore, the eigenfunctions  $\psi_{nl}^w$  cannot be in  $H^1(\mathbb{T}^2)$  unless  $n = l = 0$ .

Finally, to obtain an incompressible flow (with respect to the standard Lebesgue measure) with rough eigenfunctions, we introduce a smooth transformation of the torus

$$\bar{S} : (x, y) \rightarrow (p, q)$$

by setting

$$p = \int_0^x \bar{F}(s) ds, \quad q = \frac{1}{\bar{F}(x)} \int_0^y F(x, z) dz, \quad \text{where } \bar{F}(x) = \int_0^1 F(x, z) dz.$$

Note that  $\bar{F}(x)$  is periodic, and

$$p(x+1, y) = \int_0^{x+1} \bar{F}(s) ds = p(x, y) + \int_0^1 \bar{F}(s) ds = p(x, y) + 1.$$

We also have  $q(x+1, y) = q(x, y)$  and

$$q(x, y+1) = \frac{1}{\bar{F}(x)} \int_0^{y+1} F(x, z) dz = q(x, y) + 1.$$

Therefore, indeed,  $\bar{S}$  is a mapping of  $\mathbb{T}^2$  to itself. Since  $F(x, y)$  is positive,  $\bar{S}$  is one-to-one. It is immediate to verify that it maps the measure  $d\mu$  onto the Lebesgue measure  $dpdq$  – the Jacobian of  $\bar{S}$  is  $F(x, y)$ . Hence, the evolution group generated by the image  $u(p, q)$  of the flow  $w(x, y)$  will have the same discrete spectrum as  $U_w$ . In addition, the eigenfunctions  $\psi_{nl}^w$  of  $U_w$  are the images of the eigenfunctions  $\psi_{nl}^u$  of  $u$  under  $\bar{S}^*$ :

$$\psi_{nl}^w(x, y) = (\bar{S}^* \psi_{nl}^u)(x, y) = \psi_{nl}^u(\bar{S}(x, y)).$$

As the functions  $\psi_{nl}^w$  are not in  $H^1(\mathbb{T}^2)$  and the map  $\bar{S}$  is smooth, it follows that all the eigenfunctions of the incompressible flow  $u(p, q)$  are not in  $H^1(\mathbb{T}^2)$ . This finishes the proof of Proposition 2.8.  $\square$

## An abstract criterion for relaxation enhancement

Theorem 2.4 follows from a rather general abstract criterion, which connects us back to the abstract set-up of (1.1). We start with a self-adjoint, positive, unbounded operator  $\Gamma$  with a discrete spectrum on a separable Hilbert space  $H$ . In the setting of Theorem 2.4,  $\Gamma$  corresponds to  $-\Delta$ , with  $H$  the subspace of mean zero functions on  $L^2(\Omega)$ . We denote by

$$0 < \lambda_1 \leq \lambda_2 \leq \dots$$

the eigenvalues of  $\Gamma$ , and by  $e_j$  the corresponding orthonormal eigenvectors forming a basis in  $H$ . The (homogenous) Sobolev space  $H^m(\Gamma)$  associated with  $\Gamma$  is formed by all vectors

$$\psi = \sum_j c_j e_j \in H,$$

such that

$$\|\psi\|_{H^m(\Gamma)}^2 \equiv \sum_j \lambda_j^m |c_j|^2 < \infty.$$

Note that  $H^2(\Gamma)$  is the domain  $D(\Gamma)$  of  $\Gamma$ . The crucial assumption is that  $\lambda_n \rightarrow +\infty$  – this makes the set where the dissipation by  $\Gamma$  is not too large a compact subset of  $H$ .



The analog of the operator  $u \cdot \nabla$  in Theorem 2.4 (or, to be precise, of the self-adjoint operator generating the unitary evolution group  $U^t$  which is equal to  $iu \cdot \nabla$  on  $H^1(\Omega)$ ) is a self-adjoint operator  $L$  such that, for any  $\psi \in H^1(\Gamma)$  and  $t > 0$  we have

$$\|L\psi\|_H \leq C\|\psi\|_{H^1(\Gamma)} \quad \text{and} \quad \|e^{iLt}\psi\|_{H^1(\Gamma)} \leq B(t)\|\psi\|_{H^1(\Gamma)} \quad (2.63)$$

with both the constant  $C$  and the function  $B(t) < \infty$  independent of  $\psi$  and  $B(t) \in L^2_{\text{loc}}[0, \infty)$ . Here  $e^{iLt}$  is the unitary evolution group generated by the self-adjoint operator  $L$ . Consider a solution  $\phi^\varepsilon(t)$  of the rescaled in time version of (1.1):

$$\frac{d}{dt}\phi^\varepsilon(t) = \frac{i}{\varepsilon}L\phi^\varepsilon(t) - \Gamma\phi^\varepsilon(t), \quad \phi^\varepsilon(0) = \phi_0. \quad (2.64)$$

**Theorem 2.10** *Let  $\Gamma$  be a self-adjoint, positive, unbounded operator with a discrete spectrum and let a self-adjoint operator  $L$  satisfy conditions (2.63). Then the following two statements are equivalent:*

- (i) *For any  $\tau, \delta > 0$  there exists  $\varepsilon_0(\tau, \delta)$  such that for any  $0 < \varepsilon < \varepsilon_0(\tau, \delta)$  and any  $\phi_0 \in H$  with  $\|\phi_0\|_H = 1$ , the solution  $\phi^\varepsilon(t)$  of (2.64) satisfies  $\|\phi^\varepsilon(\tau)\|_H^2 < \delta$ .*
- (ii) *The operator  $L$  has no eigenvectors in  $H^1(\Gamma)$ .*

Theorem 2.10 provides a sharp answer to the general question of when a combination of fast unitary evolution and dissipation produces a significantly stronger dissipative effect than dissipation alone. It can be useful in any model describing a physical situation which involves fast unitary dynamics with dissipation (or, equivalently, unitary dynamics with weak dissipation). The proof of Theorem 2.10 uses ideas from quantum dynamics, in particularly the RAGE theorem (see e.g., [10]) describing the evolution of a quantum state belonging to the continuous spectral subspace of a self-adjoint operator.

**Exercise 2.11** Show that neither of the two conditions in (2.63) implies the other.

## The proof of Theorem 2.10

The general idea of the proof is quite straightforward. The dissipation balance for our problem is

$$\frac{1}{2} \frac{d}{dt} (\|\phi^\varepsilon\|_H^2) = -\langle \Gamma\phi^\varepsilon, \phi^\varepsilon \rangle = -\|\phi^\varepsilon(t)\|_{H^1(\Gamma)}^2. \quad (2.65)$$

Therefore, if  $\|\phi\|_{H^1(\Gamma)}$  is large on a time interval  $[t_1, t_2]$  then  $\|\phi^\varepsilon(t)\|_H$  will drop significantly over this time. On the other hand, we will show that if  $\|\phi^\varepsilon(\tau)\|_{H^1(\Gamma)}$  is small at some time  $\tau$  then, as  $L$  does not have  $H^1(\Gamma)$ -eigenfunctions, the free evolution

$$\frac{d\phi^0}{dt} = \frac{i}{\varepsilon}L\phi^0, \quad \phi^0(\tau) = \phi^\varepsilon(\tau), \quad t \geq \tau, \quad (2.66)$$

will make the  $H^1(\Gamma)$ -norm of  $\phi^0$  very large in a time so short that the free evolution is close to the true evolution over this short time interval. This means that the  $H^1(\Gamma)$ -norm of  $\phi^\varepsilon(t)$  will also be very large. Hence, even if the  $H^1(\Gamma)$ -norm of  $\phi^\varepsilon$  drops, it will go back up again very quickly, forcing the dissipation to be large most of the time, and reducing  $\|\phi^\varepsilon(t)\|_H$  very efficiently. Making this argument careful will take us some time, no pun intended. A crucial role is played by the fact that the unit ball in the  $H^1(\Gamma)$ -norm is compact in  $H$ .

## Preliminaries

We first collect some elementary facts and estimates for equation (2.64). For simplicity, we will denote the norm in the Hilbert space  $H$  by  $\|\cdot\|$ , the inner product in  $H$  by  $\langle \cdot, \cdot \rangle$ , the Sobolev spaces  $H^m(\Gamma)$  simply by  $H^m$  and the norms in these Sobolev spaces by  $\|\cdot\|_m$ . We have the following existence and uniqueness result.

**Proposition 2.12** *Assume that for any  $\psi \in H^1$ , we have*

$$\|L\psi\| \leq C\|\psi\|_1. \quad (2.67)$$

*Then for any  $T > 0$ , there exists a unique solution  $\phi(t)$  of the equation*

$$\frac{d\phi(t)}{dt} = (iL - \Gamma)\phi(t), \quad \phi(0) = \phi_0 \in H^1.$$

*This solution satisfies*

$$\phi(t) \in L^2([0, T], H^2) \cap C([0, T], H^1), \quad \dot{\phi}(t) \in L^2([0, T], H). \quad (2.68)$$

**Exercise 2.13** Proposition 2.12 can be proved by standard methods using Galerkin approximations and then establishing uniqueness and regularity. Fill in the details of the argument.

Next we establish a few properties that are more specific to our particular problem. It will be more convenient, in terms of notation, to rescale the time back by the factor  $\varepsilon^{-1}$ , arriving at the equation

$$\frac{d\tilde{\phi}^\varepsilon(t)}{dt} = (iL - \varepsilon\Gamma)\tilde{\phi}^\varepsilon(t), \quad \tilde{\phi}^\varepsilon(0) = \phi_0. \quad (2.69)$$

**Lemma 2.14** *Assume that (2.67) holds, then for any initial data  $\phi_0 \in H$ ,  $\|\phi_0\| = 1$ , the solution  $\tilde{\phi}^\varepsilon(t)$  of (2.69) satisfies*

$$\varepsilon \int_0^\infty \|\tilde{\phi}^\varepsilon(t)\|_1^2 dt \leq \frac{1}{2}. \quad (2.70)$$

**Proof.** Recall that if  $\phi \in H^1(\Gamma)$ , then  $\Gamma\phi \in H^{-1}(\Gamma)$  and  $\langle \Gamma\phi, \phi \rangle = \|\phi\|_1^2$ . The fact that  $L$  is self-adjoint allows us to compute

$$\frac{d}{dt} \|\tilde{\phi}^\varepsilon\|^2 = \langle \tilde{\phi}^\varepsilon, \tilde{\phi}_t^\varepsilon \rangle + \langle \tilde{\phi}_t^\varepsilon, \tilde{\phi}^\varepsilon \rangle = -2\varepsilon \|\tilde{\phi}^\varepsilon\|_1^2. \quad (2.71)$$

Integrating in time and taking into account the normalization of  $\phi_0$ , we obtain (2.70).  $\square$

An immediate consequence of (2.71) is the following result, that we state here as a separate lemma for convenience.

**Lemma 2.15** *Suppose that for all times  $t \in (a, b)$  we have  $\|\tilde{\phi}^\varepsilon(t)\|_1^2 \geq N\|\tilde{\phi}^\varepsilon(t)\|^2$ . Then the following decay estimate holds:*

$$\|\tilde{\phi}^\varepsilon(b)\|^2 \leq e^{-2\varepsilon N(b-a)} \|\tilde{\phi}^\varepsilon(a)\|^2.$$

Next we need an estimate on the growth of the difference between solutions corresponding to  $\varepsilon > 0$  and  $\varepsilon = 0$  in the  $H$ -norm.

**Lemma 2.16** *Assume, in addition to (2.67), that for any  $\psi \in H^1$  and  $t > 0$  we have*

$$\|e^{iLt}\psi\|_1 \leq B(t)\|\psi\|_1 \quad (2.72)$$

for some  $B(t) \in L^2_{\text{loc}}[0, \infty)$ . Let  $\phi^0(t), \phi^\varepsilon(t)$  be solutions of

$$\frac{d\phi^0(t)}{dt} = iL\phi^0(t), \quad \frac{d\tilde{\phi}^\varepsilon(t)}{dt} = (iL - \varepsilon\Gamma)\tilde{\phi}^\varepsilon(t),$$

satisfying  $\phi^0(0) = \phi^\varepsilon(0) = \phi_0 \in H^1$ . Then we have

$$\frac{d}{dt}\|\tilde{\phi}^\varepsilon(t) - \phi^0(t)\|^2 \leq \frac{1}{2}\varepsilon\|\phi^0(t)\|_1^2 \leq \frac{1}{2}\varepsilon B^2(t)\|\phi_0\|_1^2. \quad (2.73)$$

**Proof.** Note that  $\phi^0(t) = e^{iLt}\phi_0$  by definition. Assumption (2.72) says that this unitary evolution is bounded in the  $H^1(\Gamma)$  norm. The regularity guaranteed by conditions (2.67), (2.72) and Proposition 2.12 allows us to multiply the equation

$$\frac{d}{dt}(\tilde{\phi}^\varepsilon - \phi^0) = iL(\tilde{\phi}^\varepsilon - \phi^0) - \varepsilon\Gamma\tilde{\phi}^\varepsilon$$

by  $\tilde{\phi}^\varepsilon - \phi^0$ . We obtain

$$\frac{d}{dt}\|\tilde{\phi}^\varepsilon - \phi^0\|^2 \leq 2\varepsilon(\|\tilde{\phi}^\varepsilon\|_1\|\phi^0\|_1 - \|\tilde{\phi}^\varepsilon\|_1^2) \leq \frac{1}{2}\varepsilon\|\phi^0\|_1^2,$$

which is the first inequality in (2.73). The second inequality follows simply from the assumption (2.72).  $\square$

The following corollary is immediate.

**Corollary 2.17** *Assume that (2.67) and (2.72) are satisfied, and the initial data  $\phi_0 \in H^1$ . Then the solutions  $\tilde{\phi}^\varepsilon(t)$  and  $\phi^0(t)$  defined in Lemma 2.16 satisfy*

$$\|\tilde{\phi}^\varepsilon(t) - \phi^0(t)\|^2 \leq \frac{1}{2}\varepsilon\|\phi_0\|_1^2 \int_0^t B^2(s) ds$$

for any time  $t \leq \tau$ .

### Eigenvectors in $H^1(\Gamma)$ preclude relaxation enhancement

One direction in the proof of Theorem 2.10 is much easier: existence of  $H^1(\Gamma)$  eigenvectors of the operator  $L$  ensures existence of  $\tau_0, \delta_0 > 0$  and  $\phi_0$  with  $\|\phi_0\| = 1$  such that  $\|\phi^\varepsilon(\tau_0)\| > \delta_0$  for all  $\varepsilon$  – that is, if such eigenvectors exist, then the operator  $L$  is not relaxation enhancing.

Assume that the initial condition  $\phi_0 \in H^1$  for

$$\frac{d}{dt}\phi^\varepsilon(t) = \frac{i}{\varepsilon}L\phi^\varepsilon(t) - \Gamma\phi^\varepsilon(t), \quad \phi^\varepsilon(0) = \phi_0 \quad (2.74)$$

is an eigenvector of  $L$  corresponding to an eigenvalue  $E$ , normalized so that  $\|\phi_0\| = 1$ . Take the inner product of (2.74) with  $\phi_0$ . We arrive at

$$\frac{d}{dt}\langle\phi^\varepsilon(t),\phi_0\rangle = \frac{iE}{\varepsilon}\langle\phi^\varepsilon(t),\phi_0\rangle - \langle\Gamma\phi^\varepsilon(t),\phi_0\rangle.$$

This and the assumption  $\phi_0 \in H^1$  lead to

$$\left|\frac{d}{dt}\left(e^{-iEt/\varepsilon}\langle\phi^\varepsilon(t),\phi_0\rangle\right)\right| \leq \frac{1}{2}\left(\|\phi^\varepsilon(t)\|_1^2 + \|\phi_0\|_1^2\right).$$

Note that the value of the expression being differentiated in the left side equals to one at  $t = 0$ . By Lemma 2.14 (with a simple time rescaling) we have

$$\int_0^\infty \|\phi^\varepsilon(t)\|_1^2 dt \leq 1/2.$$

Therefore, for  $t \leq \tau = (2\|\phi_0\|_1^2)^{-1}$  we have  $|\langle\phi^\varepsilon(t),\phi_0\rangle| \geq 1/2$ . Thus,  $\|\phi^\varepsilon(\tau)\| \geq 1/2$ , uniformly in  $\varepsilon$ .  $\square$

### The RAGE theorem and the time spent in high modes

The proof of the other direction in Theorem 2.10, that the absence of eigenfunctions implies relaxation enhancement, is more subtle, and will require some preparation. We switch to the equivalent formulation (2.69), and drop the tilde (hoping that this will not cause any confusion). We need to show that if  $L$  has no  $H^1$  eigenvectors, then for all  $\tau, \delta > 0$  there exists  $\varepsilon_0(\tau, \delta) > 0$  such that if  $\varepsilon < \varepsilon_0$ , then  $\|\phi^\varepsilon(\tau/\varepsilon)\| < \delta$  whenever  $\|\phi_0\| = 1$ . Recall that the main idea of the proof can be naively described as follows. If the operator  $L$  has a purely continuous spectrum or its eigenfunctions are rough, then the  $H^1$ -norm of the free evolution (with  $\varepsilon = 0$ ) is large most of the time. On the other hand, we will show that for small  $\varepsilon$  the full evolution is close to the free evolution for a sufficiently long time. This clearly leads to dissipation enhancement.

Our first task is to get good control of the free evolution  $e^{iLt}$ . The first ingredient that we need to recall is the so-called RAGE theorem, first proved by Ruelle [43] and later generalized by Amrein and Georgescu in [1], and by Enss in [13].

**Theorem 2.18 (RAGE)** *Let  $L$  be a self-adjoint operator in a Hilbert space  $H$ . Let  $P_c$  be the spectral projection on its continuous spectral subspace, and  $\mathcal{C}$  be any compact operator. Then for any  $\phi_0 \in H$ , we have*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \|\mathcal{C}e^{iLt}P_c\phi_0\|^2 dt = 0.$$

The result can be equivalently stated for a unitary operator  $U$ , replacing  $e^{iLt}$  with  $U^t$ . The proof of the RAGE theorem can be found, for example, in [10]. This theorem is a generalization of the following classical theorem by Wiener.

**Theorem 2.19** *Let  $d\mu$  be a finite measure on  $\mathbb{R}$  with the Fourier transform*

$$F(t) = \int_{\mathbb{R}} e^{-ixt} d\mu(x).$$

*Then*

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T |F(t)|^2 dt = \sum_{x \in \mathbb{R}} |\mu(\{x\})|^2. \quad (2.75)$$

Note that the sum in the right side of (2.75) is finite since  $\mu$  is a finite measure.

A direct consequence of the RAGE theorem is the following lemma. Recall that we denote by  $0 < \lambda_1 \leq \lambda_2 \leq \dots$  the eigenvalues of the operator  $\Gamma$  and by  $e_1, e_2, \dots$  the corresponding orthonormal eigenvectors. Let us also denote by  $P_N$  the orthogonal projection on the subspace spanned by the first  $N$  eigenvectors  $e_1, \dots, e_N$  and by

$$S = \{\phi \in H : \|\phi\| = 1\}$$

the unit sphere in  $H$ . The following lemma shows that if the initial data lies in the continuous spectrum of  $L$ , then the  $L$ -evolution will spend most of time in the higher modes of  $\Gamma$ .

**Lemma 2.20** *Let  $K \subset S$  be a compact set. For any  $N, \sigma > 0$ , there exists  $T_c(N, \sigma, K)$  such that for all  $T \geq T_c(N, \sigma, K)$  and any  $\phi \in K$ , we have*

$$\frac{1}{T} \int_0^T \|P_N e^{iLt} P_c \phi\|^2 dt \leq \sigma. \quad (2.76)$$

The key observation of Lemma 2.20 is that the time  $T_c(N, \sigma, K)$  is uniform for all  $\phi \in K$ .

**Proof.** Since  $P_N$  is compact, RAGE theorem implies that for any vector  $\phi \in S$ , there exists a time  $T_c(N, \sigma, \phi)$  that depends on the function  $\phi$  such that (2.76) holds for  $T > T_c(N, \sigma, \phi)$ . To prove the uniformity in  $\phi$ , note that the function

$$f(T, \phi) = \frac{1}{T} \int_0^T \|P_N e^{iLt} P_c \phi\|^2 dt$$

is uniformly continuous on  $S$  for all  $T$  (with constants independent of  $T$ ):

$$\begin{aligned} |f(T, \phi) - f(T, \psi)| &\leq \frac{1}{T} \int_0^T \left| \|P_N e^{iLt} P_c \phi\| - \|P_N e^{iLt} P_c \psi\| \right| (\|P_N e^{iLt} P_c \phi\| + \|P_N e^{iLt} P_c \psi\|) dt \\ &\leq (\|\phi\| + \|\psi\|) \frac{1}{T} \int_0^T \|P_N e^{iLt} P_c (\phi - \psi)\| dt \leq 2\|\phi - \psi\|. \end{aligned}$$

Now, existence of a uniform  $T_c(N, \sigma, K)$  follows from compactness of  $K$  by standard arguments.  $\square$

## $H^1$ -norm of free solutions with rough eigenfunctions

We also need a lemma which controls from below the growth of the  $H^1$  norm of free solutions corresponding to rough eigenfunctions. We denote by  $P_p$  the spectral projection on the pure point spectrum of the operator  $L$ .

**Lemma 2.21** *Assume that not eigenvectors of the operator  $L$  belong to  $H^1(\Gamma)$ . Let  $K \subset S$  be a compact set. Consider the set  $K_1 \equiv \{\phi \in K : \|P_p\phi\| \geq 1/2\}$ . Then for any  $B > 0$  we can find  $N_p(B, K)$  and  $T_p(B, K)$  such that for any  $N \geq N_p(B, K)$ , any  $T \geq T_p(B, K)$  and any  $\phi \in K_1$ , we have*

$$\frac{1}{T} \int_0^T \|P_N e^{iLt} P_p \phi\|_1^2 dt \geq B. \quad (2.77)$$

Note that unlike in (2.76), we have the  $H^1$  norm in (2.77).

**Proof.** The set  $K_1$  may be empty, in which case there is nothing to prove. Otherwise, let us denote by  $E_j$  the eigenvalues of  $L$  (distinct, without repetitions) and by  $Q_j$  the orthogonal projection on the space spanned by the eigenfunctions corresponding to  $E_j$ . First, let us show that for any  $B > 0$  there is  $N(B, K)$  such that for any  $\phi \in K_1$  we have

$$\sum_j \|P_N Q_j \phi\|_1^2 \geq 2B \quad (2.78)$$

if  $N \geq N(B, K)$ . It is clear that for each fixed  $\phi$  with  $P_p \phi \neq 0$  we can find  $N(B, \phi)$  so that (2.78) holds, since by assumption  $Q_j \phi$  does not belong to  $H^1$  whenever  $Q_j \phi \neq 0$ . Assume that  $N(B, \phi)$  cannot be chosen uniformly for  $\phi \in K_1$ . This means that for any  $n$ , there exists  $\phi_n \in K_1$  such that

$$\sum_j \|P_n Q_j \phi_n\|_1^2 < 2B.$$

Since  $K_1$  is compact, we can find a subsequence  $n_l$  such that  $\phi_{n_l}$  converges to  $\bar{\phi} \in K_1$  in  $H$  as  $n_l \rightarrow \infty$ . For any  $N$  and any  $n_{l_1} > N$  we have

$$\sum_j \|P_N Q_j \bar{\phi}\|_1^2 \leq \sum_j \|P_{n_{l_1}} Q_j \bar{\phi}\|_1^2 \leq \liminf_{l \rightarrow \infty} \sum_j \|P_{n_l} Q_j \phi_{n_l}\|_1^2. \quad (2.79)$$

The last inequality follows by Fatou's Lemma from the convergence of  $\phi_{n_l}$  to  $\bar{\phi}$  in  $H$  and the fact that

$$\|P_{n_{l_1}} Q_j \psi\|_1 = \lambda_{n_{l_1}}^{1/2} \|Q_j \psi\| \leq \lambda_{n_{l_1}}^{1/2} \|\psi\|,$$

for any  $n_{l_1}$ . The expression in the right hand side of (2.79) is less than or equal to

$$\liminf_{l \rightarrow \infty} \sum_j \|P_{n_l} Q_j \phi_{n_l}\|_1^2 \leq 2B.$$

Thus, we have

$$\sum_j \|P_N Q_j \bar{\phi}\|_1^2 \leq 2B \text{ for any } N,$$

a contradiction since  $\bar{\phi} \in K_1$ . As a consequence, there exists  $N(B, K)$  so that (2.78) holds for all  $N \geq N(B, K)$  and all  $\phi \in K_1$ .

Next, take  $\phi \in K_1$  and consider

$$\frac{1}{T} \int_0^T \|P_N e^{iLt} P_p \phi\|_1^2 dt = \sum_{j,l} \frac{e^{i(E_j - E_l)T} - 1}{i(E_j - E_l)T} \langle \Gamma P_N Q_j \phi, P_N Q_l \phi \rangle. \quad (2.80)$$

In (2.80), we set

$$\frac{e^{i(E_j - E_l)T} - 1}{i(E_j - E_l)T} \equiv 1 \text{ if } j = l.$$

Notice that the sum in (2.80) converges absolutely. Indeed,

$$P_N Q_j \phi = \sum_{i=1}^N \langle Q_j \phi, e_i \rangle e_i,$$

and  $\langle \Gamma e_i, e_k \rangle = \lambda_i \delta_{ik}$ , therefore

$$\langle \Gamma P_N Q_j \phi, P_N Q_l \phi \rangle = \sum_{i=1}^N \lambda_i \langle Q_j \phi, e_i \rangle \overline{\langle Q_l \phi, e_i \rangle}.$$

Hence, the sum in the right side of (2.80) does not exceed

$$\begin{aligned} \sum_{i=1}^N \lambda_i \sum_{j,l} |\langle Q_j \phi, e_i \rangle \langle Q_l \phi, e_i \rangle| &\leq \lambda_N \sum_{i=1}^N \sum_{j,l} \|Q_j \phi\| \|Q_l \phi\| |\langle Q_j \phi / \|Q_j \phi\|, e_i \rangle \langle Q_l \phi / \|Q_l \phi\|, e_i \rangle| \\ &\leq \lambda_N \sum_{i=1}^N \sum_{j,l} \|Q_l \phi\|^2 |\langle Q_j \phi / \|Q_j \phi\|, e_i \rangle|^2 \leq \lambda_N N. \end{aligned} \quad (2.81)$$

The second step above is obtained using the Cauchy-Schwartz inequality, and the third since  $\|\phi\| = \|e_i\| = 1$ . Then for each fixed  $N$ , it follows from the dominated convergence theorem that the expression in (2.80) converges to

$$\sum_j \|\Gamma^{1/2} P_N Q_j \phi\|^2 = \sum_j \|P_N Q_j \phi\|_1^2$$

as  $T \rightarrow \infty$ .

Now assume  $N \geq N_p(B, K) \equiv N(B, K)$ , so that (2.78) holds. We claim that we can choose  $T_p(B, K)$  so that for any  $T \geq T_p(B, K)$  we have

$$\left| \frac{1}{T} \int_0^T \|P_N e^{iLt} P_p \phi\|_1^2 dt - \sum_j \|P_N Q_j \phi\|_1^2 \right| = \left| \sum_{l \neq j} \frac{e^{i(E_j - E_l)T} - 1}{i(E_j - E_l)T} \langle \Gamma P_N Q_j \phi, P_N Q_l \phi \rangle \right| \leq B \quad (2.82)$$

for all  $\phi \in K_1$ . Indeed, this follows from convergence to zero for each individual  $\phi$  as  $T \rightarrow \infty$ , compactness of  $K_1$ , and uniform continuity of the expression in the middle of (2.82) in  $\phi$  for

each  $T$  (with constants independent of  $T$ ). The latter is proved by estimating the difference of these expressions for  $\phi, \psi \in K_1$  and any  $T$  by

$$\sum_{l \neq j} |\langle \Gamma P_N Q_j \phi, P_N Q_l (\phi - \psi) \rangle| + |\langle \Gamma P_N Q_j (\phi - \psi), P_N Q_l \psi \rangle|,$$

which is then bounded by  $2\lambda_N N \|\phi - \psi\|$  using the trick from (2.81). Combining (2.78) and (2.82) proves the lemma.  $\square$

### Tracking the full dynamics with free evolution

We can now complete the proof of Theorem 2.10. Recall that given any  $\tau, \delta > 0$ , we need to show the existence of  $\varepsilon_0 > 0$  such that if  $\varepsilon < \varepsilon_0$ , then solution of the rescaled problem

$$\frac{d\phi^\varepsilon(t)}{dt} = (iL - \varepsilon\Gamma)\phi^\varepsilon(t), \quad \tilde{\phi}^\varepsilon(0) = \phi_0. \quad (2.83)$$

satisfies  $\|\phi^\varepsilon(\tau/\varepsilon)\| < \delta$  for any initial datum  $\phi_0 \in H$ ,  $\|\phi_0\| = 1$ . Let us outline the idea of the proof. Lemma 2.15 tells us that if the  $H^1$  norm of the solution  $\phi^\varepsilon(t)$  is large, relaxation is happening quickly. If, on the other hand,  $\|\phi^\varepsilon(\tau_0)\|_1^2 \leq \lambda_M \|\phi^\varepsilon(\tau_0)\|^2$ , where  $M$  is to be chosen depending on  $\tau$  and  $\delta$ , then the set of all unit vectors satisfying this inequality is compact, and so we can apply Lemma 2.20 and Lemma 2.21. Using these lemmas, we will show that even if the  $H^1$  norm is small at some moment of time  $\tau_0$ , it will be large on average in some time interval after  $\tau_0$ . Enhanced relaxation will follow.

We now provide the details. Since  $\Gamma$  is an unbounded positive operator with a discrete spectrum, we know that its eigenvalues  $\lambda_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Let us choose  $M$  large enough, so that

$$e^{-\lambda_M \tau/80} < \delta.$$

Define the sets

$$K = \{\phi \in S : \|\phi\|_1^2 \leq \lambda_M\} \subset S,$$

and as before,

$$K_1 = \{\phi \in K : \|P_p \phi\| \geq 1/2\}.$$

It is easy to see that  $K$  is compact. Choose  $N$  so that  $N \geq M$  and  $N \geq N_p(5\lambda_M, K)$  from Lemma 2.21. Define

$$\tau_1 \equiv \max \left\{ T_p(5\lambda_M, K), T_c(N, \frac{\lambda_M}{20\lambda_N}, K) \right\},$$

with  $T_p$  from Lemma 2.21, and  $T_c$  from Lemma 2.20. Finally, choose  $\varepsilon_0 > 0$  so that  $\tau_1 < \tau/2\varepsilon_0$ , and

$$\varepsilon_0 \int_0^{\tau_1} B^2(t) dt \leq \frac{1}{20\lambda_N}, \quad (2.84)$$

where  $B(t)$  is the function from condition (2.72).

Take any  $\varepsilon < \varepsilon_0$ . If we have

$$\|\phi^\varepsilon(s)\|_1^2 \geq \lambda_M \|\phi^\varepsilon(s)\|^2$$



for all  $s \in [0, \tau]$ , then Lemma 2.15 implies that

$$\|\phi^\varepsilon(\tau/\varepsilon)\| \leq e^{-2\lambda_M\tau} \leq \delta,$$

by the choice of  $M$  and we are done. Otherwise, let  $\tau_0$  be the first time in the interval  $[0, \tau/\varepsilon]$  such that

$$\|\phi^\varepsilon(\tau_0)\|_1^2 \leq \lambda_M \|\phi^\varepsilon(\tau_0)\|^2 \quad (2.85)$$

(it may be that  $\tau_0 = 0$ , of course). We claim that the following estimate holds for the decay of  $\|\phi^\varepsilon(t)\|$  on the interval  $[\tau_0, \tau_0 + \tau_1]$ :

$$\|\phi^\varepsilon(\tau_0 + \tau_1)\|^2 \leq e^{-\lambda_M\varepsilon\tau_1/20} \|\phi^\varepsilon(\tau_0)\|^2. \quad (2.86)$$

For the sake of simplicity, we will denote  $\phi^\varepsilon(\tau_0) = \phi_0$ . On the interval  $[\tau_0, \tau_0 + \tau_1]$ , consider the function  $\phi^0(t)$  satisfying

$$\frac{d}{dt}\phi^0(t) = iL\phi^0(t), \quad \phi^0(\tau_0) = \phi_0.$$

Note that by the choice of  $\varepsilon_0$ , (2.84), (2.85), and Corollary 2.17, we have

$$\|\phi^\varepsilon(t) - \phi^0(t)\|^2 \leq \frac{\lambda_M}{40\lambda_N} \|\phi_0\|^2 \quad (2.87)$$

for all  $t \in [\tau_0, \tau_0 + \tau_1]$ . Split

$$\phi^0(t) = \phi_c(t) + \phi_p(t),$$

where  $\phi_{c,p}$  also solve the free equation

$$\frac{d}{dt}\phi_{c,p}(t) = iL\phi_{c,p}(t),$$

but with initial data  $P_c\phi_0$  and  $P_p\phi_0$  at  $t = \tau_0$ , respectively. We will now consider two cases.

*Case I.* Assume that

$$\|P_c\phi_0\|^2 \geq \frac{3}{4}\|\phi_0\|^2,$$

or, equivalently,  $\|P_p\phi_0\|^2 \leq \frac{1}{4}\|\phi_0\|^2$ . Note that since  $\phi_0/\|\phi_0\| \in K$  by the hypothesis, we can apply Lemma 2.20. Our choice of  $\tau_1$  implies that

$$\frac{1}{\tau_1} \int_{\tau_0}^{\tau_0+\tau_1} \|P_N\phi_c(t)\|^2 dt \leq \frac{\lambda_M}{20\lambda_N} \|\phi_0\|^2. \quad (2.88)$$

By elementary considerations,

$$\|(I - P_N)\phi^0(t)\|^2 \geq \frac{1}{2} \|(I - P_N)\phi_c(t)\|^2 - \|(I - P_N)\phi_p(t)\|^2 \geq \frac{1}{2} \|\phi_c(t)\|^2 - \frac{1}{2} \|P_N\phi_c(t)\|^2 - \|\phi_p(t)\|^2.$$

Taking into account the fact that the free evolution  $e^{iLt}$  is unitary,  $\lambda_N \geq \lambda_M$ , our assumptions on  $\|P_{c,p}\phi_0\|$  and (2.88), we obtain

$$\frac{1}{\tau_1} \int_{\tau_0}^{\tau_0+\tau_1} \|(I - P_N)\phi^0(t)\|^2 dt \geq \frac{1}{10} \|\phi_0\|^2. \quad (2.89)$$

Using (2.87), we conclude that

$$\frac{1}{\tau_1} \int_{\tau_0}^{\tau_0+\tau_1} \|(I - P_N)\phi^\varepsilon(t)\|^2 dt \geq \frac{1}{40} \|\phi_0\|^2. \quad (2.90)$$

This estimate implies that

$$\int_{\tau_0}^{\tau_0+\tau_1} \|\phi^\varepsilon(t)\|_1^2 dt \geq \frac{\lambda_N \tau_1}{40} \|\phi_0\|^2. \quad (2.91)$$

Combining (2.91) with (2.71) yields

$$\|\phi^\varepsilon(\tau_0 + \tau_1)\|^2 \leq \left(1 - \frac{\lambda_N \varepsilon \tau_1}{20}\right) \|\phi^\varepsilon(\tau_0)\|^2 \leq e^{-\lambda_N \varepsilon \tau_1 / 20} \|\phi^\varepsilon(\tau_0)\|^2. \quad (2.92)$$

This finishes the proof of (2.86) in the first case since  $\lambda_N \geq \lambda_M$ .

*Case II.* Now suppose that  $\|P_p \phi_0\|^2 \geq \frac{1}{4} \|\phi_0\|^2$ . In this case  $\phi_0 / \|\phi_0\| \in K_1$ , and we can apply Lemma 2.21. In particular, by the choice of  $N$  and  $\tau_1$ , we have

$$\frac{1}{\tau_1} \int_{\tau_0}^{\tau_0+\tau_1} \|P_N \phi_p(t)\|_1^2 dt \geq 5\lambda_M \|\phi_0\|^2. \quad (2.93)$$

Since (2.88) still holds because of our choice of  $\tau_0$  and  $\tau_1$ , it follows that

$$\frac{1}{\tau_1} \int_{\tau_0}^{\tau_0+\tau_1} \|P_N \phi_c(t)\|_1^2 dt \leq \frac{\lambda_M}{20} \|\phi_0\|^2. \quad (2.94)$$

Note that the  $H$ -norm in (2.88) has been replaced in (2.94) by the  $H^1$ -norm at the expense of the factor of  $\lambda_N$ . Together, (2.93) and (2.94) imply

$$\frac{1}{\tau_1} \int_{\tau_0}^{\tau_0+\tau_1} \|P_N \phi^0(t)\|_1^2 dt \geq 2\lambda_M \|\phi_0\|^2. \quad (2.95)$$

Finally, (2.95) and (2.87) give

$$\int_{\tau_0}^{\tau_0+\tau_1} \|P_N \phi^\varepsilon(t)\|_1^2 dt \geq \frac{\lambda_M \tau_1}{2} \|\phi_0\|^2 \quad (2.96)$$

since  $\|P_N \phi^\varepsilon - P_N \phi^0\|_1^2 \leq \lambda_N \|\phi^\varepsilon - \phi^0\|^2$ . As before, (2.96) implies

$$\|\phi^\varepsilon(\tau_0 + \tau_1)\|^2 \leq e^{-\lambda_M \varepsilon \tau_1} \|\phi^\varepsilon(\tau_0)\|^2, \quad (2.97)$$

finishing the proof of (2.86) in the second case.

Summarizing, we see that if  $\|\phi^\varepsilon(\tau_0)\|_1^2 \leq \lambda_M \|\phi^\varepsilon(\tau_0)\|^2$ , then

$$\|\phi^\varepsilon(\tau_0 + \tau_1)\|^2 \leq e^{-\lambda_M \varepsilon \tau_1 / 20} \|\phi^\varepsilon(\tau_0)\|^2. \quad (2.98)$$

On the other hand, for any interval  $I = [a, b]$  such that  $\|\phi^\varepsilon(t)\|_1^2 \geq \lambda_M \|\phi^\varepsilon(t)\|^2$  on  $I$ , we have by Lemma 2.15 that

$$\|\phi^\varepsilon(b)\|^2 \leq e^{-2\lambda_M \varepsilon (b-a)} \|\phi^\varepsilon(a)\|^2. \quad (2.99)$$

Combining all the decay factors gained from (2.98) and (2.99), and using  $\tau_1 < \tau/2\varepsilon$ , we find that there is  $\tau_2 \in [\tau/2\varepsilon, \tau/\varepsilon]$  such that

$$\|\phi^\varepsilon(\tau_2)\|^2 \leq e^{-\lambda_M \varepsilon \tau_2 / 20} \leq e^{-\lambda_M \tau / 40} < \delta^2$$

by our choice of  $M$ . Then (2.71) gives  $\|\phi^\varepsilon(\tau/\varepsilon)\| \leq \|\phi^\varepsilon(\tau_2)\| < \delta$ , finishing the proof of Theorem 2.10.  $\square$

### 3 Dynamics in randomly perturbed problems

We now describe some results for equations of the form (1.1) when the perturbation is random, and the background operator  $L$  is not mixing at all, so mixing comes from the interplay of the simple deterministic dynamics, and randomness, as in toy example in Theorem 1.3.

#### Averaging or homogenization?

Let us briefly discuss the “general philosophy” in this class of problems. Most of the examples we will consider are in a general form

$$\phi_t = iL\phi + \varepsilon V\phi. \quad (3.1)$$

Think of  $L$  as the (linear) deterministic background problem, and  $\varepsilon V$  as an (also linear) random perturbation. As the randomness is weak, its effects will not be seen in the leading order at “short” times  $t \sim O(1)$ , but eventually the random fluctuations will manifest themselves in the leading order, on some long time scale, which we will call  $T_\varepsilon$ . The most basic questions are how large the “non-trivial behavior time”  $T_\varepsilon$  is, in terms of  $\varepsilon$ , that is, “how long do we have to wait to see anything interesting”, and what the effect of randomness on  $\phi$  will be on this time scale, in other words, what is it that we will see at the time  $T_\varepsilon$ . A typical answer is given by the central limit theorem: if the random fluctuations in the physical parameters have mean zero, then  $T_\varepsilon \sim \varepsilon^{-2}$ , and there is some sort of diffusive limit for some observable. A simple example of such behavior is in Theorem 1.2. This answer is not absolutely universal, and depends on the nature of the random fluctuations, so that in systems with slowly decaying spatial correlations, the non-trivial observation time will depend on the physical observable – various observables are affected in a non-trivial way on different time-scales. Thus, a more specific formulation of the above questions is: (1) when will a particular physical observable deviate from its deterministic behavior, and (2) how will it behave after it does that? There is an obvious necessary condition for this problem to be of any “practical” interest: the operator  $L$  has to be non-dissipative, so that the solutions of the unperturbed problem

$$\phi_t = iL\phi \quad (3.2)$$

would not essentially vanish at large times, long before  $T_\varepsilon$ . Otherwise, the random fluctuations would be acting on something that is probably too small to be of interest, though, in principle, the above questions can still be asked. In addition, it greatly helps if the dynamics of the unperturbed problem (3.2) is very well understood (preferably, very simple) so that we know very well what we are perturbing. The no-dissipation requirement is not an issue for a large class of interesting physical systems, but the second requirement is a big restriction on the class of problems we can analyze.

These questions are broadly similar to what is usually understood by random homogenization, even though the random fluctuations are typically not weak in the latter context. For example, the classical random diffusion homogenization problem

$$\begin{aligned}\phi_t &= \nabla \cdot \left( a\left(\frac{x}{\varepsilon}\right) \nabla \phi \right) \\ \phi(0, x) &= \phi_0(x),\end{aligned}\tag{3.3}$$

with a (not weakly) random, or periodic, diffusion matrix  $a(x)$ , comes from the microscopic problem with slowly varying initial data

$$\begin{aligned}\psi_t &= \nabla \cdot (a(x) \nabla \psi) \\ \psi(0, x) &= \phi_0(\varepsilon x),\end{aligned}\tag{3.4}$$

with a simple rescaling

$$\phi(t, x) = \psi\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right).$$

That is, we start the microscopic problem (3.4) with an initial condition which is locally very close to an equilibrium (a constant). Hence, we need to wait for a “long” time, of the order  $t \sim \varepsilon^{-2}$ , to observe a non-trivial deviation from the trivial deterministic dynamics for that underlying equilibrium solution. However, here the time scale  $t \sim \varepsilon^{-2}$  is not really due to the random nature of  $a(x)$  – the same result holds in the periodic case. This is really the law of large numbers rather than a central limit theorem type of a result.

The commonality between these two problems seems to be that in order to get a macroscopic description at the “large” time  $T_\varepsilon$ , one needs to know quite precisely the dynamics of the unperturbed system: be it (3.2) in the weak randomness case, or the trivial dynamics of the constant state in (3.4). The techniques in the analysis are, however, quite different. There is one important “poetic” difference between the “weak randomness” and homogenization set-ups. The non-trivial observation time  $T_\varepsilon$  is dictated by the medium in the former but by the initial condition in the latter, and is therefore beyond our control in the first problem but is controlled by the input in the second. Another obvious difference is that the underlying dynamics in the homogenization problem is trivial until the time  $t \sim \varepsilon^{-2}$ , while in the “weak randomness” case the dynamics of (3.2) is (hopefully) simple but non-trivial, and will have to be taken into account.

### The fast/slow dynamics decomposition

A useful way to look at the general dynamics (3.1)

$$\phi_t = iL\phi + \varepsilon V\phi\tag{3.5}$$

is to factor out the fast dynamics, making the interaction of the fast dynamics of  $L$  and the slow dynamics of  $V$  more visible. This is what leads to a non-trivial behavior of  $\phi$ . Let us rescale the time variable  $t \rightarrow t/\varepsilon^2$ , so that (3.5) becomes

$$\phi_t = \frac{i}{\varepsilon^2} L\phi + \frac{1}{\varepsilon} V\phi. \quad (3.6)$$

In some sense, the leading order term is

$$\tilde{\psi}_t = \frac{i}{\varepsilon^2} L\tilde{\psi}, \quad (3.7)$$

that is,

$$\tilde{\psi}(t) = e^{iLt/\varepsilon^2} \psi_0. \quad (3.8)$$

As the operator  $L$  is self-adjoint (or otherwise non-dissipative), one would expect  $\psi(t)$  to converge weakly to zero as  $\varepsilon \rightarrow 0$  but not strongly. It is convenient to write

$$\phi(t) = e^{iLt/\varepsilon^2} \psi(t). \quad (3.9)$$

The function  $\psi(t)$  satisfies

$$\psi_t = \frac{1}{\varepsilon} B\left(\frac{t}{\varepsilon^2}\right) \psi, \quad (3.10)$$

with the operator

$$B(t) = e^{-iLt} V e^{iLt}. \quad (3.11)$$

When the operators  $L$  and  $V$  do not commute, their interaction in (3.11) often leads to a non-trivial limit for the function  $\psi(t)$ . Studying the "compensated" solution  $\psi(t)$  is often an instructive approach.

## Particle motion in a random velocity field

Maybe the simplest example of a non-trivial weakly random dynamical system is particle motion in a weekly random velocity field:

$$\frac{dX(t)}{dt} = \bar{u} + \varepsilon v(X(t)), \quad X(0) = x. \quad (3.12)$$

The corresponding first order linear PDE is

$$\phi_t + (\bar{u} + \varepsilon v(x)) \cdot \nabla \phi = 0. \quad (3.13)$$

Here,  $\bar{u}$  is a non-zero mean flow, and  $v(x)$  is a mean-zero random fluctuation. We will see that appropriately re-centered particle position converges to a diffusive limit. More precisely, the process

$$Y_\varepsilon(t) = X_\varepsilon\left(\frac{t}{\varepsilon^2}\right) - \bar{u} \frac{t}{\varepsilon^2}, \quad (3.14)$$

converges to a Brownian motion. Notice that if  $\bar{u} = 0$  this result can not hold without some extra miracles or hard work – the particle would be stuck close to the zeros of  $v(x)$ . Thus, the diffusive limit is a combination of the background dynamics (with  $\varepsilon = 0$ ) and the random fluctuation that together lead to the diffusive limit, a recurrent theme in many of such problems.

## Perturbed particle motion in a 2D Hamiltonian flow

The next example is a particle moving in weakly perturbed two-dimensional Hamiltonian flow, described by a stochastic differential equation

$$dX_t = -\nabla^\perp H(X_t)dt + \varepsilon dB_t. \quad (3.15)$$

Here,  $X_t = (x_t, y_t)$  is the particle position,  $B_t$  is the standard two-dimensional Brownian motion, and  $H(x)$  is a background Hamiltonian. We use the notation  $\nabla^\perp H(x, y) = (H_y, -H_x)$ . Think of  $H(x, y)$  as a function growing at infinity, with finitely many non-degenerate critical points, and there is at most one critical point on each level set of  $H$ . If one would like to avoid the language of the stochastic differential equations, the corresponding ODE analog is

$$\frac{dX}{dt} = -\nabla^\perp H(X_t)dt + \varepsilon v(t), \quad (3.16)$$

with a random flow  $v(t)$ . The results would be very similar to those for (3.15).

If  $\varepsilon = 0$ , then the Hamiltonian is constant along the flow trajectories:  $H(X(t)) = H(X(0))$ . However, as can be immediately seen both from (3.15) and (3.16), when  $\varepsilon > 0$ , the Hamiltonian is not preserved by the perturbed dynamics, and is slowly changing in time. One has to wait for a long time for the evolution of the Hamiltonian to become non-trivial, and we will show that the process

$$h_\varepsilon(t) = H\left(X\left(\frac{t}{\varepsilon^2}\right)\right), \quad (3.17)$$

converges to a certain diffusion process on the Reeb graph of the function  $H(x)$ .

If  $H(x, y)$  has many critical points on the same level set, such as the function

$$H(x, y) = \cos x \cos y,$$

then the dynamics is quite different – particles perform a diffusion process between various cells (invariant regions) of the flow.

## Particle in a weakly random Hamiltonian flow

The next example is a particle moving in a weakly random Hamiltonian field. The simplest example is a classical particle in a weak random potential force field:

$$\frac{dX(t)}{dt} = K(t), \quad \frac{dK(t)}{dt} = -\varepsilon \nabla V(X(t)), \quad (3.18)$$

in dimensions  $d \geq 3$ . This is the motion corresponding to the classical Hamiltonian

$$H_\varepsilon(x, k) = \frac{k^2}{2} + \varepsilon V(x). \quad (3.19)$$

One may consider more general Hamiltonians of the form

$$H_\varepsilon(x, k) = H_0(k) + \varepsilon H_1(x, k), \quad (3.20)$$

with a deterministic background Hamiltonian  $H_0(k)$  and a random perturbation  $H_1(x, k)$ , but we will not do this here.

As in (3.15) and (3.16), if  $\varepsilon = 0$  in (3.18), the Hamiltonian  $H_0(k) = k^2/2$  is preserved along the trajectories, and the process  $K(t)$  stays on the sphere  $|K(t)| = |K(0)|$ . Of course, when  $\varepsilon = 0$  in (3.18) we actually have  $K(t) = K(0)$  but that is beside the point. Now, unlike for (3.15) or (3.16), the perturbation of the Hamiltonian  $H_0(k)$  in (3.20) is time-independent and Hamiltonian, thus the full Hamiltonian  $H_\varepsilon(x, k)$  is preserved by the dynamics for  $\varepsilon > 0$ . Hence, you do not expect to see the diffusion across the level sets of the Hamiltonian  $H_0(k)$  as the limit process, which is what we obtained for (3.15). Rather, as  $K(t)$  is a process on the level sets of  $H_\varepsilon(x, k)$ , in the limit  $\varepsilon \rightarrow 0$ , we expect it to converge to a diffusion on the level set  $H_0(K(t)) = H_0(K(0))$ . For the classical mechanics Hamiltonian (3.19), this is simply the sphere  $|K(t)| = |K(0)|$ .

Thus, we will consider the process

$$K_\varepsilon(t) = K\left(\frac{t}{\varepsilon^2}\right), \quad (3.21)$$

and we will see that it converges to a Brownian motion  $\bar{K}(t)$  on the sphere  $|K(t)| = |K(0)|$ , with a certain diffusivity matrix  $D(k)$ . If we were to consider a more general background Hamiltonian than  $H_0(k) = k^2/2$ , we would see in the limit a diffusion process on its level set rather than on the sphere  $\{|k| = \text{const}\}$ , which is a level set of  $H_0(k) = k^2/2$ . As we are dealing with long times, and  $X(t) = K(t)$  is not small, by the times of the order  $O(\varepsilon^{-2})$  the particle will be "very far" in space. Accordingly, the spatial component needs to be re-scaled: the process

$$X_\varepsilon(t) = \varepsilon^2 X\left(\frac{t}{\varepsilon^2}\right) \quad (3.22)$$

converges to

$$\bar{X}(t) = \int_0^t \bar{K}(s) ds. \quad (3.23)$$

In terms of PDEs, the Hamiltonian dynamics corresponds to the Liouville equation

$$\phi_t + k \cdot \nabla_x \phi - \varepsilon \nabla V(x) \cdot \nabla_k \phi = 0. \quad (3.24)$$

The above result says that

$$\phi_\varepsilon(t, x) = \mathbb{E}\left(\phi\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon^2}, k\right)\right)$$

converges to  $\bar{\phi}(t, x, k)$ , solution of

$$\phi_t + k \cdot \nabla_x \phi - \frac{\partial}{\partial k_j} \left( D_{jm}(k) \frac{\partial \bar{\phi}}{\partial k_m} \right) = 0. \quad (3.25)$$

Note, again, that the background dynamics is very important here: if we would consider the (very artificial) Hamiltonian

$$H_\varepsilon(x, k) = \varepsilon V(x),$$

with a random function  $V(x)$ , then the dynamics would be

$$\frac{dX}{dt} = 0, \quad \frac{dK}{dt} = \varepsilon \nabla V(X).$$

Its solution is trivial:  $X(t) = X(0)$ , and

$$K(t) = \varepsilon \nabla V(X(0))t.$$

Thus, the long time limit of  $K(t)$  would not be diffusive at all.

## The Schrödinger equation

The last example we will consider in this family is the Schrödinger equation with a weak random potential:

$$i\phi_t + \frac{1}{2}\Delta\phi - \varepsilon V(x)\phi = 0. \quad (3.26)$$

One may think of this equation as an example of the infinite dimensional Hamiltonian dynamics with the Hamiltonian

$$H[\psi] = \frac{1}{2} \int (|\nabla\psi(x)|^2 + \varepsilon V(x)\psi^2(x)) dx,$$

and view it as an infinite-dimensional version of (3.18). However, we will use completely different tools for the analysis here, that are on the surface much less natural. The "good" tools are missing at the moment.

## The Liouville-Green (Wentzel-Kramers-Brillouin) approximation

One natural way to look at weakly perturbed non-dissipative problems is provided by the WKB approximation developed originally, and independently from each other, by Liouville and Green in 1837. The idea is very simple, and we illustrate it for the Schrödinger equation with a slowly varying potential

$$i\psi_t + \frac{1}{2}\Delta\psi - V(\varepsilon x)\psi = 0. \quad (3.27)$$

Rescaling time and space as  $t \rightarrow t/\varepsilon$  and  $x \rightarrow x/\varepsilon$  leads to

$$i\varepsilon\psi_t + \frac{\varepsilon^2}{2}\Delta\psi - V(x)\psi = 0. \quad (3.28)$$

Let us seek an oscillatory solution of the wave equation in the form

$$\psi(t, x) = A(t, x)e^{iS(t, x)/\varepsilon}.$$

Here,  $A(t, x)$  is the wave amplitude and  $S(t, x)$  is its phase. Note that the amplitude and the phase vary on the macroscopic scale (independent of  $\varepsilon$ ) but, for  $\varepsilon \ll 1$ , the function  $\psi(t, x)$  oscillates on the scale  $\varepsilon$ , which is the original microscopic scale in (3.27). Inserting this ansatz for  $\psi(t, x)$  into (3.28), we get, in the leading order the eikonal equation

$$S_t + \frac{|\nabla S|^2}{2} + V(x) = 0. \quad (3.29)$$

The next order of powers in  $\varepsilon$  gives the amplitude equation

$$(|A|^2)_t + \nabla \cdot (|A|^2 \nabla S) = 0. \quad (3.30)$$

Let us make the following observation: consider the measure

$$W(t, x, \xi) = |A(t, x)|^2 \delta(\xi - \nabla S(t, x)),$$



with  $A(t, x)$  and  $S(t, x)$  which solve the eikonal equation (3.29), and the amplitude equation (3.30). Then, a direct computation shows that  $W(t, x, \xi)$  satisfies the Liouville equation of the classical mechanics

$$\frac{\partial W}{\partial t} + \xi \cdot \nabla_x W - \nabla_x V \cdot \nabla_\xi W = 0. \quad (3.31)$$

The corresponding characteristics are

$$\frac{dX}{dt} = K, \quad \frac{dK}{dt} = -\nabla V(X),$$

which is nothing but Newton's equations of motion. This is probably the simplest connection between the classical and quantum mechanics. We can think of  $W(t, x, \xi)$  as the phase space energy density of the quantum particle: note that

$$\rho(t, x) := |A(t, x)|^2 = \int W(t, x, \xi) d\xi,$$

and if we think of (3.30) as the fluid equation

$$\rho_t + \nabla \cdot (\rho v) = 0,$$

with the velocity  $v(t, x) = \nabla S(t, x)$ , then the support of  $W(t, x, \xi)$  is exactly at  $\xi = v(t, x)$ .

### A caustic: seductiveness of the kinetic approach

Let us now explain how we can see the formation of a caustic in terms of the Liouville equation. To be concrete and simple, consider the Schrödinger equation with  $V = 0$ , and let us take the initial phase as  $S_0(x) = -x^2/2$  with a smooth initial amplitude  $A_0(x)$ . Solution of the eikonal equation

$$S_t + \frac{1}{2} S_x^2 = 0,$$

is given explicitly by  $S(t, x) = -x^2/(2(1-t))$  – a caustic appears at  $t = 1$ . The corresponding characteristics for the amplitude equation satisfy

$$\dot{X} = -\frac{X}{1-t}, \quad X(0) = x$$

and are given by  $X(t) = x(1-t)$  – hence all characteristics arrive to the point  $x = 0$  at the time  $t = 1$ . This is the caustic point. At this time the geometric optics approximation breaks down and is no longer valid.

On the other hand, the Liouville equation (3.31) is linear, and its solutions should not break down. Let us see what happens: as  $V = 0$ , the Liouville equation is

$$W_t + k \cdot \nabla_x W = 0, \quad W(0, x, k) = W_0(x, k). \quad (3.32)$$

Its solution is  $W(t, x, k) = W_0(x - kt, k)$  and clearly exists for all time. Since the initial phase is  $S_0(x) = -x^2/2$ , at  $t = 0$  we have

$$W_0(x, k) = |A_0(x)|^2 \delta(k + x),$$

so that the solution of (3.32) is

$$W(t, x, k) = |A_0(x - kt)|^2 \delta(k + x - kt).$$

This means that at the time  $t = 1$  the solution

$$W(t = 1, x, k) = |A_0(x - k)|^2 \delta(x)$$

is no longer singular in momenta  $k$  but rather in space, being concentrated at  $x = 0$ . On the other hand, the solution of the Liouville equation exists beyond this time, unlike that of the eikonal equation, and from the Liouville point of view nothing particularly dramatic happens at  $t = 0$ .

Anticipating the need to study problems in a random medium, a natural question then is the following: suppose that the initial condition for the eikonal equation is  $S_0(x) = k_0 \cdot x$  – this is a plane wave, and the medium is weakly random. How long will it take for the solution to form a caustic? If it happens very quickly, then the geometric optics ansatz in a random medium can not be used for too long – this is a very important point, as it gives the need to very interesting mathematics!

### A stochastic caustic

Let us now make a jump of more than a hundred years and look at the question of when a caustic would appear in a weakly random medium. This is the work of Kulkarny and White in 1982 in 2D, and White in 1984 in 3D. Assume that the sound speed in the medium is weakly fluctuating: it has the form

$$c(x) = 1 + \varepsilon \mu(x),$$

where  $\mu(x)$  is a mean zero random process, stationary in space, and  $\varepsilon \ll 1$  is a small parameter measuring the strength of the fluctuations. Typically, a mean zero random fluctuation of size  $\varepsilon$  will produce a non-trivial effect on a time scale  $T_\varepsilon \sim \varepsilon^{-2}$  – this is the central limit theorem, bringing us back to the 18th century (de Moivre (1667-1754) stated it in 1733). White (and with Kulkarny) considered the ray equations in such medium, and in 2D it was shown (after lengthy calculations) that on the time scale  $t \sim \varepsilon^{-2/3}$  the ray curvature behaves as the solution of the stochastic differential equation

$$dZ = -Z^2 dt + dB_t.$$

Solutions of this stochastic differential equation blow up in a finite (but random) time, almost surely. This means that (in the original time variables), a caustic will form at a random time of the order  $T \sim O(\varepsilon^{-2/3})$  which is much shorter than the "interesting" central limit time scale  $O(\varepsilon^{-2})$ . Thus, a caustic happens relatively quickly, before one would expect the macroscopic observables to be affected. This time is even much shorter than the naive "non-trivial effect" time  $O(\varepsilon^{-1})$ . Thus, a straightforward geometric optics ansatz in a weakly random medium can be expected to hold only for times which are much shorter than times of "real" interest. A different description has to be used if we want to understand what happens on longer time scales, and this is accomplished by the kinetic theory.

## The kinetic models of wave propagation in heterogenous media

We will not discuss much about the kinetic models in these notes but some comments are in order.

We have at least three basic length scales in wave propagation problems:  $L$  – the overall propagation distance from the source to our observation point,  $\lambda$  – the scale on which the initial source is localized, and  $l_c$  – the typical scale of variations of the medium. The latter two scales are often not defined in a precise way, and we will explain later what exactly we mean by them. Generally, we will be interested in the situations when the propagation distance  $L$  is much larger than both  $\lambda$  and  $l_c$ , giving even small variations in the microstructure a chance to have a strong effect on the macroscopic features of the wave. This brings us to the next important parameter:  $\varepsilon \ll 1$  is the relative strength of the microscopic fluctuations in the parameters of the medium. We will always assume that this strength is small.

Note that  $\lambda$  can often be chosen – this is, essentially, the wave length of the probing signal, and we may modify it to suit a particular application. The propagation distance  $L$  can also be chosen – this is the observation scale, that the observer can often (but not always) control. On the other hand, the scale of the medium variations  $l_c$  is typically outside of our control – the medium is usually given to us, and we can not modify it. The same is true for  $\varepsilon$  – this parameter is a feature of the medium and not of a particular setting of the physical experiment. A typical question we will be facing is “Given the strength of the microscopic fluctuations  $\varepsilon \ll 1$ , and the medium variations scale  $l_c$ , as well as the probing signal wave length  $\lambda$ , how large can the propagation distance  $L$  be, so that we can still have an effective macroscopic model for the wave, and what will that model be?” The answer will, broadly speaking, depend on two factors: the relative size of  $l_c$  and  $\lambda$ , and on the statistics of the small scale fluctuations of the medium. The three regimes we would ideally describe in some detail are random geometric optics, radiative transport, and random homogenization. It is not very likely we will have time for all of them in these lectures.

The macroscopic models are often written in terms of the energy density in the phase space. The underlying premise is that the multiple scattering of the waves by the medium inhomogeneities will create “waves going in all directions at each point”. Thus, the primary object is now not the wave field but the (empirical) wave energy density  $W(t, x, \xi)$  at the time  $t > 0$ , at a position  $x \in \mathbb{R}^n$ , with the wave vector  $\xi \in \mathbb{R}^n$ . The wave energy evolution is described in terms of the kinetic equation

$$\frac{\partial W(t, x, \xi)}{\partial t} + \nabla_{\xi} \omega(\xi) \cdot \nabla_x W(t, x, \xi) = \mathcal{L}_{sc} W(t, x, \xi). \quad (3.33)$$

Here,  $\omega(\xi)$  is the dispersion relation of the wave and depends on the particular type of the wave. The left side of (3.33) has nothing to do with the inhomogeneities of the medium<sup>1</sup> and represents the free transport of the wave energy along the characteristics  $\dot{X} = \nabla_{\xi} \omega(\xi)$  (which are straight lines). On the other hand, the scattering operator  $\mathcal{L}_{sc}$  incorporates the macroscopic effects of the small scale inhomogeneities, and involves the possibility for waves to scatter in different directions at a given point. Its exact form depends on the physical regime of the problem, and the task of modeling is typically two-fold: to find the relation of

---

<sup>1</sup>Strictly speaking, this statement assumes that the fluctuations are sufficiently weak so that they do not modify the wave dispersion relation.

the phase space energy density  $W(t, x, \xi)$  to the underlying wave field that can be directly measured (pressure, electric and magnetic fields, elastic displacements, and so on, depending on the problem), and to identify the scattering operator  $\mathcal{L}_{sc}$  for a particular physical problem.

## 4 A limit theorem for a particle in a random flow

We now go just one level up in difficulty compared to the toy problem (1.33), and consider the same background dynamics: a uniform flow but instead of somewhat artificially adding a random potential, we add a random fluctuation to the flow itself. That is, we look at a particle moving in a random flow with a large mean:

$$\frac{dX}{dt} = \bar{u} + \varepsilon v(X), \quad (4.1)$$

with the corresponding first order PDE

$$\phi_t + (\bar{u} + \varepsilon v(x)) \cdot \nabla \phi = 0. \quad (4.2)$$

Rescaling the time variable  $t \rightarrow t/\varepsilon^2$  gives

$$\phi_t + \frac{1}{\varepsilon^2}(\bar{u} + \varepsilon v(x)) \cdot \nabla \phi = 0. \quad (4.3)$$

As in the previous example, the background dynamics is very simple:

$$\bar{\phi}_t + \frac{1}{\varepsilon^2} \bar{u} \cdot \nabla \bar{\phi} = 0, \quad (4.4)$$

or

$$\bar{\phi}(t, x) = \phi_0(x - \bar{u} \frac{t}{\varepsilon^2}).$$

Accordingly, we take out the background dynamics, as in the fast/slow dynamics decomposition (3.9):

$$\phi(t, x) = \psi(t, x - \bar{u} \frac{t}{\varepsilon^2}). \quad (4.5)$$

The function  $\psi(t, x)$  satisfies

$$\psi_t + \frac{1}{\varepsilon} v(x + \bar{u} \frac{t}{\varepsilon^2}) \cdot \nabla \psi = 0. \quad (4.6)$$

A convenient approach to this problem is via understanding the general problem of the behavior of a particle in a rapidly varying in time random flow:

$$\dot{X} = \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, X\right), \quad X(0) = x, \quad (4.7)$$

with a random field  $V$  when  $\varepsilon \ll 1$ . When the random flow is spatially uniform,  $V = V(t)$ , then

$$X(t) = \frac{1}{\varepsilon} \int_0^t V\left(\frac{s}{\varepsilon^2}\right) ds = \varepsilon \int_0^{t/\varepsilon^2} V(s) ds. \quad (4.8)$$

Thus,  $X(t)$  converges in law to a Brownian motion, according to Theorem 1.2. In the general case, when  $V(t, x)$  is not spatially uniform, this question goes back to the papers by Khasminskii [32] from the 60's with subsequent contributions by various authors: without any attempt at completeness we mention the work of Papanicolaou and Kohler [42], and Kesten and Papanicolaou [30]. We present a version of the limit theorem due to T. Komorowski [36].

Let us explain where the scaling in (4.7) comes from – why the time dependence of the particle velocity is "fast" and the space-dependence is "slow". To see that let us start with a dynamical system

$$\frac{dY}{dT} = v_0 V\left(\frac{T}{t_0}, \frac{Y}{x_0}\right)$$

with a random time-dependent field  $V(s, x)$  and introduce non-dimensional space-time variables  $X = Y/x_0$ ,  $s = T/t_0$ :

$$\frac{dX}{ds} = \varepsilon V(s, X), \quad \varepsilon = \frac{v_0 t_0}{x_0}.$$

Let us now assume that  $\varepsilon \ll 1$  is a small parameter – physically, this means that the time it takes the particle to pass one spatial correlation length is much larger than the correlation time of the random fluctuations. Therefore, in this regime the temporal randomness of  $V(s, x)$  "dominates" the spatial variations. If we now introduce a slow time  $t$  so that  $t = \varepsilon^2 s$ , then in the variables  $(t, x)$  the particle obeys (4.7). The limit  $\varepsilon \rightarrow 0$  now corresponds to observing the particle at times much larger than the correlation time of the random fluctuations and on the spatial scale of the order of the correlation length of the medium.

The first order equation corresponding to (4.7) is

$$\frac{\partial \phi}{\partial t} - \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, x\right) \cdot \nabla \phi = 0, \quad \phi(0, x) = \phi_0(x). \quad (4.9)$$

Its solution is  $\phi(t, x) = \phi_0(X(t; x))$ , where  $X(t; x)$  is the solution of (4.7).

When does one expect the trajectories of (4.7) to behave diffusively? First of all,  $V$  has to have mean zero so that the mean displacement would not be clearly biased. Second,  $V$  should "mix things around" which means that the flow should be incompressible. It helps if dynamics at "far away" points is nearly independent: this is formalized by the mixing assumption below that eliminates the memory effect. Finally, there should be no distinguished times – this requires stationarity of  $V$  in time.

### Assumptions on the random field

We now list the formal assumptions on the random field that we will use to prove the diffusive limit.

**Stationarity.** The random field  $V(t, x)$  is strictly stationary in time and space. This means that for any  $t_1, t_2, \dots, t_m \in \mathbb{R}$ ,  $x_1, \dots, x_m \in \mathbb{R}^n$ , and each  $h \in \mathbb{R}$  and  $y \in \mathbb{R}^n$  the joint distribution of

$$V(t_1 + h, x + y), V(t_2 + h, x + y), \dots, V(t_m + h, x + y)$$

is the same as that of

$$V(t_1, x), V(t_2, x), \dots, V(t_m, x).$$

We will denote by  $R_{nm}(t, x)$  the two-point correlation tensor of  $V(t, x)$ :

$$R_{nm}(t, x) = \mathbb{E} \{V_n(s, y)V_m(t + s, y + x)\}. \quad (4.10)$$

The spatial stationarity of  $V(t, x)$  is not necessary but it allows to simplify a few expressions in what follows. It can, however, be dropped and we adopt it here simply for convenience. On the other hand, stationarity in time is essential for the limit theorem.

**Mixing: attempt 1.** We will assume that the field  $V(t, x)$  is mixing. Roughly speaking, this means that the values of  $V(t, x)$  are sufficiently independent at different times. One possible way to formulate this assumption is to say that  $V(t, x)$  and  $V(t + h, y)$  are nearly independent if the time increment  $h$  is large enough, no matter what  $x$  and  $y$  are. This is formalized in terms of the  $\sigma$ -algebras  $\tilde{\mathcal{V}}_a^b$  generated by the sets of the form

$$\{\omega : V(t, x, \omega) \in A\},$$

where  $a \leq t \leq b$ ,  $x \in \mathbb{R}^n$ , and  $A$  is a Borel set in  $\mathbb{R}^n$ . The corresponding mixing coefficient is

$$\tilde{\beta}(h) = \sup_{t \geq 0} \sup_{A \in \tilde{\mathcal{V}}_{t+h}^\infty, B \in \tilde{\mathcal{V}}_0^t} \frac{|P(A \cap B) - P(A)P(B)|}{P(B)}.$$

The mixing assumption would be that for any  $m \geq 0$  the mixing coefficient satisfies

$$h^m \tilde{\beta}(h) \leq C_m \text{ for all } h \geq 0.$$

Heuristically, this means that events in  $\tilde{\mathcal{V}}_0^t$  and  $\tilde{\mathcal{V}}_{t+h}$  are basically independent.

The problem with this definition of mixing is that it would not apply to random fields of the form

$$V(t, x) = v(x + \bar{u}t), \quad (4.11)$$

with  $\bar{u} \neq 0$ , and a random field  $v(x)$ , which is our original motivation. Indeed, for such  $V(t, x)$ , we have

$$V(t, x) = V(t + h, y - h\bar{u}), \quad (4.12)$$

for all  $x$  and  $y$  such that  $y = x - h\bar{u}$ . Thus, the assumption that  $V(t, x)$  and  $V(t + h, y)$  are nearly independent for all  $x, y \in \mathbb{R}^n$  can not hold for  $V(t, x)$  given by (4.11). On the formal level, this is reflected in the fact that all  $\sigma$ -algebras  $\tilde{\mathcal{V}}_a^b$  are the same in this case, no matter what  $a, b$  are.

**Mixing: attempt 2.** Thus, to allow for random fields as in (4.11), we need to modify the definition of the mixing coefficient. One natural way is to assume that  $V(t, x)$  and  $V(t + h, y)$  are nearly independent only for "nearby"  $x$  and  $y$  if  $h$  is large. Identity (4.12) hints that it suffices to have "near independence" of  $V(t, x)$  and  $V(t + h, y)$  for  $x$  and  $y$  such that  $|x - y| \ll h$ .

To make this formal, we fix  $C > 0$  and, given a time interval  $I_{ab} = a \leq t \leq b$ , consider the sets

$$S_a^b = \{(t, x) : a \leq t \leq b, |x| \leq C(1 + \sqrt{t})\}.$$

We denote by  $\mathcal{V}_a^b$  the  $\sigma$ -algebra generated by the sets of the form  $\{\omega : V(t, x, \omega) \in A\}$ , with  $(t, x) \in S_a^b$ , and  $A$  is a Borel set in  $\mathbb{R}^n$ . The mixing coefficient is now defined as

$$\beta(h) = \sup_{0 \leq t \leq 1+h^{3/2}} \sup_{A \in \mathcal{V}_{t+h}^\infty, B \in \mathcal{V}_0^t} \frac{|P(A \cap B) - P(A)P(B)|}{P(B)}, \quad (4.13)$$

and our mixing assumption is that it satisfies

$$h^m \beta(h) \leq C_m \text{ for all } h \geq 0.$$

Let us see why this mixing condition is reasonable for velocity fields of the form (4.11), and why we have the restriction

$$0 \leq t \leq 1 + h^{3/2}$$

in the supremum in (4.13). Consider two space-time points  $(s_1, x_1) \in S_0^t$ , and  $(s_2, x_2) \in S_{t+h}^\infty$ , with  $0 \leq t \leq 1 + h^{3/2}$ , then

$$V(s_1, x_1) = v(x_1 + s_1 \bar{u}), \quad V(s_2, x_2) = v(x_2 + s_2 \bar{u}), \quad (4.14)$$

and

$$\begin{aligned} d &= |x_2 + s_2 \bar{u} - (x_1 + s_1 \bar{u})| \geq (s_2 - s_1) |\bar{u}| - |x_2| - |x_1| \\ &\geq (s_2 - s_1) |\bar{u}| - C(1 + \sqrt{s_2}) - C(1 + \sqrt{s_1}) \geq (s_2 - s_1) |\bar{u}| - 2C(1 + \sqrt{s_2}). \end{aligned} \quad (4.15)$$

Now, if  $s_2 \leq Ch^{5/3}$ , then

$$d \geq h |\bar{u}| - C(1 + h^{5/6}) \geq ch,$$

for  $h > C$ . On the other hand, if  $s_2 \geq Ch^{5/3}$ , then, as  $s_1 \leq t \leq C(1 + h^{3/2})$ , we have

$$d \geq cs_2 - C\sqrt{s_2} \geq cs_2 \geq ch^{5/3}.$$

Thus, the distance between the points entering  $V(s_1, x_1)$  and  $V(s_2, x_2)$  in (4.14) is large, and spatial decorrelation of  $v(x)$  would imply that our mixing assumption on  $V(t, x)$  holds.

**Exercise 4.1** Formulate carefully a mixing condition on the field  $v(x)$  that would imply the mixing assumption on  $V(t, x)$ . See [30] for the precise details.

**Boundedness.** We assume that the random field  $V(t, x)$  has three spatial derivatives and there exists a deterministic constant  $C > 0$  so that with probability one we have

$$|V(t, x)| + \left| \frac{\partial V(t, x)}{\partial x_j} \right| + \left| \frac{\partial^2 V}{\partial x_i \partial x_j} \right| + \left| \frac{\partial^3 V}{\partial x_i \partial x_j \partial x_k} \right| \leq C < +\infty$$

for all  $1 \leq i, j, l \leq n$ . This assumption can be weakened considerably.

**Incompressibility.** The field  $V$  is divergence free, that is, almost surely

$$\nabla \cdot V(t, x) = \sum_{j=1}^n \frac{\partial V_j}{\partial x_j} = 0.$$

The reason this assumption is made is to avoid sinks and sources that may exist in non-divergence free flows. Random incompressible flows, on the other hand, act more like measure-preserving random rearrangements.

## The limit theorem

Let us define the diffusion matrix

$$a_{pq} = \int_0^\infty E \{V_q(t, 0)V_p(0, 0) + V_p(t, 0)V_q(0, 0)\} dt = \int_0^\infty [R_{pq}(t, 0) + R_{qp}(t, 0)] dt$$

and its symmetric non-negative definite square-root matrix  $\sigma$ :  $\sigma^2 = a$ . Then the following theorem holds.

**Theorem 4.2** *Suppose that the random field  $V(t, x)$  satisfies the assumptions above, and that the matrix  $a_{pq}$  is strictly positive definite. Then the process  $X_\varepsilon(t)$ , which satisfies*

$$\dot{X}_\varepsilon = \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, X_\varepsilon\right), \quad X_\varepsilon(0) = x, \quad (4.16)$$

*converges weakly as  $\varepsilon \rightarrow 0$  to the limit process  $\bar{X}(t) = x + \sigma B(t)$ . Here,  $B(t)$  is the standard Brownian motion.*

The main result of [36] is actually much more general – it applies also to non-divergence free velocities and allows for a mean drift. Then the large time behavior is a sum of a large (order  $1/\varepsilon$ ) deterministic component that comes from the flow compressibility and an order one diffusive process. One can also account for the possible small scale variations of the random field looking at equations of the form

$$\frac{dX}{dt} = \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, \frac{X(t)}{\varepsilon^\alpha}\right)$$

with  $0 \leq \alpha < 1$ . We will not describe these generalizations in detail here. We should also mention that when  $\alpha = 1$  a new regime arises – the time it takes the particle to pass one spatial correlation length is no longer much larger than the correlation time of the random fluctuations. This seriously changes the analysis.

## 5 Basic facts on weak convergence in $C$ and $D$

### Weak convergence

Before we present the proof of Theorem 4.2, we recall in this section basic facts from [6] on weak convergence of probability measures. All the proofs of the results of this section can be found there as well as a wealth of other information. Recall that a sequence of Borel measures  $P_n$  defined on a space  $\Omega$  converges weakly to a Borel measure  $P$  on  $\Omega$  if for every bounded continuous real function  $f$  we have

$$\int_\Omega f dP_n \rightarrow \int_\Omega f dP.$$

Equivalently, for every set  $A$  with  $P(\partial A) = 0$  we have  $P_n(A) \rightarrow P(A)$ . A family  $F$  of (Borel) probability measures on  $\Omega$  is relatively weakly compact if every sequence  $P_n$  of elements in  $F$  contains a weakly convergent subsequence  $P_{n_k}$  which converges weakly to a probability measure  $Q$ .



## Weak convergence in $C$

An effective way to verify weak compactness in the space  $C = C([0, T]; \mathbb{R}^n)$  of continuous functions (paths) is provided by Prokhorov's theorem. Recall that a family  $F$  of probability measures is tight if for every  $\varepsilon > 0$  there exists a compact set  $K$  so that  $P(K) > 1 - \varepsilon$  for all measures  $P \in F$ .

**Theorem 5.1** *If a family  $F$  is tight then it is relatively compact.*

As a corollary we have the following basic criterion for weak convergence.

**Corollary 5.2** *Let  $P_n$  and  $P$  be probability measures on  $C$ . If the finite-dimensional distributions of  $P_n$  converge weakly to those of  $P$  and  $\{P_n\}$  is a tight family then  $P_n$  converge weakly to  $P$ .*

It is important to note that convergence of finite-dimensional distributions in  $C$  in itself does not imply weak convergence and tightness assumption in Corollary 5.2 can not be dropped. Indeed, consider a sequence of piece-wise linear functions  $z_n$  which increase from 0 to 1 on the interval  $[0, 1/n]$ , decrease from 1 to 0 on the interval  $[1/n, 2/n]$  and are equal to zero for  $t \geq 2/n$ . Set the measure  $P_n = \delta_{z_n}$  and let  $P = \delta_0$ , the delta-function concentrated on the function  $z = 0$ . Suppose that  $A$  is a finite-dimensional subset of  $C$ , that is, there exists a finite set of times  $t_1, \dots, t_k$  so that if a path  $x(t)$  lies in  $A$  then so do all paths  $y(t)$  such that  $x(t_i) = y(t_i)$  for all  $1 \leq i \leq k$ . Then, as soon as  $n$  is so large that  $1/n < t_i$  for all  $i = 1, \dots, k$  such that  $t_i > 0$  we have  $P_n(A) = P(A)$  simply because  $z_n(t_j) = z(t_j)$  for all  $j = 1, \dots, k$  (including the time  $t_i = 0$  if there is such an  $i$ ) and thus  $z_n$  lies in  $A$  if and only if  $z \in A$ . On the other hand, if we define  $f(x) = \min[2, \|x\|]$  with the uniform norm

$$\|x\| = \sup_{0 \leq t \leq 1} |x(t)|$$

then  $f$  is a continuous function on  $C$  but

$$\int f dP_n = 1$$

while

$$\int f dP = 0.$$

Therefore  $P_n$  does not converge weakly to  $P$ . This example shows that convergence of finite-dimensional distributions is not sufficient for weak convergence.

The advantage of tightness is that it is a verifiable notion by means of various moduli of continuity. The usual modulus of continuity of a function  $x(t)$ ,  $t \in [0, 1]$  is defined as

$$w_x(\varepsilon) = \sup_{|t-s| \leq \varepsilon} |x(s) - x(t)|, \quad 0 < \varepsilon \leq 1.$$

The Arzela-Ascoli theorem implies that a set  $A$  is relatively compact in  $C$  if and only if both

$$\sup_{x \in A} |x(0)| < +\infty,$$

and

$$\limsup_{\varepsilon \rightarrow 0} \sup_{x \in A} w_x(\varepsilon) = 0.$$

The following theorem (Theorem 7.3 in [6]) is the most basic criterion for tightness in  $C$ .

**Theorem 5.3** *A sequence of probability measures  $P_n$  on  $C$  is tight if and only if the following two conditions hold: (i) for each  $\eta > 0$  there exist  $n_0$  and  $a > 0$  so that*

$$P_n[x : x(0) \geq a] \leq \eta \text{ for all } n \geq n_0, \quad (5.1)$$

*and (ii) for each  $\delta > 0$  and  $\eta > 0$  there exists  $0 < \varepsilon < 1$  and  $n_0$  so that*

$$P_n[x : w_x(\varepsilon) \geq \delta] \leq \eta \text{ for all } n \geq n_0. \quad (5.2)$$

Condition (5.1) is usually easy to verify, especially so when we the measures  $P_n$  are generated by solutions of differential equations (with coefficients that depend on the parameter  $n$ ) with a prescribed initial point – then  $x(0)$  does not depend on  $n$ . On the other hand, verifying (5.2) is the heart of the proof of many limit theorems. Some criteria for (5.2) to hold will be given in the next section.

### The space $D$

It is quite common that one has to deal with convergence of processes that have jumps but are “nice” otherwise. The appropriate space to work with is of functions that have limits on the left and are continuous on the right:

- (i) For  $0 \leq t < 1$  the right limit  $x(t^+) = \lim_{s \rightarrow t^+} x(s)$  exists and  $x(t) = x(t^+)$ .
  - (ii) For  $0 < t \leq 1$  the left limit  $x(t^-) = \lim_{s \rightarrow t^-} x(s)$ .
- (5.3)

Such functions are often called cadlag functions (“continu á droite, limites á gauche”).

Cadlag functions can not be too bad: for instance, it is easy to check that for any cadlag function  $x(t)$  and any  $\varepsilon > 0$  one can find a finite partition  $0 = t_0 < t_1 < \dots < t_n = 1$  of the interval  $[0, 1]$  such that the oscillation  $w_x[t_{i-1}, t_i] < \varepsilon$ . Here the oscillation of a function  $x(t)$  on a set  $S$  is defined as

$$w_x(S) = \sup_{s, t \in S} |x(s) - x(t)|. \quad (5.4)$$

It follows that any cadlag function  $x(t)$  is uniformly bounded and, moreover, has at most countably many discontinuities since the number of points where the jump magnitude exceeds  $1/n$  is finite for all  $n \in \mathbb{N}$ . We will continue to denote the usual uniform norm by

$$\|x\| = \sup_{0 \leq t \leq 1} |x(t)|.$$

The usual uniform topology is too rigid to work in the space  $D$ . If we think of functions in  $D$  as, for instance, realizations of a random jump process, then we would like to think of two realizations as close even if the jumps occur not at exactly the same time but rather at close times. The uniform norm does not capture this idea. Instead, for two functions  $x$  and  $y$  in  $D$  we define the distance  $d(x, y)$  as the smallest number  $\varepsilon > 0$  so that we may find an increasing continuous function (“time change”)  $\lambda(t)$  such that  $\lambda(0) = 0$ ,  $\lambda(1) = 1$  and both

$$\sup_{t \in [0, 1]} |\lambda(t) - t| < \varepsilon$$

and

$$\sup_{t \in [0,1]} |x(t) - y(\lambda(t))| = \sup_{t \in [0,1]} |x(\lambda^{-1}(t)) - y(t)| < \varepsilon. \quad (5.5)$$

This metric defines the Skorohod topology.

Let  $\Lambda$  be the set of increasing continuous functions  $\lambda(t)$  such that  $\lambda(0) = 0$ ,  $\lambda(1) = 1$ . A sequence  $x_n(t)$  converges to  $x(t)$  in the Skorohod topology in  $D$  if there exists a sequence  $\lambda_n \in \Lambda$  such that  $\tilde{x}_n(t) = x_n(\lambda_n(t))$  converges to  $x(t)$  and  $\lambda_n(t)$  converges to  $t$  – both in the uniform topology of  $[0, 1]$ . In particular, the usual uniform convergence implies convergence in the Skorohod topology – simply take  $\lambda_n(t) = t$ . Moreover, as

$$|x_n(t) - x(t)| \leq |x_n(t) - x(\lambda_n(t))| + |x(\lambda_n(t)) - x(t)|, \quad (5.6)$$

it follows that  $x_n(t)$  converges pointwise to  $x(t)$  at the points where  $x(t)$  is continuous. Since  $x(t)$  is continuous for all but countably many points, the Skorohod convergence implies pointwise convergence except on a countable set of points. In addition (5.6) implies that if the limit  $x(t)$  is continuous on  $[0, 1]$  (and hence uniformly continuous) then the Skorohod convergence implies the uniform convergence.

The problem is that the space  $D$  is not complete under the metric  $d$  as can be seen on the following example. Let  $x_n(t) = 1$  for  $0 \leq t \leq 1/2^n$  and  $x_n(t) = 0$  otherwise. Let  $\lambda_n \in \Lambda$  be a (piecewise) linear function:

$$\lambda_n(t) = \frac{t}{2}$$

on the interval  $[0, 1/2^n]$  and

$$\lambda_n(t) = \frac{1}{2^{n+1}} + \frac{1 - \frac{1}{2^{n+1}}}{1 - \frac{1}{2^n}} \left( t - \frac{1}{2^n} \right)$$

on the interval  $[1/2^n, 1]$  so that  $\lambda_n$  maps  $[0, 1/2^n]$  onto  $[0, 1/2^{n+1}]$ . Then  $x_{n+1}(\lambda_n(t)) = x_n(t)$  and  $|\lambda_n(t) - t| \leq 1/2^{n+1}$ . This means that  $d(x_n, x_{n+1}) \leq 1/2^{n+1}$  and the sequence  $x_n(t)$  is Cauchy in the metric  $d$ . On the other hand,  $x_n(t)$  converges pointwise to  $x(t) = 0$  for all  $t > 0$ . Therefore, if  $x_n$  converges in the Skorohod topology the only possible limit function is  $x(t) = 0$  (because the Skorohod convergence implies pointwise convergence except on a countable set). However, the distance from each  $x_n(t)$  to  $x = 0$  is equal to one (simply because  $x(\lambda(t)) \equiv 0$  for all  $\lambda \in \Lambda$  and  $x_n(0) = 1$  for all  $n$ ) and thus  $x_n(t)$  does not converge in the Skorohod topology.

The way to make the space  $D$  complete is to introduce a different metric  $d_0$  defined as follows. For  $\lambda \in \Lambda$  define

$$\|\lambda\|_0 = \sup_{s < t} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right|.$$

This means that the slopes of  $\lambda$  are bounded away from zero and infinity if  $\|\lambda\|_0 < \infty$ . The distance  $d_0(x, y)$  for  $x, y \in D$  is the smallest number  $\varepsilon \geq 0$  so that there exists  $\lambda \in \Lambda$  such that  $\|\lambda\|_0 < \varepsilon$  and (5.5) holds. This is more restrictive than  $d$ : it requires that not only  $\lambda$  is close to identity in the uniform norm but the slopes of  $\lambda$  are all close to one. In particular, the above example of a non-converging Cauchy sequence involves  $\lambda_n$  which are not close to identity in this norm. We have the following proposition.

**Proposition 5.4** *The metrics  $d$  and  $d_0$  are equivalent on  $D$  in the sense that  $d(x_n, x) \rightarrow 0$  if and only if  $d_0(x_n, x) \rightarrow 0$ . Moreover, the space  $D$  is separable under both  $d$  and  $d_0$  and complete under  $d_0$ .*

There is no contradiction in this proposition to the above example of a sequence  $x_n$  which is  $d$ -Cauchy in  $D$  but does not converge. This sequence is simply not  $d_0$ -Cauchy:

$$d_0(x_n, x_{n+1}) = \|\lambda_n\|_0 = \log 2.$$

### Compactness in $D$

Modulus of continuity is not a right notion for a function in  $D$  as  $w_x(\delta)$  does not vanish in the limit  $\delta \rightarrow 0$ . An alternative modulus which allows for jumps is defined as follows. We have mentioned that for any function  $x(t) \in D$  and any  $\varepsilon > 0$  one can find a finite partition

$$0 = t_0 < t_1 < \dots < t_n = 1,$$

such that on each sub-interval the oscillation  $w_x[t_{i-1}, t_i] < \varepsilon$ . We say that a partition  $\{t_i\}$  is  $\delta$ -sparse if  $t_i - t_{i-1} > \delta$  for all  $i$ . Define the modulus

$$w'_x(\delta) = \inf_{\{t_i\}} \max_{1 \leq i \leq n} w_x[t_{i-1}, t_i]$$

with the infimum taken over all  $\delta$ -sparse partitions  $\{t_i\}$ . The previous argument shows that

$$\lim_{\delta \rightarrow 0} w'_x(\delta) = 0,$$

for any cadlag function  $x \in D$ . It is straightforward to check that we always have

$$w'_x(\delta) \leq w_x(2\delta).$$

There can be no inequality in the opposite direction because the usual modulus of continuity  $w_x(\delta)$  does not go to zero as  $\delta \rightarrow 0$  for a discontinuous function from  $D$ . However, for a continuous function  $x(t)$  we do have an inequality  $w_x(\delta) \leq 2w'_x(\delta)$  so for continuous functions the two moduli are equivalent.

The most basic criterion for compactness in  $D$  is the following analog of the Arzela-Ascoli theorem.

**Theorem 5.5** *A necessary and sufficient condition for a set  $A$  to be relatively compact in the Skorohod topology is that  $\sup_{x \in A} \|x\| < \infty$  and  $\lim_{\delta \rightarrow 0} \sup_{x \in A} w'_x(\delta) = 0$ .*

Since the space  $D$  is separable and complete, an immediate consequence of this theorem is the following tightness criterion.

**Theorem 5.6** *A necessary and sufficient condition for a sequence  $P_n$  of probability measures on  $D$  to be tight is that*

$$(i) \quad \lim_{a \rightarrow \infty} \limsup_n P_n [x : \|x\| \geq a] = 0,$$

and

$$(ii) \quad \lim_{\delta \rightarrow 0} \limsup_n P_n [x : w'_x(\delta) \geq \varepsilon] = 0 \text{ for all } \varepsilon > 0.$$

Another useful generalization of the modulus of continuity is the following modulus

$$w_x''(\delta) = \sup_{0 \leq u-s \leq \delta} \left[ \sup_{s \leq t \leq u} (\min[|x(u) - x(t)|, |x(t) - x(s)|]) \right].$$

This is yet another relaxation as it is not hard to see that  $w_x''(\delta) \leq w_x'(\delta)$ . However, once again, there is no inequality in the opposite direction: for the functions

$$x_n(t) = \begin{cases} 1, & \text{for } 0 \leq t < 1/n, \\ 0, & \text{for } 1/n \leq t \leq 1 \end{cases},$$

we have  $w_{x_n}''(\delta) = 0$  while  $w_{x_n}'(\delta) = 1$  for  $\delta > 1/n$  because any  $\delta$ -sparse partition will still contain an interval  $[0, t_1]$  with  $t_1 > \delta > 1/n$  where the oscillation is equal to one. This is an end-point phenomenon which also happens for the functions

$$y_n(t) = \begin{cases} 0, & \text{for } 0 \leq t < 1 - 1/n, \\ 1, & \text{for } 1 - 1/n \leq t \leq 1. \end{cases}$$

Nevertheless, this is the only obstacle for a compactness criterion in terms of  $w_x''(\delta)$  alone. The following result takes this problem into account.

**Theorem 5.7** *A necessary and sufficient condition for a set  $A$  to have a compact closure in the Skorohod topology is that  $\sup_{x \in A} \|x\| < \infty$ ,  $\lim_{\delta \rightarrow 0} \sup_{x \in A} w_x''(\delta) = 0$  and*

$$\lim_{\delta \rightarrow 0} \sup_{x \in A} |x(\delta) - x(0)| = 0, \text{ and } \lim_{\delta \rightarrow 0} \sup_{x \in A} |x(1^-) - x(1 - \delta)| = 0.$$

A direct analog of Theorem 5.6 is then the following.

**Theorem 5.8** *A necessary and sufficient condition for a sequence  $P_n$  of probability measures on  $D$  to be tight is that*

$$(i) \quad \lim_{a \rightarrow \infty} \limsup_n P_n [x : \|x\| \geq a] = 0,$$

and

$$(ii.1) \quad \lim_{\delta \rightarrow 0} \limsup_n P_n [x : w_x''(\delta) \geq \varepsilon] = 0 \text{ for all } \varepsilon > 0,$$

and

$$(ii.2) \quad \begin{cases} \lim_{\delta \rightarrow 0} \limsup_n P_n [x : |x(\delta) - x(0)| \geq \varepsilon] = 0 \\ \lim_{\delta \rightarrow 0} \limsup_n P_n [x : |x(1^-) - x(1 - \delta)| \geq \varepsilon] = 0. \end{cases}$$

A convenient and more practical criterion for weak convergence is the following. Given a probability measure  $P$  we denote by  $T_P$  the set of all times  $t$  such that  $P[J_t] = 0$  where

$$J_t = \{x \in D : x(t) \neq x(t^-)\}$$

is the set of all functions that have a jump at time  $t$ . If  $X$  is a random variable on  $D$  then we write  $T_X$  for  $T_P$  where  $T_P$  is the law of  $X$ .

**Theorem 5.9** *Suppose that the finite-dimensional distributions  $(X_{t_1}^n, \dots, X_{t_k}^n)$  of random variables  $X^n$  defined on  $D$  converge weakly as  $n \rightarrow \infty$  to  $(X_{t_1}, \dots, X_{t_k})$  whenever all  $t_i$  lie in  $T_X$ , and  $X_1 - X_{1-\delta}$  goes weakly to zero as  $\delta \rightarrow 0$ . Assume also that there exists  $\beta \geq 0$  and  $\alpha > 1/2$  so that for all  $r \leq s \leq t$  and  $\lambda > 0$  we have*

$$P[\min\{|X_s^n - X_r^n|, |X_t^n - X_s^n|\} \geq \lambda] \leq \frac{C}{\lambda^{4\beta}} |F(t) - F(s)|^{2\alpha}, \quad (5.7)$$

where  $F$  is a non-decreasing continuous function on  $[0, 1]$ . Then  $X_n$  converge weakly to  $X$  as  $n \rightarrow \infty$ .

The key estimate in the proof of Theorem 5.9 is that (5.7) implies that there exists a constant  $K$  that depends only on  $C$ ,  $\alpha$  and  $\beta$  so that

$$P[w_{X^n}''(\delta) \geq \varepsilon] \leq \frac{K}{\varepsilon^{4\beta}} (F(1) - F(0)) [w_F(2\delta)]^{2\alpha-1}, \quad (5.8)$$

where  $w_F$  is the modulus of continuity of the function  $F$ . This means that (5.7) ensures that condition (ii.1) of Theorem 5.8 holds. A useful and verifiable condition that guarantees (5.7) is that there exist  $\beta > 0$ ,  $\alpha > 1/2$  and  $C > 0$  so that

$$\mathbb{E}\{|X_s^n - X_r^n|^{2\beta} |X_t^n - X_s^n|^{2\beta}\} \leq C|t - r|^{2\alpha} \quad (5.9)$$

for all  $n$ . Then we may take  $F(t) = t$  and (5.8) becomes

$$P[w_{X^n}''(\delta) \geq \varepsilon] \leq \frac{K}{\varepsilon^{4\beta}} \delta^{2\alpha-1}. \quad (5.10)$$

This is why we need  $\alpha > 1/2$  in (5.9). It follows that we may use (5.9) as a substitute for condition (ii.1) in Theorem 5.8.

In turn, the following condition is sufficient to ensure that (5.9) holds: for any  $T > 0$  and  $\nu > 0$  there exists a constant  $C(T, \nu)$  so that for all  $n$ , and all  $0 \leq s \leq t \leq u \leq T$ , we have

$$\mathbb{E}\{|X_n(u) - X_n(t)|^2 |X_n(t) - X_n(s)|^\nu\} \leq C(T, \nu)(u - t) \mathbb{E}\{|X_n(t) - X_n(s)|^\nu\}. \quad (5.11)$$

Indeed, when  $\nu = 0$  in (5.11) we have

$$\mathbb{E}\{|X_n(u) - X_n(t)|^2\} \leq C(T, \nu)(u - t) \text{ for all } n \text{ and all } 0 \leq t \leq u \leq T.$$

Taking  $\nu = 2$  in (5.11) we get, using the above:

$$\begin{aligned} \mathbb{E}\{|X_n(u) - X_n(t)|^2 |X_n(t) - X_n(s)|^2\} &\leq C(T, \nu)(u - t) \mathbb{E}\{|X_n(t) - X_n(s)|^2\} \\ &\leq C(T, \nu)(u - t)(t - s) \leq C(T, \nu)(u - s)^2 \end{aligned} \quad (5.12)$$

and thus (5.9) indeed holds. A somewhat more general estimate than (5.11) is a reformulation in terms of the conditional expectation

$$\mathbb{E}\{|X_n(t) - X_n(s)|^2 \mid \mathcal{F}_s\} \leq C(T)(t - s). \quad (5.13)$$

A practical advantage of working with the conditional expectation in (5.13) is that the power of  $(t - s)$  on the right is equal to one, not larger than one as in (5.9).

## 6 The proof of the limit theorem

We now come to the proof of the limit Theorem 4.2 for the solutions of

$$\dot{X}_\varepsilon = \frac{1}{\varepsilon} V \left( \frac{t}{\varepsilon^2}, X_\varepsilon \right), \quad X_\varepsilon(0) = x. \quad (6.1)$$

The "difficult to believe at first" aspect of Theorem 4.2 is simply the fact that  $X_\varepsilon(t)$  is of the size  $O(1)$  despite the coefficient  $\varepsilon^{-1}$  in front of the velocity field. That is the key to the proof, and is a result of cancellation due to the mixing properties of the velocity field.

The proof proceeds in two steps. The key step is to establish tightness of the processes  $X_\varepsilon(t)$ , so that a limit in law along a subsequence  $\varepsilon_k \rightarrow 0$  exists. This is done in the space  $D$  of cadlag functions. However, as the processes  $X_\varepsilon(t)$  are all continuous the limit process also has to be continuous and convergence take place in the space  $C$  of continuous functions. In the last step we show that the only possible limit along a subsequence is a Brownian motion multiplied by the matrix  $\sigma$ . This uses the martingale characterization of the Brownian motion.

Tightness of  $X_\varepsilon$  is a consequence of the following.

**Proposition 6.1** *There exist  $C > 0$  and  $\nu > 0$  so that*

$$\mathbb{E} \left\{ |X_\varepsilon(t) - X_\varepsilon(s)|^2 |X_\varepsilon(s) - X_\varepsilon(u)|^2 \right\} \leq C(t-u)^{1+\nu}, \quad (6.2)$$

for all  $0 \leq u \leq s \leq t \leq T$ .

This is criterion (5.9) for tightness in the space  $D$  with  $\beta = 1$  and  $\alpha = (1 + \nu)/2$ . The main step in the proof of (6.2) is to find  $\gamma \in (1, 2)$  such that we have the following estimate for the conditional expectation

$$\mathbb{E} \left\{ |X_\varepsilon(t) - X_\varepsilon(s)|^2 | \mathcal{F}_s \right\} \leq C(t-s) \text{ for } t-s > 10\varepsilon^\gamma. \quad (6.3)$$

As we have explained, the estimate (6.3) itself is sufficient to establish tightness in  $D$  for the family  $X_\varepsilon(t)$  if it were to hold for all  $t > s$  – see (5.13). As we will prove it only for pairs of time with a gap:  $t-s > 10\varepsilon^\gamma$ , we may at the moment conclude only that

$$\mathbb{E} \left\{ |X_\varepsilon(t) - X_\varepsilon(s)|^2 |X_\varepsilon(s) - X_\varepsilon(u)|^2 \right\} \leq C(t-u)^2 \text{ for } t-s > 10\varepsilon^\gamma \text{ and } s-u > 10\varepsilon^\gamma.$$

Our first step is to establish that, with an appropriate choice of  $\gamma \in (1, 2)$ , if either  $t-s \leq 10\varepsilon^\gamma$  or  $s-u \leq 10\varepsilon^\gamma$ , the estimate (6.2) follows from (6.3) together with the dynamical system (6.1) governing  $X_\varepsilon(t)$ . If both  $t-s \leq 10\varepsilon^\gamma$  and  $s-u \leq 10\varepsilon^\gamma$  then we have directly from (6.1):

$$\mathbb{E} \left\{ |X_\varepsilon(t) - X_\varepsilon(s)|^2 |X_\varepsilon(s) - X_\varepsilon(u)|^2 \right\} \leq \frac{C(t-s)^2(s-u)^2}{\varepsilon^4} \leq \frac{C\varepsilon^{11\gamma/4}(t-u)^{5/4}}{\varepsilon^4} \leq C(t-u)^{5/4},$$

provided that  $\gamma > 16/11$ . On the other hand, if, say,  $t-s \leq 10\varepsilon^\gamma$  but  $s-u > 10\varepsilon^\gamma$ , (6.3) implies that

$$\mathbb{E} \left\{ |X_\varepsilon(s) - X_\varepsilon(u)|^2 \right\} \leq C(s-u),$$

and (6.1) implies that with probability one

$$|X_\varepsilon(t) - X_\varepsilon(s)| \leq \frac{C(t-s)}{\varepsilon}.$$

Therefore, the following estimate holds for such times  $t, s$  and  $u$ :

$$\begin{aligned} \mathbb{E} \{ |X_\varepsilon(t) - X_\varepsilon(s)|^2 |X_\varepsilon(s) - X_\varepsilon(u)|^2 \} &\leq \frac{C}{\varepsilon^2} (t-s)^2 (s-u) \\ &\leq C\varepsilon^{7\gamma/4-2} (t-u)^{5/4} \leq C(t-u)^{5/4}, \end{aligned}$$

provided that  $\gamma > 8/7$ . We see that, indeed, (6.3) together with (6.1) are sufficient to prove the tightness criterion (6.2). The rest of the proof of tightness of the processes  $X_\varepsilon(t)$  is concerned with verifying (6.3).

### The mixing lemmas

A crucial component in many proofs of this kind is some sort of a mixing lemma that is needed to establish the tightness of the dynamics. It translates the mixing properties of the random field into the mixing properties of the trajectories.

We set  $G_0(s, x) = V(s, x)$  and

$$G_{1,j}(s_1, s_2, x) = \sum_{p=1}^n V_p(s_2, x) \frac{\partial V_j(s_1, x)}{\partial x_p}, \quad j = 1, \dots, n.$$

**Exercise 6.2** Show that incompressibility of  $V(t, x)$  and its spatial stationarity imply that

$$\mathbb{E}\{G_1(s_1, s_2, x)\} = 0,$$

for all  $s_1, s_2$  and  $x$ .

The next lemma quantifies the mixing of the trajectories.

**Lemma 6.3** Fix  $T \geq 0$  and let  $0 \leq u \leq s \leq T$ . Assume that  $Y$  is a  $\mathcal{V}_0^{s/\varepsilon^2}$ -measurable random vector function. Then there exists  $\varepsilon_0 > 0$  and a constant  $C > 0$  such that for any  $0 \leq u \leq s \leq s_2 \leq s_1 \leq T$  and  $0 < \varepsilon < \varepsilon_0$  we have

$$\left| \mathbb{E} \left\{ V \left( \frac{s_1}{\varepsilon^2}, X_\varepsilon(u) \right) Y \left( \frac{s}{\varepsilon^2} \right) \right\} \right| \leq C\beta(s_1 - s) \mathbb{E} \left| Y \left( \frac{s}{\varepsilon^2} \right) \right|, \quad (6.4)$$

$$\left| \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[ V \left( \frac{s_1}{\varepsilon^2}, X_\varepsilon(u) \right) \right] Y \left( \frac{s}{\varepsilon^2} \right) \right\} \right| \leq C\beta(s_1 - s) \mathbb{E} \left| Y \left( \frac{s}{\varepsilon^2} \right) \right|, \quad (6.5)$$

and

$$\left| \mathbb{E} \left\{ G_1 \left( \frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u) \right) Y \left( \frac{s}{\varepsilon^2} \right) \right\} \right| \leq C\beta^{1/2}(s_1 - s_2)\beta^{1/2}(s_2 - s) \mathbb{E} \left| Y \left( \frac{s}{\varepsilon^2} \right) \right|, \quad (6.6)$$

$$\left| \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[ G_1 \left( \frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u) \right) \right] Y \left( \frac{s}{\varepsilon^2} \right) \right\} \right| \leq C\beta^{1/2}(s_1 - s_2)\beta^{1/2}(s_2 - s) \mathbb{E} \left| Y \left( \frac{s}{\varepsilon^2} \right) \right|, \quad (6.7)$$

for all  $1 \leq k \leq n$ .



**Proof.** First of all, we note that for  $\rho > 1/2$ ,  $C > 1 + \sup |V(t, x)|$  and  $0 < \varepsilon < \varepsilon_0(T)$  the process  $X_\varepsilon(t)$ ,  $0 \leq t \leq u \leq T$  does not leave the ball of the radius  $C(1 + u^\rho/\varepsilon^{2\rho})$  centered at the origin, and hence is  $\mathcal{V}_0^{u/\varepsilon^2}(C, \rho)$ -measurable:

$$|X_\varepsilon(t)| \leq \frac{1}{\varepsilon} \int_0^u \left| V\left(\frac{s}{\varepsilon^2}, X_\varepsilon(s)\right) \right| ds \leq \frac{Cu}{\varepsilon} \leq C \left(1 + \frac{u^\rho}{\varepsilon^{2\rho}}\right)$$

for all  $0 \leq t \leq u$ .

We first prove (6.4)-(6.5). We prove only the second inequality, (6.5) as the proof of (6.4) is identical. The idea is to replace the random variable  $X_\varepsilon(u)$  by a deterministic value and use the mixing properties of the field  $V(t, x)$  in time. Let  $M \in \mathbb{N}$  be a fixed positive integer and  $l \in \mathbb{Z}^n$ . Define the event

$$A(l) = \left[ \omega : \frac{l_j}{M} \leq X_j^\varepsilon(u) < \frac{l_j + 1}{M}, \quad j = 1, \dots, n \right], \quad l = (l_1, \dots, l_n).$$

The event  $A(l)$  is  $\mathcal{V}_0^{s/\varepsilon^2}$  measurable since  $u \leq s$ . We may decompose the expectation in (6.5) using the fact that the random variable  $X_\varepsilon(u)$  is close to the non-random value  $l/M$  on the event  $A(l)$  as follows:

$$\begin{aligned} & \left| \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[ V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right) \right] Y\left(\frac{s}{\varepsilon^2}\right) \right\} \right| = \left| \sum_l \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[ V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right) \right] Y\left(\frac{s}{\varepsilon^2}\right) \chi_{A(l)} \right\} \right| \\ & \leq \left| \sum_l \mathbb{E} \left\{ \left[ \frac{\partial}{\partial x_k} \left[ V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right) \right] - \frac{\partial}{\partial x_k} \left[ V\left(\frac{s_1}{\varepsilon^2}, \frac{l}{M}\right) \right] \right] Y\left(\frac{s}{\varepsilon^2}\right) \chi_{A(l)} \right\} \right| \\ & + \left| \sum_l \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[ V\left(\frac{s_1}{\varepsilon^2}, \frac{l}{M}\right) \right] Y\left(\frac{s}{\varepsilon^2}\right) \chi_{A(l)} \right\} \right| = I + II. \end{aligned}$$

As the points  $l/M$  are deterministic, the second term above may be now estimated using the mixing property (4.13) and the fact that  $\mathbb{E} \{ \partial V / \partial x_k \} = 0$  by

$$II \leq 2K\beta \left( \frac{s_1 - s}{\varepsilon^2} \right) \sum_l \mathbb{E} \left\{ \left| Y\left(\frac{s}{\varepsilon^2}\right) \right| \chi_{A(l)} \right\} = 2K\beta \left( \frac{s_1 - s}{\varepsilon^2} \right) \mathbb{E} \left\{ \left| Y\left(\frac{s}{\varepsilon^2}\right) \right| \right\},$$

uniformly in  $M$ .

As far as  $I$  is concerned, we have assumed that two spatial derivatives of the field  $V(t, x)$  are bounded by a deterministic constant, hence  $\partial V / \partial x_k$  is uniformly continuous in space. Therefore, using the Lebesgue dominated convergence theorem we conclude that

$$I \leq \frac{C}{M} \sum_l \mathbb{E} \left\{ \left| Y\left(\frac{s}{\varepsilon^2}\right) \right| \chi_{A(l)} \right\} = \frac{C}{M} \sum_l \mathbb{E} \left\{ \left| Y\left(\frac{s}{\varepsilon^2}\right) \right| \right\} \rightarrow 0 \text{ as } M \rightarrow +\infty,$$

and (6.5) follows. An identical proof shows that in addition we have the same bound for the second derivatives of the random field  $V$ :

$$\left| \mathbb{E} \left\{ \frac{\partial^2}{\partial x_k \partial x_m} \left[ V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right) \right] Y\left(\frac{s}{\varepsilon^2}\right) \right\} \right| \leq C\beta(s_1 - s) \mathbb{E} \left| Y\left(\frac{s}{\varepsilon^2}\right) \right|. \quad (6.8)$$

We now prove (6.7) – the proof of (6.6) is identical. Let us first write out the expression for  $G_1$  :

$$\begin{aligned} & \left| \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[ G_1 \left( \frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u) \right) \right] Y \left( \frac{s}{\varepsilon^2} \right) \right\} \right| \\ & \leq \sum_{p=1}^n \left| \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[ V_p \left( \frac{s_2}{\varepsilon^2}, X_\varepsilon(u) \right) \frac{\partial}{\partial x_p} \left( V \left( \frac{s_1}{\varepsilon^2}, X_\varepsilon(u) \right) \right) \right] Y \left( \frac{s}{\varepsilon^2} \right) \right\} \right| \end{aligned}$$

Now we may apply (6.5), (6.8) in two different ways using different parts of the inequality

$$0 \leq u \leq s \leq s_2 \leq s_1.$$

First, we may use (6.5), (6.8) with the gap between  $s_1$  and  $s_2$ , that is, we group into “ $Y$ ” in (6.5), (6.8) all terms that involve  $s$  and  $s_2$ . Using in addition the uniform bounds on  $V$  and its derivatives this leads to

$$\left| \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[ G_1 \left( \frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u) \right) \right] Y \left( \frac{s}{\varepsilon^2} \right) \right\} \right| \leq C\beta \left( \frac{s_1 - s_2}{\varepsilon^2} \right) \mathbb{E} \left\{ \left| Y \left( \frac{s}{\varepsilon^2} \right) \right| \right\}.$$

Second, note that (6.5) may be slightly generalized to apply with  $\partial V/\partial x_k$  replaced by a sufficiently smooth in space  $\mathcal{V}_{s_1}^T$  random variable with an expectation equal to zero. As  $\mathbb{E}\{G_1\} = 0$  indeed, we can use this modified version of (6.5) with the gap between  $s_2$  and  $s$ , taking “ $Y$ ” in (6.5) to be simply  $Y(s/\varepsilon^2)$ :

$$\left| \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[ G_1 \left( \frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u) \right) \right] Y \left( \frac{s}{\varepsilon^2} \right) \right\} \right| \leq C\beta \left( \frac{s_2 - s}{\varepsilon^2} \right) \mathbb{E} \left\{ \left| Y \left( \frac{s}{\varepsilon^2} \right) \right| \right\}.$$

Multiplying these two inequalities and taking the square root we conclude that (6.7) holds. This finishes the proof of Lemma 6.3.  $\square$

## The proof of Proposition 6.1

**Step 1. Taking a time-step backward.** We start with a pair of times  $t > s$  with a gap between them:  $t - s > 10\varepsilon^\gamma$ . Consider a partition  $s = t_0 < t_1 < \dots < t_{M+1} = t$  of the interval  $[s, t]$  into subintervals of the length

$$\Delta t = l_\varepsilon = (t - s) \left( \left[ \frac{t - s}{\varepsilon^\gamma} \right] \right)^{-1},$$

where  $[x]$  is the integer part of  $x$ , so that  $\varepsilon^\gamma/2 \leq l_\varepsilon \leq 2\varepsilon^\gamma$ . The parameter  $\gamma \in (1, 2)$  is to be defined later. The important aspect is that  $\gamma < 2$  so that  $\Delta t$  is much larger than the velocity correlation time  $\varepsilon^2$ . The basic idea in the proof of (6.3) is “to expand  $X_\varepsilon(t) - X_\varepsilon(s)$  in a Taylor series” with a “large” time step  $O(\Delta t)$ . The first two terms in this expansion will be explicitly computable. The error terms which are nominally large are shown to be small using the mixing Lemma 6.3. The last point is the key to the whole argument.

Dropping the subscript  $\varepsilon$  of  $X_\varepsilon$  we write for  $t > s$ :

$$X(t) - X(s) = \frac{1}{\varepsilon} \int_s^t V \left( \frac{u}{\varepsilon^2}, X(u) \right) du = \frac{1}{\varepsilon} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} V \left( \frac{u}{\varepsilon^2}, X(u) \right) du \quad (6.9)$$

Our task is to estimate the integral inside the summation in the right side of (6.9). In the preparation for the application of the mixing lemma, on the interval  $t_i \leq u \leq t_{i+1}$  the integrand can be rewritten as

$$\begin{aligned} V\left(\frac{u}{\varepsilon^2}, X(u)\right) &= V\left(\frac{u}{\varepsilon^2}, X(t_{i-1})\right) + \int_{t_{i-1}}^u \frac{d}{du_1} V\left(\frac{u}{\varepsilon^2}, X(u_1)\right) du \\ &= V\left(\frac{u}{\varepsilon^2}, X(t_{i-1})\right) + \int_{t_{i-1}}^u \sum_{p=1}^n \frac{\partial}{\partial x_p} \left[ V\left(\frac{u}{\varepsilon^2}, X(u_1)\right) \right] \left( \frac{1}{\varepsilon} V_p\left(\frac{u_1}{\varepsilon^2}, X(u_1)\right) \right) du_1 \\ &= V\left(\frac{u}{\varepsilon^2}, X(t_{i-1})\right) + \frac{1}{\varepsilon} \int_{t_{i-1}}^u G_1\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(u_1)\right) du_1. \end{aligned}$$

The next step is to expand  $G_1$  as well, also around the ‘‘one-step-backward’’ time  $t_{i-1}$ :

$$G_1\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(u_1)\right) = G_1\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1})\right) + \frac{1}{\varepsilon} \int_{t_{i-1}}^{u_1} G_2\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, \frac{u_2}{\varepsilon^2}, X(u_2)\right) du_2$$

with

$$G_2(u, u_1, u_2, x) = \sum_{q=1}^n \frac{\partial}{\partial x_q} [G_1(u, u_1, x)] V_q(u_2, x).$$

Putting together the above calculations we see that

$$\begin{aligned} X(t) - X(s) &= \frac{1}{\varepsilon} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} V\left(\frac{u}{\varepsilon^2}, X(u)\right) du = \frac{1}{\varepsilon} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} V\left(\frac{u}{\varepsilon^2}, X(t_{i-1})\right) du \\ &+ \frac{1}{\varepsilon^2} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} \left[ \int_{t_{i-1}}^u G_1\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(u_1)\right) du_1 \right] du \\ &= \frac{1}{\varepsilon} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} V\left(\frac{u}{\varepsilon^2}, X(t_{i-1})\right) du + \frac{1}{\varepsilon^2} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} \left[ \int_{t_{i-1}}^u G_1\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1})\right) du_1 \right] du \\ &+ \frac{1}{\varepsilon^3} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} \left[ \int_{t_{i-1}}^u \left[ \int_{t_{i-1}}^{u_1} G_2\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, \frac{u_2}{\varepsilon^2}, X(u_2)\right) du_2 \right] du_1 \right] du. \end{aligned}$$

The triple integral on the last line is deterministically small with an appropriate choice of  $\gamma$ : the time interval in each integration is smaller than  $\varepsilon^\gamma$  and the total number of terms is at most  $2(t-s)/\varepsilon^\gamma$  as we have assumed that  $t-s \geq 10\varepsilon^\gamma$ . Therefore, the last integral is bounded by

$$\frac{1}{\varepsilon^3} \left| \sum_{i=0}^M \int_{t_i}^{t_{i+1}} \left[ \int_{t_{i-1}}^u \left[ \int_{t_{i-1}}^{u_1} G_2\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, \frac{u_2}{\varepsilon^2}, X(u_2)\right) du_2 \right] du_1 \right] du \right| \leq C\varepsilon^{2\gamma-3}(t-s)$$

which is small if  $\gamma > 3/2$ . This is a general idea in proofs of weak coupling limits: pull back one time step and expand the integrands until they become almost surely small, then compute the limit of the (very) finite number of surviving terms. In our present case we have shown that, for  $3/2 < \gamma < 2$ ,

$$X(t) - X(s) = L_1(s, t) + L_2(s, t) + E(s, t)$$

where

$$L_1(s, t) = \frac{1}{\varepsilon} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} V \left( \frac{u}{\varepsilon^2}, X(t_{i-1}) \right) du$$

and

$$L_2(s, t) = \frac{1}{\varepsilon^2} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} \left[ \int_{t_{i-1}}^u G_1 \left( \frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1}) \right) du_1 \right] du,$$

while  $|E(s, t)| \leq C\varepsilon^\alpha(t - s)$  with some  $\alpha > 0$  and a deterministic constant  $C > 0$ . This finishes the first preliminary step in the proof of tightness.

**Step 2. Application of the tightness criterion.** Now we are ready to prove (6.3). That is, we have to verify that for any non-negative and  $\mathcal{V}_0^{s/\varepsilon^2}$ -measurable random variable  $Y$  we have for all  $0 \leq s \leq t \leq T$  such that  $t \geq s + 10\varepsilon^\gamma$ :

$$\mathbb{E} \{ |X(t) - X(s)|^2 Y \} \leq C(T)(t - s) \mathbb{E} \{ Y \}.$$

Our estimates in Step 1 show that it is actually enough to verify that

$$\mathbb{E} \{ (L_j(s, t))^2 Y \} \leq C(t - s) \mathbb{E} \{ Y \}, \quad j = 1, 2.$$

**An estimate for  $L_1$ .** We first look at the term corresponding to  $L_1$ : it is equal to

$$\begin{aligned} \mathbb{E} \{ (L_1(s, t))^2 Y \} &= \frac{2}{\varepsilon^2} \sum_{i < j}^n \sum_{p=1}^n \int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} \mathbb{E} \left\{ V_p \left( \frac{u}{\varepsilon^2}, X(t_{i-1}) \right) V_p \left( \frac{u'}{\varepsilon^2}, X(t_{j-1}) \right) Y \right\} du' du \\ &+ \frac{1}{\varepsilon^2} \sum_j^n \sum_{p=1}^n \int_{t_j}^{t_{j+1}} \int_{t_j}^{t_{j+1}} \mathbb{E} \left\{ V_p \left( \frac{u}{\varepsilon^2}, X(t_{j-1}) \right) V_p \left( \frac{u'}{\varepsilon^2}, X(t_{j-1}) \right) Y \right\} du' du = \sum_{i \leq j} I_{ij}. \end{aligned}$$

The idea is to use separation between  $t_{i-1}$  and  $t_{j-1}$  and apply the mixing lemma. Accordingly we look at the cases  $i \leq j - 2$ ,  $i = j - 1$  and  $i = j$  separately as the terms end up being of a different order. The terms with  $i \leq j - 2$  may be estimated with the help of the mixing Lemma 6.3 using the time gap between the times  $u'$  and  $t_{j-1} \geq t_{i+1} \geq u$  which is much larger than the correlation time  $\varepsilon^2$ :

$$\begin{aligned} \sum_{j=0}^M \sum_{i \leq j-2} |I_{ij}| &\leq \frac{C}{\varepsilon^2} \sum_{j=0}^M \sum_{i \leq j-2} \int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} \beta \left( \frac{u' - t_{j-1}}{\varepsilon^2} \right) \mathbb{E} \{ Y \} du' du \\ &\leq \frac{C}{\varepsilon^2} \beta(\varepsilon^{-2+\gamma}) (t - s)^2 E \{ Y \} \leq C\varepsilon^p (t - s) E \{ Y \} \end{aligned}$$

for any  $p > 0$  since  $\gamma < 2$  and  $\beta(s)$  decays faster than any power of  $s$ . The term  $I_3$  corresponding to  $i = j$  can be estimated using the mixing lemma again, using the fact that  $t_{j-1}$  is

smaller than both  $u$  and  $u'$ :

$$\begin{aligned} \sum_{j=0}^M |I_{jj}| &\leq \frac{C}{\varepsilon^2} \sum_{j=0}^M \int_{t_j}^{t_{j+1}} \int_{t_j}^{t_{j+1}} \mathbb{E} \left\{ V_p \left( \frac{u}{\varepsilon^2}, X(t_{j-1}) \right) V_p \left( \frac{u'}{\varepsilon^2}, X(t_{j-1}) \right) Y \right\} du' du \quad (6.10) \\ &\leq \frac{2C}{\varepsilon^2} \sum_{j \in I} \int_{t_j}^{t_{j+1}} \int_{u'}^{t_{j+1}} \beta \left( \frac{u - u'}{\varepsilon^2} \right) dud u' E \{Y\} \leq C(t - s) E \{Y\} \int_0^\infty \beta(u) du. \end{aligned}$$

The integral  $I_2$  with  $i = j - 1$  is estimated similarly.

**A better estimate estimate for  $L_1$ .** Let us now go one step further and actually identify the limit of  $E\{L_{1,j}(s, t)L_{1,m}(s, t)Y\}$  with  $1 \leq j, m \leq n$ . The previous calculations already show that the term corresponding to the previous  $I_1$  (but now with  $V_j$  and  $V_m$  replacing  $V_p$  and  $V_p$ ) satisfies

$$|I_1| \leq C\varepsilon^\alpha(t - s)E\{Y\},$$

with  $\alpha > 0$  so we are interested only in the limit of  $I_2$  and  $I_3$ . The term  $I_3$  is computed as in (6.10) with the help of the mixing lemma:

$$\begin{aligned} \sum_{j \in I} |I_{jj}| &= \frac{1}{\varepsilon^2} \sum_{j=0}^M \int_{t_j}^{t_{j+1}} \int_{t_j}^{t_{j+1}} \mathbb{E} \left\{ V_j \left( \frac{u}{\varepsilon^2}, X(t_{j-1}) \right) V_m \left( \frac{u'}{\varepsilon^2}, X(t_{j-1}) \right) Y \right\} du' du \quad (6.11) \\ &= \frac{1}{\varepsilon^2} \sum_{j=0}^M \int_{t_j}^{t_{j+1}} \int_{t_j}^{t_{j+1}} R_{jm} \left( \frac{u - u'}{\varepsilon^2}, 0 \right) dud u' E \{Y\} + o(1)(t - s) E \{Y\} \\ &= \left[ \int_{-\infty}^\infty R_{jm}(\tau, 0) d\tau + o(1) \right] (t - s) E \{Y\}. \end{aligned}$$

Finally,  $I_2$  corresponding to  $i = j - 1$  is computed as

$$\begin{aligned} \sum_{j \in I} |I_{j-1,j}| &= \frac{1}{\varepsilon^2} \sum_{j=0}^M \int_{t_j}^{t_{j+1}} \int_{t_{j-1}}^{t_j} \mathbb{E} \left\{ V_j \left( \frac{u}{\varepsilon^2}, X(t_{j-1}) \right) V_m \left( \frac{u'}{\varepsilon^2}, X(t_{j-2}) \right) Y \right\} du' du \quad (6.12) \\ &= \frac{1}{\varepsilon^2} \sum_{j \in I} \int_{t_j}^{t_{j+1}} \int_{t_{j-1}}^{t_j} R_{jm} \left( \frac{u - u'}{\varepsilon^2}, 0 \right) dud u' E \{Y\} + o(1)(t - s) E \{Y\} = o(1)(t - s) E \{Y\}. \end{aligned}$$

because  $t_{j+1} - t_j = \varepsilon^\gamma \gg \varepsilon^2$ . Therefore we actually have a more precise estimate

$$\mathbb{E} \{ (L_{1,j}(s, t)L_{1,m}(s, t))Y \} = \left[ \int_{-\infty}^\infty R_{jm}(\tau, 0) d\tau + o(1) \right] (t - s) E \{Y\}. \quad (6.13)$$

**An estimate for  $L_2$ .** Following the above steps one also establishes the required estimate for  $L_2$ :

$$\mathbb{E} \{ (L_2(s, t))^2 Y \} \leq C(t - s) E \{Y\}. \quad (6.14)$$

There is no reason to repeat these calculations separately for  $L_2$  except that an even stronger estimate than (6.14) holds with an appropriate choice of  $\gamma$ :

$$\mathbb{E} \{ (L_2(s, t))^2 Y \} \leq C \varepsilon^\alpha (t - s) E \{ Y \} \quad (6.15)$$

with  $\alpha > 0$ . We will need (6.15) in the identification of the limit, thus we will show it now:

$$\begin{aligned} & \mathbb{E} \{ (L_2(s, t))^2 Y \} \\ &= \frac{1}{\varepsilon^4} \sum_{i,j} \int_{t_i}^{t_{i+1}} du \int_{t_j}^{t_{j+1}} du' \int_{t_{i-1}}^u du_1 \int_{t_{j-1}}^{u'} du'_1 \mathbb{E} \left\{ G_1 \left( \frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1}) \right) G_1 \left( \frac{u'}{\varepsilon^2}, \frac{u'_1}{\varepsilon^2}, X(t_{j-1}) \right) Y \right\}. \end{aligned}$$

Once again, you split the sum above into terms with  $i \leq j - 2$ ,  $i = j - 1$  and  $i = j$ : those with  $i \leq j - 2$  add up to

$$\begin{aligned} & \frac{1}{\varepsilon^4} \sum_{i \leq j-2} \int_{t_i}^{t_{i+1}} du \int_{t_j}^{t_{j+1}} du' \int_{t_{i-1}}^u du_1 \int_{t_{j-1}}^{u'} du'_1 \mathbb{E} \left\{ G_1 \left( \frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1}) \right) G_1 \left( \frac{u'}{\varepsilon^2}, \frac{u'_1}{\varepsilon^2}, X(t_{j-1}) \right) Y \right\} \\ & \leq C \varepsilon^{2\gamma-4} \beta (\varepsilon^{\gamma-2}) (t - s)^2 \mathbb{E} \{ Y \}. \end{aligned}$$

We used in the above estimate the mixing lemma with the gap between  $t_{i-1}$  and  $t_{j-1}$  as well as the fact that the length of each time interval is  $\varepsilon^\gamma$  while the total number of terms in the sum is not more than  $(2(t - s)/\varepsilon^\gamma)^2$ . The important difference with  $L_1$  is that the term with  $i = j$  is also small:

$$\begin{aligned} & \frac{1}{\varepsilon^4} \sum_i \int_{t_i}^{t_{i+1}} du \int_{t_i}^{t_{i+1}} du' \int_{t_{i-1}}^u du_1 \int_{t_{i-1}}^{u'} du'_1 \mathbb{E} \left\{ G_1 \left( \frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1}) \right) G_1 \left( \frac{u'}{\varepsilon^2}, \frac{u'_1}{\varepsilon^2}, X(t_{i-1}) \right) Y \right\} \\ & \leq C \varepsilon^{3\gamma-4} (t - s) \mathbb{E} \{ Y \} \end{aligned}$$

simply because now the number of summands is bounded by  $(2(t - s)/\varepsilon^\gamma)$  (without the square). This means that if we take  $\gamma > 4/3$  this term is bounded by the right side of (6.15). The contribution of the terms with  $i = j - 1$  is estimated identically – hence (6.15) indeed holds.

Summarizing our work so far (and restoring the missing indices) we have shown that

$$\mathbb{E} \{ (X_m(t) - X_m(s))(X_n(t) - X_n(s))Y \} = \left[ \int_{-\infty}^{\infty} R_{mn}(\tau, 0) d\tau + o(1) \right] (t - s) E \{ Y \} \quad (6.16)$$

for all  $t > s$  with  $t - s \geq 10\varepsilon^\gamma$ . This, of course, implies (6.3) and hence the tightness of the family  $X_\varepsilon(t)$  follows.

### Identification of the limit

In order to identify the limit all we have to do is verify that the limit is continuous (that we already know) and the following two conditions hold: first,

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \{ [(X_j^\varepsilon(t) - X_j^\varepsilon(s))(X_m^\varepsilon(t) - X_m^\varepsilon(s)) - a_{jm}(t - s)] \Psi \} = 0$$

for all bounded non-negative continuous functions

$$\Psi = \Psi(X_\varepsilon(t_1), \dots, X_\varepsilon(t_n)),$$

with  $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq s < t \leq T$ . Second, we need

$$\limsup_{\varepsilon \rightarrow 0} \mathbb{E} \{ (X_j^\varepsilon(t))^4 \} < +\infty$$

for all  $t > 0$ . The former condition we have already verified in the previous section in the proof of tightness. The latter may be checked using very similar arguments. This finishes the proof of Theorem 4.2.  $\square$

## 7 Mixing in strong shear flows

In the previous sections, we have considered a flow with a uniform mean. Let us now consider what happens when the mean flow is not uniform and leads to some stretching. A simple example of such dynamics is advection by a shear flow in a channel

$$D = \{(x, y) : x \in \mathbb{R}, y \in \Omega \subset \mathbb{R}^d\} \subset \mathbb{R}^{d+1}. \quad (7.1)$$

Here,  $\Omega$  is a smooth bounded domain – the channel cross-section. The flow trajectories are straight lines along the channel:

$$\frac{dX}{dt} = u(Y(t)), \quad \frac{dY}{dt} = 0, \quad X(0) = x, \quad Y(0) = y. \quad (7.2)$$

This, of course, has an explicit solution

$$X(t) = x + u(y)t, \quad Y(t) = y. \quad (7.3)$$

In order to randomly perturb this flow, we will add a diffusive perturbation to the flow. This avoids many of the technicalities we have encountered when dealing with a time-independent random flow. Thus, we consider a system of two stochastic differential equations

$$\begin{aligned} dX_t &= -u(Y_t) + \sqrt{2\varepsilon} dB_t^{(1)}, \\ dY_t &= \sqrt{2\varepsilon} dB_t^{(2)}. \end{aligned} \quad (7.4)$$

The Brownian motion  $B_t^{(1)}$  is one-dimensional, while  $B_t^{(2)}$  is  $d$ -dimensional. The corresponding Kolmogorov equation is

$$\frac{\partial \phi}{\partial t} + u(y) \frac{\partial \phi}{\partial x} = \varepsilon^2 \Delta \phi. \quad (7.5)$$

We will consider this problem on long time scales, of the order  $t \sim \varepsilon^2$ , to make the effect of the random perturbation non-trivial. The corresponding time-rescaling gives

$$\frac{\partial \phi}{\partial t} + \frac{1}{\varepsilon^2} u(y) \frac{\partial \phi}{\partial x} = \Delta \phi. \quad (7.6)$$

This problem is posed in the channel  $D$  in (7.1), with the Neumann boundary condition at the boundary:

$$\frac{\partial \phi}{\partial \nu} = 0 \text{ at } \partial D = \mathbb{R} \times \partial \Omega. \quad (7.7)$$

Note that if  $u(y) \equiv \bar{u} = \text{const}$ , then solution of (7.7) is simply a translate of the solution of the heat equation:

$$\phi(t, x, y) = \bar{\phi}(t, x - \bar{u} \frac{t}{\varepsilon^2}, y), \quad (7.8)$$

where  $\bar{\phi}(t, x, y)$  is the solution of the standard heat equation

$$\frac{\partial \bar{\phi}}{\partial t} = \Delta \bar{\phi}, \quad (7.9)$$

with the Neumann boundary conditions (7.7). Therefore, if  $u(y)$  is a uniform flow, then the solution of (7.6) behaves as the solution of the standard heat equation.

## Quenching by a shear flow

The goal of this section is to investigate what happens if the flow is not uniform – there is a speed mismatch when moving along the trajectories. This will connect the Kesten-Papanicolaou theorem on the behavior of the particles in a flow with a uniform mean to the relaxation enhancement results. Thus, we will ask the same question as in the relaxation enhancement setting: consider the solution of

$$\frac{\partial \phi}{\partial t} + \frac{1}{\varepsilon^2} u(y) \frac{\partial \phi}{\partial x} = \Delta \phi. \quad (7.10)$$

$$\begin{aligned} \frac{\partial \phi}{\partial \nu} &= 0 \text{ at } \partial D = \mathbb{R} \times \partial \Omega, \\ \phi(0, x, y) &= \phi_0(x, y), \end{aligned} \quad (7.11)$$

with a rapidly decaying initial condition  $\phi_0(x, y)$ . When is it true that for any time  $\tau > 0$  and  $\delta > 0$  we have

$$\|\phi(\tau, \cdot)\|_{L^\infty(D)} \leq \delta, \quad (7.12)$$

provided that  $\varepsilon < \varepsilon_0(\tau, \delta)$ ?

**Definition 7.1** *We say that the profile  $u(y)$  is quenching if for any  $L$  and any initial condition  $\phi_0(x, y)$  supported inside the interval  $[-L, L] \times \Omega$ , with  $0 \leq \phi_0(x, y) \leq 1$ , there exists  $\varepsilon_0$  such that the solution of (7.10) satisfies (7.12) for all  $\varepsilon \in (0, \varepsilon_0)$ .*

The key feature that distinguishes quenching from non-quenching velocities is the absence or presence of large enough flat parts in the profile  $u(y)$ .

**Definition 7.2** *We say that the profile  $u(y) \in C^\infty(\Omega)$  satisfies the H-condition if*

$$\text{there is no point } y \in \Omega, \text{ where all derivatives of } u(y) \text{ vanish.} \quad (7.13)$$



The H-condition guarantees that the operator

$$\frac{\partial}{\partial t} + u(y)\frac{\partial}{\partial x} - \Delta_y \quad (7.14)$$

is hypoelliptic [26]. The study of existence of smooth fundamental solutions for such operators was initiated by Kolmogorov [35]. Kolmogorov's work with  $\Omega = \mathbb{R}$  and  $u(y) = y$  served in part as a motivation for the fundamental result on characterization of hypoelliptic operators of Hörmander [26]. The hypoellipticity of the operator (7.14) plays a key role in some of our considerations. The next result shows that the H-condition implies quenching.

**Theorem 7.3** *Let  $u \in C^\infty(\Omega)$  satisfy the H-condition. Then  $u(y)$  is quenching. That is, for any  $\delta > 0$  and any  $\tau > 0$  there exists a constant  $C(u, \Omega, \tau, \delta) > 0$  that is independent of  $\varepsilon \in (0, 1)$  such that*

$$\|\phi(\tau, \cdot)\|_{L^\infty(D)} \leq \delta \quad (7.15)$$

whenever the initial condition  $\phi_0(x, y)$  is supported in an interval  $[-L, L] \times \Omega$ , with  $L < C/\varepsilon$ .

More precise refinements of Theorem 7.3 can be found in [7, 34].

We now prove Theorem 7.3. Let  $\phi(t, x, y)$  be the solution of

$$\begin{aligned} \phi_t + \frac{1}{\varepsilon}u(y)\phi_x &= \Delta\phi \\ \phi(0, x, y) &= \phi_0(x, y) \\ \frac{\partial\phi}{\partial n} &= 0 \quad \text{on } \partial D. \end{aligned} \quad (7.16)$$

Let us write

$$\phi(t, x, y) = \int_{-\infty}^{\infty} dz G(t, x - z)\Psi(t, z, y),$$

with the function  $\Psi(t, x, y)$  satisfying the degenerate parabolic equation

$$\begin{aligned} \Psi_t + \frac{1}{\varepsilon}u(y)\Psi_x &= \Delta_y\Psi \\ \Psi(0, x, y) &= \phi_0(x, y) \\ \frac{\partial\Psi}{\partial n} &= 0 \quad \text{on } \partial D. \end{aligned} \quad (7.17)$$

Here,  $G(t, x)$  is the standard heat kernel

$$G(t, x) = \frac{1}{\sqrt{4\pi\kappa t}} \exp\left(-\frac{x^2}{4\kappa t}\right).$$

If  $u(y)$  satisfies the H-condition (7.13) then the diffusion process defined by (7.17) has a unique smooth transition probability density. Indeed, the Lie algebra generated by the operators  $\nabla_y$  and  $\partial_t + \varepsilon^{-2}u(y)\partial_x$  consists of vector fields of the form

$$\nabla_y, \frac{\partial}{\partial t} + u(y)\frac{\partial}{\partial x}, \frac{\partial u(y)}{\partial y_k}\frac{\partial}{\partial x}, \frac{\partial^2 u(y)}{\partial y_i \partial y_j}\frac{\partial}{\partial x}, \dots, u^{(n)}(y)\frac{\partial}{\partial x}, \dots$$

which span  $\mathbb{R}^2$  if  $u(y)$  satisfies (7.13). Then the theory of Hörmander [26], and the results of Ichihara and Kunita [25] imply that there exists a smooth transition probability density  $p_\varepsilon(t, x, y, y')$  such that

$$\Psi(t, x, y) = \int_{\mathbb{R}} dx' \int_{\Omega} dy' p_\varepsilon(t, x - x', y, y') \phi_0(x', y').$$

In particular, the function  $p_A(t)$  is uniformly bounded from above for any  $t > 0$  [25]. Then we have

$$\|\phi(t)\|_{L^\infty} \leq \|p_\varepsilon(t)\|_{L^\infty(D)} \|\phi_0\|_{L^1(D)}.$$

It is straightforward to observe that

$$p_\varepsilon(t, x, y, y') = \varepsilon p_0(t, \varepsilon x, y, y')$$

with  $p_0$  being the transition probability density for (7.17) with  $\varepsilon = 1$ . That is,  $p_0$  satisfies

$$\begin{aligned} \frac{\partial p_0}{\partial t} + u(y) \frac{\partial p_0}{\partial x} &= \Delta_y p_0, \\ p_0(0, x, y, y') &= \delta(x) \delta(y - y'), \\ \frac{\partial p_0}{\partial n} &= 0 \text{ for } x \in \partial\Omega. \end{aligned}$$

Therefore, we obtain

$$\phi(t, x, y) \leq \varepsilon \|p_0(t)\|_{L^\infty(D)} \|\phi_0\|_{L^1(D)},$$

and the conclusion of Theorem 7.3 follows.

## Non-quenching by flows with plateaus

The next result shows that a plateau in the profile  $u(y)$  prohibits quenching. Therefore, the conditions in Theorem 7.3 are natural.

**Theorem 7.4** *There exists a universal constant  $C_0 > 0$ , such that, if  $u(y) = \bar{u} = \text{const}$  in a ball  $y \in B(a, h) \subset \Omega$  for some  $a \in \Omega$  and  $h > 0$ , then there exist initial conditions supported in  $[-1, 1] \times \Omega$  such that*

$$\|\phi(t = 1, \cdot)\|_{L^\infty(D)} \geq C_0, \tag{7.18}$$

for all  $\varepsilon \in (0, 1)$ .

The proof is quite simple: solution of (7.16) is above the solution of the Dirichlet problem in the smaller channel  $D' = \{(x, y) : x \in \mathbb{R}, y \in B(a, h)\}$ :

$$\begin{aligned} \phi_t + \frac{1}{\varepsilon} u(y) \phi_x &= \Delta \phi \text{ in } D', \\ \phi(0, x, y) &= \phi_0(x, y) \\ \phi &= 0 \text{ on } \partial D'. \end{aligned} \tag{7.19}$$

However, as  $u(y) = \bar{u}$  in  $D'$ , we have

$$\phi(t, x, y) = \psi(t, x - \bar{u}\frac{t}{\varepsilon^2}, y),$$

with the function  $\psi(t, x, y)$  that solves

$$\begin{aligned} \psi_t &= \Delta\phi \text{ in } D', \\ \phi(0, x, y) &= \phi_0(x, y) \\ \psi &= 0 \text{ on } \partial D'. \end{aligned} \tag{7.20}$$

The conclusion of Theorem 7.4 follows simply from the fact that the function  $\psi$  does not depend on  $\varepsilon$ .

## 8 Particles in randomly perturbed incompressible flows

Let us now consider particles moving in an incompressible flow with stochastic perturbations, so that the underlying dynamics is governed by a stochastic differential equation

$$dX_t = -u(X)dt + \varepsilon dB_t. \tag{8.1}$$

Here,  $u(x)$  is an incompressible flow in a domain  $\Omega$ :

$$\nabla \cdot u = 0, \text{ for all } x \in \Omega, \tag{8.2}$$

and  $B_t$  is the standard Brownian motion. One example are Hamiltonian flows, and in two dimensions, if  $\Omega$  is simply connected, these are the only examples:  $u = \nabla^\perp H$ . The underlying time-dependent PDE is

$$\frac{\partial\phi}{\partial t} + u \cdot \nabla\phi = \varepsilon^2\Delta\phi, \tag{8.3}$$

and its time-rescaled version

$$\frac{\partial\phi}{\partial t} + \frac{1}{\varepsilon^2}u \cdot \nabla\phi = \Delta\phi, \tag{8.4}$$

We will assume that the flow  $u$  does not penetrate the boundary of  $\Omega$ :

$$u \cdot \nu = 0 \text{ on } \partial\Omega. \tag{8.5}$$

Here,  $\nu$  is the normal to the boundary.

Let us make one additional observation. Recall that the probabilistic interpretation for the solutions of (8.4) is as follows. Consider the solution of the stochastic differential equation

$$dX_t = -\frac{1}{\varepsilon^2}u(X_t)dt + \sqrt{2}dB_t, \quad X_0 = x. \tag{8.6}$$

Then  $\phi(t, x)$  is given by

$$\phi(t, x) = \mathbb{E}_x[\phi_0(X_{\min(t, \tau)}), \tag{8.7}$$

with the convention that  $\phi_0(X_\tau) = 0$ . Here,  $\tau$  is the first exit time from the domain  $\Omega$  for the process  $X_t$  starting at  $X_0 = x$ . Assume now that  $\Phi(x)$  is a first integral of the flow  $u(x)$ :

$$u \cdot \nabla \Phi = 0, \quad (8.8)$$

and consider  $Y_t = \Phi(X_t)$ . The process  $Y_t$  satisfies a stochastic differential equation

$$dY_t = \nabla \Phi(X_t) \cdot dX_t + \Delta \Phi(X_t) dt = \Delta \Phi(X_t) dt + \sqrt{2} \nabla \Phi(X_t) \cdot dB_t. \quad (8.9)$$

In particular, there is no large term in (8.9) – the process  $\Phi(X_t)$  remains of the order  $O(1)$ , and undergoes a slow evolution. Thus, even though the evolution of  $X_t$  is fast, the first integrals evolve slowly, meaning that  $X_t$  moves very fast on the level sets of the efforts integrals but not across the level sets.

## Oscillation on streamlines

To keep the presentation simple, we will first consider a steady version of this problem:

$$\begin{aligned} -\Delta \phi^\varepsilon + \frac{1}{\varepsilon^2} u \cdot \nabla \phi^\varepsilon &= g(x), \\ \phi^\varepsilon(x) &= 0, \quad x \in \partial\Omega. \end{aligned} \quad (8.10)$$

Here,  $\Omega$  is a bounded smooth domain. First, note that multiplying by  $\phi^\varepsilon(x)$  and integrating by parts gives

$$\int_{\Omega} |\nabla \phi^\varepsilon|^2 dx = \int_{\Omega} f \phi^\varepsilon dx \leq \|g\|_{L^2} \|\phi^\varepsilon\|_{L^2}. \quad (8.11)$$

The Poincaré inequality implies that

$$\|\phi\|_{L^2} \leq C_p \|\nabla \phi^\varepsilon\|_{L^2}. \quad (8.12)$$

It follows from (8.11) and (8.12) that

$$\|\phi\|_{L^2} \leq C \|\nabla \phi^\varepsilon\|_{L^2} \leq C \|g\|_{L^2}. \quad (8.13)$$

Next, we multiply (8.10) by  $u \cdot \nabla \phi^\varepsilon$  and integrate to get

$$\begin{aligned} \int_{\Omega} |u \cdot \nabla \phi^\varepsilon|^2 dx &= \varepsilon^2 \int_{\Omega} (u \cdot \nabla \phi^\varepsilon) \Delta \phi^\varepsilon dx + \varepsilon^2 \int_{\Omega} g(x) (u \cdot \nabla \phi) \\ &\leq -\varepsilon^2 \int_{\Omega} \nabla(u \cdot \nabla \phi^\varepsilon) \cdot \nabla \phi^\varepsilon dx + \frac{\varepsilon^2}{2} \|g\|_{L^2}^2 + \frac{\varepsilon^2}{2} \int_{\Omega} |u \cdot \nabla \phi^\varepsilon|^2 dx. \end{aligned} \quad (8.14)$$

We rewrite the integrand in the second line above as:

$$\nabla(u \cdot \nabla \phi^\varepsilon) \cdot \nabla \phi^\varepsilon = \frac{\partial u_k}{\partial x_i} \frac{\partial \phi^\varepsilon}{\partial x_k} \frac{\partial \phi^\varepsilon}{\partial x_i} + u_k \frac{\partial^2 \phi^\varepsilon}{\partial x_k \partial x_i} \frac{\partial \phi^\varepsilon}{\partial x_i} = \frac{\partial u_k}{\partial x_i} \frac{\partial \phi^\varepsilon}{\partial x_k} \frac{\partial \phi^\varepsilon}{\partial x_i} + \frac{1}{2} u \cdot \nabla (|\nabla \phi^\varepsilon|^2). \quad (8.15)$$

Once again using incompressibility of  $u$ , we obtain from the above

$$\begin{aligned} - \int_{\Omega} \nabla(u \cdot \nabla \phi^\varepsilon) \cdot \nabla \phi^\varepsilon dx &= \frac{1}{2} \int_{\Omega} (u \cdot \nabla (|\nabla \phi^\varepsilon|^2)) dx - \int_{\Omega} \frac{\partial u_n}{\partial x_m} \frac{\partial \phi^\varepsilon}{\partial x_m} \frac{\partial \phi^\varepsilon}{\partial x_n} dx \\ &\leq M \varepsilon^2 \int_{\Omega} |\nabla \phi^\varepsilon|^2 dx \leq CM \varepsilon^2 \|g\|_{L^2}, \end{aligned} \quad (8.16)$$

where  $M = \|\nabla u\|_{L^\infty(\Omega)}$ . We deduce that

$$\int_{\Omega} |u \cdot \nabla \phi^\varepsilon|^2 dx \leq C\varepsilon^2 \|g\|_{L^2}^2. \quad (8.17)$$

Informally, this means that the oscillation of  $\phi^\varepsilon$  along the stream lines of  $u$  is small. There are, of course, ways to make this more precise but it says, roughly, that as  $\varepsilon \rightarrow 0$  the function  $\phi^\varepsilon(x)$  converges to a limit  $\bar{\phi}(x)$  which is constant on the streamlines of  $u$ .

## An unfortunate toy example: a radially symmetric Hamiltonian

Let us first consider a special situation when the Hamiltonian  $H(x, y) = (x^2 + y^2)/2$ , so that

$$u(x, y) = \nabla^\perp H(x, y) = (H_y, -H_x) = (y, -x),$$

and (8.1) becomes

$$\begin{aligned} dX_t &= Y_t dt + \varepsilon dB_t^{(1)}, \\ dY_t &= -X_t dt + \varepsilon dB_t^{(2)}. \end{aligned} \quad (8.18)$$

The time rescaling  $t \rightarrow t/\varepsilon^2$  leads to

$$\begin{aligned} dX_t &= \frac{1}{\varepsilon^2} Y_t dt + dB_t^{(1)}, & X_0 &= x, \\ dY_t &= -\frac{1}{\varepsilon^2} X_t dt + dB_t^{(2)}, & Y_0 &= y. \end{aligned} \quad (8.19)$$

The corresponding PDE is

$$\begin{aligned} \frac{\partial v}{\partial t} &= \frac{y}{\varepsilon^2} \frac{\partial v}{\partial x} - \frac{x}{\varepsilon^2} \frac{\partial v}{\partial y} + \frac{1}{2} \Delta v, \\ v(0, x) &= v_0(x), \end{aligned} \quad (8.20)$$

in the sense that

$$v(t, x) = \mathbb{E}_{x,y}[v_0(X_t, Y_t)]. \quad (8.21)$$

Switching to the polar coordinates  $x = r \cos \theta$ ,  $y = r \sin \theta$  gives

$$\begin{aligned} \frac{\partial v}{\partial t} &= -\frac{1}{\varepsilon^2} \frac{\partial v}{\partial \theta} + \frac{1}{2} \left[ \frac{\partial^2 v}{\partial r^2} + \frac{1}{r} \frac{\partial v}{\partial r} + \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} \right], \\ v(0, r, \theta) &= v_0(r, \theta). \end{aligned} \quad (8.22)$$

We see that the average along the streamlines

$$\bar{v}(t, r) = \frac{1}{2\pi} \int_0^{2\pi} v(t, r, \theta) d\theta$$

satisfies a parabolic equation

$$\begin{aligned} \frac{\partial \bar{v}}{\partial t} &= \frac{1}{2} \left[ \frac{\partial^2 \bar{v}}{\partial r^2} + \frac{1}{r} \frac{\partial \bar{v}}{\partial r} \right], \\ \bar{v}(0, r) &= \bar{v}_0(r). \end{aligned} \quad (8.23)$$

The corresponding process  $R_t = (X_t^2 + Y_t^2)^{1/2}$  converges to a diffusion with the generator as in (8.23). There is an additional feature here, specific to the quadratic Hamiltonian: we may write

$$v(t, r, \theta) = w(t, r, \theta - \frac{t}{\varepsilon^2}). \quad (8.24)$$

The function  $w$  satisfies

$$\begin{aligned} \frac{\partial w}{\partial t} &= \frac{1}{2} \left[ \frac{\partial^2 w}{\partial r^2} + \frac{1}{r} \frac{\partial w}{\partial r} + \frac{1}{r^2} \frac{\partial^2 w}{\partial \theta^2} \right], \\ w(0, r, \theta) &= v_0(r, \theta). \end{aligned} \quad (8.25)$$

In other words, by factoring out the background dynamics –a fast rotation, we arrive exactly at a diffusion equation, without any need to pass to the limit  $\varepsilon \rightarrow 0$ .

This example may look somewhat unfortunate since we do not see the uniformization along the streamlines that we have observed in the general steady version of this problem. To see it, one should consider the time-averages of the solution. In other words, the weak limit (as a function in time) of the function  $v(t, r, \theta)$  in (8.24) is

$$\bar{v}(t, r) = \frac{1}{2\pi} \int_0^{2\pi} w(t, r, \theta) d\theta, \quad (8.26)$$

and is independent of  $\theta$ .

## The Freidlin problem in two dimensions: one cell

Let us now consider the more general two-dimensional case, with  $u(x, y) = \nabla^\perp H(x, y)$ . We first assume that the Hamiltonian  $H(x, y)$  is a convex function growing at infinity, with a minimum at some point  $(x_0, y_0) \in \mathbb{R}^2$ . We denote  $H_0 = H(x_0, y_0)$ . The assumption that  $u$  is parallel to the boundary of  $\Omega$  means that  $\partial\Omega$  is a level set of  $H(x, y)$ . As we have seen, solutions of (8.10)

$$\begin{aligned} -\Delta \phi^\varepsilon + \frac{1}{\varepsilon^2} u \cdot \nabla \phi^\varepsilon &= g(x), \\ \phi^\varepsilon(x) &= 0, \quad x \in \partial\Omega. \end{aligned} \quad (8.27)$$

are uniformly bounded in  $L^\infty(\Omega) \cap H_0^1(\Omega)$ :

$$0 \leq \phi^\varepsilon \leq C, \quad \int |\nabla \phi^\varepsilon|^2 dx \leq C, \quad (8.28)$$

with the constant  $C > 0$  independent of  $\varepsilon > 0$ . Hence, the family  $\phi_\varepsilon$  converges weakly in  $H^1(\Omega)$  (after extracting a subsequence) and strongly in  $L^2(\Omega)$  to a function  $\bar{\phi}$ .

We claim that  $\bar{\phi}$  depends only on the variable  $h = H(x, y)$ . Indeed, if we multiply (8.27) by  $\varepsilon^2$  and let  $\varepsilon \rightarrow +\infty$ , we get

$$u \cdot \nabla \bar{\phi} = 0 \quad (8.29)$$

in the sense of distributions. It is convenient to introduce the curvilinear coordinates  $(h, \theta)$ . The coordinates are chosen so that  $h(x, y) = H(x, y)$ , and the streamlines of the flow  $u(x, y)$  are  $\{h = \text{const}\}$ . The level lines of the coordinate  $\theta = \Theta(x, y)$  are orthogonal to the flow lines:

$$\nabla \Theta \cdot \nabla H = 0.$$

We normalize  $\theta$  so that  $0 \leq \theta \leq 2\pi$ . As we have mentioned, the boundary  $\partial\Omega$  is a level set

$$\partial\Omega = \{h = h_0\},$$

and we will assume without loss of generality that  $h_0 = 0$ . Then (8.29) implies that  $\bar{\phi}$  depends only on the variable  $h$ . The  $L^\infty$ -bound in (8.28) implies that

$$0 \leq \bar{\phi}(h) \leq C.$$

In addition, we have

$$\int |\nabla_x \bar{\phi}|^2 dx = \int |\bar{\phi}_h|^2 |\nabla h|^2 dx = \int_0^{H_0} |\bar{\phi}_h|^2 \left( \int_0^{2\pi} \frac{|\nabla H|^2}{J} d\theta \right) dh.$$

Here  $J = H_y \Theta_x - H_x \Theta_y$  is the Jacobian of the coordinate change. Note that  $\nabla \Theta = \rho \nabla^\perp H$  with some scalar function  $\rho > 0$ , so that

$$J = \rho |\nabla H|^2, \quad |\nabla \Theta| = \rho |\nabla H| \text{ and } dl = d\theta / |\nabla \Theta|.$$

Therefore, we have

$$\int_0^{2\pi} \frac{|\nabla H|^2}{J} d\theta = \oint_{H(x,y)=h} |\nabla H| dl := a(h), \quad (8.30)$$

and thus we have a weighted  $H^1$ -bound

$$\int_0^{H_0} a(h) |\bar{\phi}_h|^2 dh < +\infty,$$

which follows from (8.28), and hence  $\bar{\phi}(h)$  is continuous for  $h < H_0$ , as  $p(h) \sim C(H_0 - h)$  for  $h$  close to  $H_0$ .

Next, we re-write (8.27) in the curvilinear coordinates:

$$\begin{aligned} & -\frac{|\nabla H|^2}{J} \frac{\partial^2 \phi^\varepsilon}{\partial h^2} - \frac{|\nabla \Theta|^2}{J} \frac{\partial^2 \phi^\varepsilon}{\partial \theta^2} - \frac{(\Delta H)}{J} \frac{\partial \phi^\varepsilon}{\partial h} - \frac{(\Delta \Theta)}{J} \frac{\partial \phi^\varepsilon}{\partial \theta} + \frac{1}{\varepsilon^2} \frac{\partial \phi^\varepsilon}{\partial \theta} = \frac{1}{J} g(h, \theta), \quad (8.31) \\ & \phi^\varepsilon(0, \theta) = 0, \quad \phi^\varepsilon(h, \theta) \text{ is bounded for } 0 \leq h \leq H_0. \end{aligned}$$

Integrating this equation in  $\theta$  and passing to the limit  $\varepsilon \rightarrow 0$  we obtain the limit problem for the function  $\bar{\phi}$ :

$$\begin{aligned} & -a(h) \bar{\phi}''(h) - b(h) \bar{\phi}'(h) = \bar{g}(h), \quad (8.32) \\ & \bar{\phi}(0) = 0, \quad \bar{\phi}(h) \text{ is bounded for } 0 \leq h \leq H_0, \end{aligned}$$

with  $a(h)$  as in (8.30), and

$$b(h) = \int_0^{2\pi} \frac{\Delta H}{J} d\theta, \quad \bar{g}(h) = \int_0^{2\pi} \frac{g(h, \theta) d\theta}{J}.$$

The problem (8.32) is called the Freidlin problem. We note that

$$b(h) = \int_0^{2\pi} \frac{\Delta H}{J} d\theta = \oint_{H(x,y)=h} \frac{\Delta H}{|\nabla H|} dl = a'(h).$$

The last equality above follows from the fact that

$$a(h) = \oint_{H(x,y)=h} |\nabla H| dl = \int_{G_h} \Delta H dx dy.$$

Here  $G_h = \{h \leq H(x, y) \leq H_0\}$  is the interior of the streamline  $\{H(x, y) = h\}$ . Hence, the Freidlin problem (8.32) can be re-written in a self-adjoint form as

$$\begin{aligned} -\frac{d}{dh} \left( a(h) \frac{d\bar{\phi}(h)}{dh} \right) &= \bar{g}(h), \\ \bar{\phi}(0) &= 0, \quad \bar{\phi}(h) \text{ is bounded for } 0 \leq h \leq H_0. \end{aligned} \tag{8.33}$$

### The Freidlin problem: gluing conditions

Let us now consider a more general situation when the function  $H(x, y)$  may have many critical points. We will assume that each level set can contain just one critical point. Then the limiting diffusion is defined on the Reeb graph of the function  $H(x, y)$ . The Reeb graph can be informally described as follows. Its vertices correspond to the level sets of  $H$  containing the saddle points  $\bar{x}_1, \dots, \bar{x}_N$  of  $H$ . The level sets containing a saddle point  $\bar{x}_k$  of  $H(x)$  are topologically "figure eights", with one critical point of  $H(x)$  inside each of the two "circles"  $C_{k1}$  and  $C_{k2}$ . The level sets inside each of  $C_{k1}$  and  $C_{k2}$  correspond to an edge of the Reeb graph,  $e_{k1}$  and  $e_{k2}$ . These two edges are joined at the vertex corresponding to  $\bar{x}_k$ . Each edge is parametrized by the values of  $H$  inside the corresponding "circle". The limit function  $\bar{\phi}(h)$  satisfies the Freidlin problem (8.33) along each edge, with the coefficients  $a(h)$  computed as in (8.30), inside the cell of the flow that corresponds to that edge.

Let us now obtain the gluing conditions at a vertex  $\bar{x}_k$ , where three edges  $e_{out}$  and  $e_{i1}$ ,  $e_{i2}$  join, corresponding to the "outside" region  $C_{out,k}$ , and two inside circles  $C_{k1}$  and  $C_{k2}$ . We integrate (8.27)

$$-\Delta \phi^\varepsilon + \frac{1}{\varepsilon^2} u \cdot \nabla \phi^\varepsilon = g(x), \tag{8.34}$$

over a domain bounded by a "just outside circle"  $C_{out}$  and two just circles  $C_{k1}$  and  $C_{k2}$ . It follows that, to the leading order in the thickness of this annular region, we have

$$\oint_{C_{out}} \frac{\partial \phi}{\partial n} dl = \oint_{C_{k1}} \frac{\partial \phi}{\partial n} dl + \oint_{C_{k2}} \frac{\partial \phi}{\partial n} dl + \text{l.o.t.} \tag{8.35}$$

Note that, since all contours in (8.35) are level sets, we have

$$\frac{\partial \phi}{\partial n} \approx \frac{\partial \bar{\phi}}{\partial h} |\nabla H|.$$

Using this in (8.35) leads to

$$\frac{\partial \bar{\phi}_{out}}{\partial h} \oint_{C_{out}} |\nabla H| dl = \frac{\partial \bar{\phi}_{k,1}}{\partial h} \oint_{C_{k1}} |\nabla H| dl + \frac{\partial \bar{\phi}_{k,2}}{\partial h} \oint_{C_{k2}} |\nabla H| dl. \tag{8.36}$$



In other words, the gluing condition at vertex  $k$  corresponding to the level set  $H(x, y) = H_k$ , is:

$$a_{k,out} \frac{\partial \bar{\phi}_{out}}{\partial h}(H_k) = a_{k,in1} \frac{\partial \bar{\phi}_{k,1}}{\partial h}(H_k) + a_{k,in1} \frac{\partial \bar{\phi}_{k,1}}{\partial h}(H_k). \quad (8.37)$$

These gluing conditions together with the Poisson equations (8.33) completely describe the limit problem.

## Particles in a random force field: “short” times

We now assume that the potential  $V(t, x)$  is random, weak and varies on the scale much larger than the initial data. More precisely, we consider the semiclassical Schrödinger equation

$$i\varepsilon\phi_t + \frac{\varepsilon^2}{2}\Delta\phi - \delta V(x)\phi = 0 \quad (8.38)$$

with the  $\varepsilon$ -oscillatory initial data  $\phi(0, x) = \phi_0^\varepsilon(x)$ . This equation is written on the scale of the variations of the random potential, and  $\delta \ll 1$  is the parameter measuring its strength. Passing to the high frequency limit  $\varepsilon \rightarrow 0$  we obtain the Liouville equation for the Wigner measure of the family  $\phi_\varepsilon(t, x)$ :

$$\frac{\partial W}{\partial t} + k \cdot \nabla_x W - \delta \nabla_x V(x) \cdot \nabla_k W = 0, \quad (8.39)$$

with the initial data  $W(0, x, k) = W_0(x, k)$ , the Wigner measure of the family  $\phi_0^\varepsilon(x)$ . As the parameter  $\delta \ll 1$  is small, the effect of the randomness will be felt only after long times. We will build our analysis of (8.39) slowly, starting with relatively short times, and later for the long times. We will assume that  $V(x)$  is a spatially homogeneous random process with mean zero and the correlation function  $R(x)$ :

$$\langle V(x) \rangle = 0, \quad R(x) = \langle V(y)V(x+y) \rangle. \quad (8.40)$$

It will be convenient for us to use the correlation matrix for the force  $\nabla V$ :

$$\left\langle \frac{\partial V(y)}{\partial y_i} \frac{\partial V(x+y)}{\partial y_j} \right\rangle = -\frac{\partial^2 R(x)}{\partial x_i \partial x_j}. \quad (8.41)$$

### The characteristics at short times

We begin with the very basic theory of characteristics in a weakly random medium – this material originated in the classical paper by J.B. Keller [29]. The characteristics for the Liouville equation (8.39) are

$$\frac{dX}{dt} = -K(t), \quad \frac{dK}{dt} = \delta \nabla V(X(t)), \quad X(0) = x, \quad K(0) = k. \quad (8.42)$$

Let us seek the trajectories  $X(t), K(t)$  as a formal perturbation expansion

$$X(t) = X_0(t) + \delta X_1(t) + \delta^2 X_2(t) + \dots, \quad K(t) = K_0(t) + \delta K_1(t) + \delta^2 K_2(t) + \dots$$

We insert this expansion into the characteristics (8.42), and get in the leading order:

$$X_0(t) = x - k_0 t, \quad K_0(t) = k.$$

As expected, in the leading order the characteristics are straight lines. The first order correction in  $\delta$  is

$$K_1(t) = \int_0^t \nabla V(X_0(s)) ds = \int_0^t \nabla V(x - ks) ds, \quad (8.43)$$

and

$$X_1(t) = \int_0^t K_1(s) ds = \int_0^t (t-s) \nabla V(x - ks) ds. \quad (8.44)$$

Naively, in order to see how long this approximation should hold, we estimate that during a time  $T$  we would get  $K_1(T) \sim T$ , and  $X_1(T)$  of the order  $T^2$  meaning that we would need  $\delta T^2 \ll 1$ , or  $T \ll \delta^{-1/2}$  for the spatial trajectory to stay close to the straight line. Let us now see how randomness affects this ballpark estimate – we have, as in (1.30):

$$\begin{aligned} \langle K_1^2(t) \rangle &= \int_0^t \int_0^t \langle \nabla V(x - ks) \cdot \nabla V(x - ks') \rangle ds ds' \\ &= - \int_0^t \int_0^t \Delta R(k(s-s')) ds ds' = Dt + O(1), \quad \text{as } t \rightarrow +\infty, \end{aligned}$$

with the diffusion coefficient

$$D = - \int_{-\infty}^{\infty} \Delta R(ks) ds. \quad (8.45)$$

With a little bit more work, one can show that an appropriate rescaling of  $K_1(t)$  converges to a Brownian motion with the diffusion matrix

$$D_{ij} = - \int_{-\infty}^{\infty} \frac{\partial^2 R(ks)}{\partial x_i \partial x_j} ds. \quad (8.46)$$

The variance of  $X_1(t)$  can also be computed explicitly:

$$\begin{aligned} \langle X_1^2(t) \rangle &= \int_0^t \int_0^t (t-s)(t-s') \langle \nabla V(x - ks) \cdot \nabla V(x - ks') \rangle ds ds' \\ &= - \int_0^t \int_0^t (t-s)(t-s') \Delta R(k(s-s')) ds ds' = \frac{Dt^3}{3} + O(1), \quad \text{as } t \rightarrow +\infty, \end{aligned}$$

and, once again, with a bit more work it can be shown that an appropriate rescaling of  $X(t)$  converges, at large times to the time integral of the Brownian motion with the diffusion matrix  $D_{ij}$ . The above computations indicate that the simple perturbation expansion should hold for times  $T$  such that

$$\delta^2 T^3 \sim O(1),$$

that is, for times of the order  $T \sim \delta^{-2/3}$ , which is much longer than the “deterministic prediction”  $T \sim \delta^{-1/2}$ .

Formally, this means that for large times (but much smaller than  $\delta^{-2/3}$ ), the expected value of the solutions of the Liouville equation (8.39) is well-approximated by the solutions of the Fokker-Planck kinetic equation

$$\frac{\partial \bar{W}}{\partial t} + k \cdot \nabla_x \bar{W} = \delta^2 \sum_{i,j=1}^n D_{ij} \frac{\partial^2 \bar{W}}{\partial k_i \partial k_j}, \quad (8.47)$$

that is,  $\langle W(t, x, k) \rangle \approx \bar{W}(t, x, k)$ . This is probably the simplest way to get to a kinetic description of waves in random media. Instead of trying to make this approximation result precise, for times  $t \ll \delta^{-2/3}$ , let us explain why such result, while providing a very nice “hooligan’s derivation of the kinetic limit”, can not “truly hold” for longer times, when the deviation of the characteristics from straight lines will be not small. The problem is that the original characteristics (8.42) preserve the classical Hamiltonian:

$$\omega(x, k) = \frac{k^2}{2} + V(x),$$

that is,  $\omega(X(t), K(t)) = \omega(X(0), K(0))$ . In particular, if, say,  $V(x)$  is a bounded random potential, it is impossible for  $K(t)$  to behave as a Brownian motion for large times. Nevertheless, the overall picture described above is not too wrong, and in the next step we will see how it can be naturally modified to see what happens at large times.

## 8.1 Random geometric optics: the long time limit

### A particle in a random Hamiltonian

We will now study the “truly” long time asymptotics of geometric optics in a weakly random medium. This problem can be analyzed in the general setting of a particle in a weakly random Hamiltonian field:

$$\frac{dX^\delta}{dt} = \nabla_k H_\delta, \quad \frac{dK^\delta}{dt} = -\nabla_x H_\delta, \quad X^\delta(0) = 0, \quad K^\delta(0) = k_0, \quad (8.48)$$

with a random Hamiltonian of the form  $H_\delta(x, k) = H_0(k) + \delta H_1(x, k)$ . Here  $H_0(k)$  is the background Hamiltonian and  $H_1(x, k)$  is a random perturbation, while the small parameter  $\delta \ll 1$  measures the relative strength of random fluctuations. This was done in [3] and [37]. Here, we will resist the temptation to describe the general results, and restrict ourselves to the case at hand, with  $H_0(k) = |k|^2/2$  and  $H_1(x, k) = V(x)$ , which simplifies some considerations. Thus, we are interested in the Liouville equations

$$\frac{\partial \phi}{\partial t} + k \cdot \nabla_x \phi - \delta \nabla V(x) \cdot \nabla_k \phi = 0, \quad (8.49)$$

and the corresponding characteristics

$$\frac{dX}{dt} = K, \quad \frac{dK}{dt} = -\delta \nabla_x V(X), \quad X(0) = 0, \quad K(0) = k_0, \quad (8.50)$$

on the time scale  $t \sim \delta^{-2}$ . As usual, we will assume that the random potential  $V(x)$  is a mean-zero statistically homogeneous random field, with a rapidly decaying correlation function  $R(x)$ :

$$\langle V(x) \rangle = 0, \quad \langle V(y)V(x+y) \rangle = R(x), \quad (8.51)$$

We have already seen that at relatively short times  $t \ll \delta^{-2/3}$  the “boosted” deviation  $(K(t) - k_0)/\delta$  behaves as a Brownian motion. At the longer times, we are interested not in the deviation from the original direction but in the particle momentum itself. An important simple observation is that (8.50) preserves the Hamiltonian

$$H(x, k) = \frac{k^2}{2} + \delta V(x). \quad (8.52)$$

Hence, the law of any possible limit for the process  $K_\delta(t) = K(t/\delta^2)$ , as  $\delta \rightarrow 0$ , has to be supported on the sphere  $|K(t)| = |k_0|$  (and can not be a regular Brownian motion). Moreover, one would expect the law of the limit process to be isotropic – there is no preferred direction in the problem. One possibility is that  $K_\delta(t)$  tends to a uniform distribution on the sphere  $\{|k| = |k_0|\}$  – and this is, indeed, what happens at times  $t \gg \delta^{-2}$ . However, at an intermediate stage, at times of the order  $\delta^{-2}$ , the process  $K_\delta(t)$  converges to the Brownian motion  $B_s(t)$  on the sphere (this is an isotropic diffusion such that  $|B_s(t)| = 1$  for all  $t$ ). This intuitive result has been first proved in [31] in dimensions higher than two, and later extended to two dimensions with the Poisson distribution of scatterers in [12], and in a general two-dimensional setting in [38]. The rescaled spatial component  $X^\delta(t) = \delta^2 X(t/\delta^2)$  converges to the time integral of the Brownian motion on the sphere:

$$X(t) = \int_0^t B_s(\tau) d\tau.$$

In turn, the long time limit of a momentum diffusion is the standard spatial Brownian motion, and we will see that on the times longer than  $\delta^{-2}$  the spatial component  $X(t)$  converges to the Brownian motion, while  $K(t)$  becomes uniformly distributed on the sphere  $\{|k| = |k_0|\}$ .

Let us mention that another important, (in the context of waves in random media) Hamiltonian

$$H_\delta(x, k) = (c_0 + \delta c_1(x))|k|, \quad (8.53)$$

arises in the geometrical optics limit of the wave equation. We will not address it directly here, but, as we have mentioned, the analysis of the classical Hamiltonian (8.52) can be generalized in a relatively straightforward way – see [37] for details. We stick here with (8.52) solely for the sake of simplicity of presentation.

### The Fokker-Planck limit

Let the function  $\phi_\delta(t, x, k)$  satisfy the Liouville equation

$$\begin{aligned} \frac{\partial \phi^\delta}{\partial t} + k \cdot \nabla_x \phi^\delta - \delta \nabla V(x) \cdot \nabla_k \phi^\delta &= 0, \\ \phi^\delta(0, x, k) &= \phi_0(\delta^2 x, k). \end{aligned} \quad (8.54)$$

There are two assumptions implicitly made here: first is that the random potential is weak, and the second is that the initial data varies on the scale  $1/\delta^2$  relative to the scale of the variations of the potential. In the terminology of the introduction, this means that  $l_c/L = \delta^2$  – or, we choose the particular observation scale  $L = l_c/\delta^2$ . One may wonder also as to what happens on other observation scales – we will address this further below.

Let us define the diffusion matrix  $D_{mn}$  by

$$D_{ml}(k) = -\frac{1}{|k|} \int_{-\infty}^{\infty} \frac{\partial^2 R(s\hat{k})}{\partial x_n \partial x_m} ds, \quad m, l = 1, \dots, n. \quad (8.55)$$

Note that if the correlation function is isotropic:  $R = R(|x|)$ , then  $D_{mn}$  has a particularly simple form:

$$D_{ml}(k) = D(\delta_{mn} - \hat{k}_l \hat{k}_m), \quad D = -\frac{2}{|k|} \int_0^{\infty} \frac{R'(r)}{r} dr, \quad m, l = 1, \dots, n. \quad (8.56)$$

We have the following result.

**Theorem 8.1** *Let  $\phi^\delta$  be the solution of (8.54), with the initial data  $\phi_0 \in C_c^\infty(\mathbb{R}^{2d})$ , whose support is contained inside a spherical shell  $\mathcal{A}(M) = \{(x, k) : M^{-1} < |k| < M\}$  for some positive  $M > 0$ , and let  $\bar{\phi}$  satisfy*

$$\begin{aligned} \frac{\partial \bar{\phi}}{\partial t} + k \cdot \nabla_x \bar{\phi} &= \sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left( D_{mn}(k) \frac{\partial \bar{\phi}}{\partial k_n} \right) \\ \bar{\phi}(0, x, k) &= \phi_0(x, k). \end{aligned} \quad (8.57)$$

*Suppose that  $M \geq M_0 > 0$  and  $T \geq T_0 > 0$ . Then, there exist two constants  $C, \alpha_0 > 0$  such that for all  $T \geq T_0$*

$$\sup_{(t,x,k) \in [0,T] \times K} \left| \mathbb{E} \phi^\delta \left( \frac{t}{\delta^2}, \frac{x}{\delta^2}, k \right) - \bar{\phi}(t, x, k) \right| \leq CT(1 + \|\phi_0\|_{1,4}) \delta^{\alpha_0} \quad (8.58)$$

*for all compact sets  $K \subset \mathcal{A}(M)$ .*

Note that

$$\sum_{m=1}^d D_{nm}(\hat{k}, k) \hat{k}_m = -\sum_{m=1}^d \frac{1}{2|k|} \int_{-\infty}^{\infty} \frac{\partial^2 R(s\hat{k})}{\partial x_n \partial x_m} \hat{k}_m ds = -\sum_{m=1}^d \frac{1}{2|k|} \int_{-\infty}^{\infty} \frac{d}{ds} \left( \frac{\partial R(s\hat{k})}{\partial x_n} \right) ds = 0$$

and thus the  $K$ -process generated by (8.57) is indeed a diffusion process on a sphere  $|k| = \text{const}$ , or, equivalently, equations (8.57) for different values of  $|k|$  are decoupled. Another important point is that the assumption that the initial data does not concentrate close to  $k = 0$  is important – if  $|k|$  is very small, the particle moves very slowly, and does not have a sufficient time to sample enough of the random medium by the time  $\delta^{-2}$ .

## Beyond the Fokker-Planck limit

Let us now return to the question of what happens to the solutions of the Liouville equation with the initial data that varies on a scale much longer than  $\delta^{-2}$  – in other words, the observation is taken on even larger scales than described by the Fokker-Planck limit. It is straightforward to see that solutions of the Fokker-Planck equation (8.57) themselves converge in the long time limit to the solutions of the spatial diffusion equation. More, precisely, we have the following result. Let  $\bar{\phi}_\gamma(t, x, k) = \bar{\phi}(t/\gamma^2, x/\gamma, k)$ , where  $\bar{\phi}$  satisfies (8.57) with slowly varying initial data  $\bar{\phi}_\gamma(0, t, x, k) = \phi_0(\gamma x, k)$ . We also let  $w(t, x, |k|)$  be the solution of the spatial diffusion equation:

$$\begin{aligned} \frac{\partial w}{\partial t} &= \sum_{m,n=1}^d a_{mn}(|k|) \frac{\partial^2 w}{\partial x_n \partial x_m}, \\ w(0, x, |k|) &= \bar{\phi}_0(x, |k|) \end{aligned} \quad (8.59)$$

with the averaged initial data

$$\bar{\phi}_0(x, k) = \frac{1}{\Gamma_{n-1}} \int_{\mathbb{S}^{n-1}} \phi_0(x, k) d\Omega(\hat{k}).$$

Here,  $d\Omega(\hat{k})$  is the surface measure on the unit sphere  $\mathbb{S}^{n-1}$  and  $\Gamma_n$  is the area of an  $n$ -dimensional sphere. The diffusion matrix  $A := [a_{nm}]$  in (8.59) is given explicitly as

$$a_{ij}(k) = \frac{|k|^2}{\Gamma_{n-1}} \int_{\mathbb{S}^{n-1}} \hat{k}_i \chi_j(k) d\Omega(\hat{k}). \quad (8.60)$$

The functions  $\chi_j$  appearing above are the mean-zero solutions of

$$\sum_{m,i=1}^d \frac{\partial}{\partial k_m} \left( D_{mi}(k) \frac{\partial \chi_j}{\partial k_i} \right) = -\hat{k}_j, \quad (8.61)$$

and when the correlation function  $R(x)$  is isotropic, so that  $D_{mi}$  is given by (8.56), they are just multiples of  $\hat{k}_j$ :  $a_j(k) = c(|k|)\hat{k}_j$ , with an appropriate constant  $c(|k|)$  that can be computed explicitly. In that case, the matrix  $a_{nm}$  is a multiple of identity, and (8.59) becomes the standard diffusion equation

$$\frac{\partial w}{\partial t} = \bar{a}(|k|) \Delta_x w, \quad (8.62)$$

with an appropriate diffusion constant  $\bar{a}$ .

**Theorem 8.2** *For every pair of times  $0 < T_* < T < +\infty$  the re-scaled solution  $\bar{\phi}_\gamma(t, x, k) = \bar{\phi}(t/\gamma^2, x/\gamma, k)$  of (8.57) converges as  $\gamma \rightarrow 0$  in  $C([T_*, T]; L^\infty(\mathbb{R}^{2d}))$  to  $w(t, x, k)$ . Moreover, there exists a constant  $C_0 > 0$ , so that we have*

$$\|w(t, \cdot) - \bar{\phi}_\gamma(t, \cdot)\|_{L^\infty} \leq C_0 (\gamma T + \sqrt{\gamma}) \|\phi_0\|_{C^1}, \quad (8.63)$$

for all  $T_* \leq t \leq T$ .

The proof of Theorem 8.2 is based on classical asymptotic expansions and is quite straightforward. As an immediate corollary of Theorems 8.1 and 8.2, we obtain the following result.

**Theorem 8.3** *Let  $\phi_\delta$  be solution of (8.54) with the initial data  $\phi_\delta(0, x, k) = \phi_0(\delta^{2+\alpha}x, k)$  and let  $\bar{w}(t, x)$  be the solution of the diffusion equation (8.59) with the initial data  $w(0, x, k) = \bar{\phi}_0(x, k)$ . Then, there exists  $\alpha_0 > 0$  and a constant  $C > 0$  so that for all  $0 \leq \alpha < \alpha_0$  and all  $0 < T_* \leq T$  we have for all compact sets  $K \subset \mathcal{A}(M)$ :*

$$\sup_{(t,x,k) \in [T_*, T] \times K} |w(t, x, k) - \mathbb{E}\bar{\phi}_\delta(t, x, k)| \leq CT\delta^{\alpha_0 - \alpha}, \quad (8.64)$$

where  $\bar{\phi}_\delta(t, x, k) := \phi_\delta(t/\delta^{2+2\alpha}, x/\delta^{2+\alpha}, k)$ .

Theorem 8.3 shows that if the initial data varies on a scale slightly larger than  $\delta^{-2}$  then we observe spatial diffusion for the solution (and uniform distribution in  $k$ ) on the appropriate time scale. The requirement that  $\alpha$  is small is most likely technical and a constraint of a “perturbative” proof – the result should hold for any  $\alpha > 0$ .

To summarize: if the initial data for the random Liouville equation

$$\frac{\partial \phi}{\partial t} + k \cdot \nabla_x \phi - \delta V(x) \cdot \nabla_x \phi = 0, \quad (8.65)$$

varies on the scale  $\delta^{-2}$ :  $\phi(0, x) = \phi_0(\delta^2 x, k)$ , then on the time scale  $t \sim \delta^{-2}$  the expectation of the rescaled solution  $\phi_\delta(t, x, k) = \phi(t/\delta^2, x/\delta^2, k)$  converges to the solution of the Fokker-Planck equation. On the other hand, if the initial data varies on an even larger scale:  $\phi(0, x, k) = \phi(\delta^{2+\alpha}x, k)$  then on the time scale  $t \sim \delta^{-2-2\alpha}$  the expectation of the rescaled field  $\phi_\delta(t, x, k) = \phi(t/\delta^{2+2\alpha}, x/\delta^{2+\alpha}, k)$  converges to the solution of the spatial diffusion equation and is uniformly distributed in the directions  $\hat{k}$  for each  $|k|$  fixed. Thus, the appropriate kinetic limit depends on the scale of the probing signal, which, in turn, determines the proper time scale of the observations.

## A formal derivation of the momentum diffusion

We now describe how the momentum diffusion operator in (8.57) can be derived in a quick formal way. We represent the solution of (8.54) as  $\phi^\delta(t, x, k) = \psi^\delta(\delta^2 t, \delta^2 x, k)$  and write an asymptotic multiple scale expansion for  $\psi^\delta$

$$\psi^\delta(t, x, k) = \bar{\phi}(t, x, k) + \delta \phi_1\left(t, x, \frac{x}{\delta^2}, k\right) + \delta^2 \phi_2\left(t, x, \frac{x}{\delta^2}, k\right) + \dots \quad (8.66)$$

We assume formally that the leading order term  $\bar{\phi}$  is deterministic and independent of the fast variable  $z = x/\delta^2$ . We insert this expansion into (8.54) and obtain in the order  $O(\delta^{-1})$ :

$$\nabla V(z) \cdot \nabla_k \bar{\phi} - k \cdot \nabla_z \phi_1 = 0. \quad (8.67)$$

Let  $\theta \ll 1$  be a small positive regularization parameter that will be later sent to zero, and consider a regularized version of (8.67):

$$\frac{1}{|k|} \nabla V(z) \cdot \nabla_k \bar{\phi} - \hat{k} \cdot \nabla_z \phi_1 + \theta \phi_1 = 0,$$

Its solution is

$$\phi_1(z, k) = -\frac{1}{|k|} \int_0^\infty \sum_{m=1}^d \frac{\partial V(z + s\hat{k})}{\partial z_m} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_m} e^{-\theta s} ds, \quad (8.68)$$

and the role of  $\theta > 0$  is to ensure that the integral in the right side converges. The next order equation becomes upon averaging

$$\frac{\partial \bar{\phi}}{\partial t} + k \cdot \nabla_x \bar{\phi} = \langle \nabla V(z) \cdot \nabla_k \phi_1 \rangle. \quad (8.69)$$

The term in the right side above may be written using expression (8.68) for  $\phi_1$ :

$$\langle \nabla V(z) \cdot \nabla_k \phi_1 \rangle = \left\langle \sum_{m,n=1}^d \frac{\partial V(z)}{\partial z_m} \frac{\partial}{\partial k_m} \left( \frac{1}{|k|} \int_0^\infty \frac{\partial V(z + s\hat{k})}{\partial z_n} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right) \right\rangle.$$

Using spatial stationarity of  $H_1(z, k)$  we may rewrite the above as

$$\begin{aligned} & - \left\langle \sum_{m,n=1}^d V(z) \frac{\partial}{\partial z_m} \frac{\partial}{\partial k_m} \left( \frac{1}{|k|} \int_0^\infty \frac{\partial V(z + s\hat{k})}{\partial z_n} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right) \right\rangle \\ &= - \sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left( \frac{1}{|k|} \int_0^\infty \left\langle V(z, k) \frac{\partial^2 V(z + s\hat{k})}{\partial z_n \partial z_m} \right\rangle \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right) \\ &= - \sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left( \frac{1}{|k|} \int_0^\infty \frac{\partial^2 R(s\hat{k})}{\partial x_n \partial x_m} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right) \\ &\rightarrow -\frac{1}{2} \sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left( \frac{1}{|k|} \int_{-\infty}^\infty \frac{\partial^2 R(s\hat{k})}{\partial x_n \partial x_m} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} ds \right), \text{ as } \theta \rightarrow 0^+. \end{aligned}$$

We insert the above expression into (8.69) and obtain

$$\frac{\partial \bar{\phi}}{\partial t} = \sum_{m,n=1}^d \frac{\partial}{\partial k_n} \left( D_{nm}(k) \frac{\partial \bar{\phi}}{\partial k_m} \right) + k \cdot \nabla_x \bar{\phi} \quad (8.70)$$

with the diffusion matrix  $D(\hat{k}, k)$  as in (8.55). Observe that (8.70) is nothing but (8.57). However, the naive asymptotic expansion (8.66) may not be justified directly, to the best of my knowledge. The rigorous proof is based on a completely different method.

## References

- [1] W.O. Amrein, and V. Georgescu, On the characterization of bound states and scattering states in quantum mechanics. *Helv. Phys. Acta* **46**, 1973/74, 635–658.
- [2] D.V. Anosov, *The additive functional homology equation that is connected with an ergodic rotation of the circles (Russian)*, *Izv. Akad. Nauk SSSR Ser. Mat.* **37**, 1973, 1259–1274.



- [3] G. Bal, T. Komorowski and L. Ryzhik, Self-averaging of the Wigner transform in random media, *Comm. Math. Phys.*, 2003, **242**, 81–135.
- [4] H. Berestycki, F. Hamel and N. Nadirashvili, Elliptic eigenvalue problems with large drift and applications to nonlinear propagation phenomena. *Comm. Math. Phys.* **253**, 2005, 451–480.
- [5] H. Berestycki, A. Kiselev, A. Novikov and L. Ryzhik, The explosion problem in a flow, *Jour. d’Anal. Mathématique*, **110**, 2010, 31–65.
- [6] P. Billingsley, *Convergence of Probability Measures*, Wiley, 1999.
- [7] P. Constantin, A. Kiselev and L. Ryzhik, Quenching of flames by fluid advection, *Comm. Pure Appl. Math.*, **54**, 2001, 1320–1342.
- [8] P. Constantin, A. Kiselev, L. Ryzhik and A. Zlatoš, Diffusion and mixing in fluid flow. *Ann. of Math.*, **168**, 2008, 643–674.
- [9] I.P. Cornfeld, S.V. Fomin and Ya.G. Sinai, *Ergodic Theory*, Springer-Verlag, New York, 1982.
- [10] H. Cycon, R. Froese, W. Kirsch and B. Simon, *Schrödinger Operators*, Springer-Verlag, 1987.
- [11] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, Vol. 3: Spectral Theory and Applications, Springer, 2000.
- [12] D. Dürr, S. Goldstein and J. Lebowitz, Asymptotic motion of a classical particle in a random potential in two dimensions: Landau model, *Comm. Math. Phys.*, **113**, 1987, 209–230.
- [13] V. Enss, Asymptotic completeness for quantum mechanical potential scattering. I. Short range potentials. *Comm. Math. Phys.* **61**, 1978, 285–291.
- [14] L.C. Evans, *Partial Differential Equations*, Second Edition, AMS, 2010.
- [15] A. Fannjiang and L. Wołowski, Noise induced dissipation in Lebesgue-measure preserving maps on d-dimensional torus. *J. Statist. Phys.* **113**, 2003, 335–378.
- [16] A. Fannjiang, S. Nonnenmacher, L. Wołowski, Relaxation time of quantized toral maps. *Ann. Henri Poincaré* **7**, 2006, 161–198.
- [17] A. Fannjiang, S. Nonnenmacher, L. Wołowski, Dissipation time and decay of correlations. *Nonlinearity* **17**, 2004, 1481–1508.
- [18] B. Fayad, *Weak mixing for reparameterized linear flows on the torus*, *Ergodic Theory Dynam. Systems* **22**, 2002, 187–201.
- [19] B. Fayad, Smooth mixing flows with purely singular spectra. *Duke Math. J.* **132**, 2006, 371–391.

- [20] B. Franke, Integral inequalities for the fundamental solutions of diffusions on manifolds with divergence-free drift. (English summary) *Math. Z.* **246**, 2004, n 373–403.
- [21] J. Garnier and G. Papanicolaou, Passive sensor imaging using cross correlations of noisy signals in a scattering medium, *SIAM J. Imaging Sciences*, **2**, 2009, 396–437.
- [22] J. Garnier, G. Papanicolaou, A. Semin, and C. Tsogka, Signal-to-noise ratio estimation in passive correlation-based imaging, *SIAM J. Imaging Sciences*, **6**, 2013, 1092–1110.
- [23] P. Gérard, Microlocal defect measures, *Comm. PDEs*, **16**, 1991, 1761-1794.
- [24] P.Gérard, P.Markovich, N.Mauser and F.Poupaud, Homogenization limits and Wigner transforms, *Comm.Pure Appl. Math.*, **50**, 1997, 323-380.
- [25] Ichihara, K. and Kunita, H., *A classification of the second order degenerate elliptic operators and its probabilistic characterization*, *Z. Wahrscheinlichkeitstheorie verw.*, **30**, 1974, 235-254.
- [26] Hörmander, L., Hypoelliptic second order differential equations, *Acta Math.*, **119**, 1967, 147-171.
- [27] A. Katok, *Cocycles, cohomology and combinatorial constructions in ergodic theory*, In collaboration with E. A. Robinson, Jr. *Proc. Sympos. Pure Math.*, 69, Smooth ergodic theory and its applications (Seattle, WA, 1999), 107–173, Amer. Math. Soc., Providence, RI, 2001.
- [28] Y. Katznelson, *An Introduction to Harmonic Analysis*, Cambridge University Press, 2004.
- [29] J.B. Keller, Stochastic equations and wave propagation in random media, *Proc. Sympos. Appl. Math. AMS*, Vol. XVI, 1964, 145–170.
- [30] H. Kesten and G. Papanicolaou, A limit theorem for turbulent diffusion, *Comm. Math. Phys.*, **65**, 1979, 97-128.
- [31] H. Kesten, G. C. Papanicolaou, A Limit Theorem for Stochastic Acceleration, *Comm. Math. Phys.*, **78**, 1980, 19–63.
- [32] R. Khasminskii, A limit theorem for solutions of differential equations with a random right hand side, *Theory Probab. Appl.* **11**, 390-406.
- [33] Y. Kifer, *Random Perturbations of Dynamical Systems*, Birkhäuser, Boston, 1988.
- [34] A. Kiselev and A. Zlatoš, Quenching of combustion by shear flows, *Duke Math. J.* **132**, 2006, 49–72.
- [35] A.N. Kolmogorov, *On dynamical systems with an integral invariant on the torus (Russian)*, *Doklady Akad. Nauk SSSR (N.S.)*, **93**, 1953, 763–766.
- [36] T. Komorowski, Diffusion approximation for the advection of particles in a strongly turbulent random environment, *Ann. Probab.*, **24**, 1996, 346–376.

- [37] T. Komorowski and L. Ryzhik, Diffusion in a weakly random Hamiltonian flow, *Comm. Math. Phys.*, **262**, 2006, 277–323.
- [38] T. Komorowski and L. Ryzhik, The stochastic acceleration problem in two dimensions, *Israel Jour. Math.*, **155**, 2006, 157–204.
- [39] J. Moser, *On the volume elements on a manifold*, *Trans. Amer. Math. Soc.*, **120**, 1965, 286–294.
- [40] J. von Neumann, *Zur operatorenmethode in der classischen mechanik*, *Annals of Math (2)*, **33**, 1932, 587–642.
- [41] J.R. Norris, Long-time behaviour of heat flow: global estimates and exact asymptotics. *Arch. Rational Mech. Anal.* **140**, 1997, 161–195.
- [42] G. Papanicolaou and W. Kohler, Asymptotic theory of mixing stochastic ordinary differential equations, *Comm. Pure Appl. Math.*, **27**, 1974, 641–668.
- [43] D. Ruelle, A remark on bound states in potential-scattering theory. *Nuovo Cimento A (10)* **61**, 1969, 655–662.
- [44] M.D. Šklover, *Classical dynamical systems on the torus with continuous spectrum*, *Izv. Vuzov*, **10**, 1967, 113–124.
- [45] S. Sternberg, *On differential equations on the torus*, *Amer. J. Math.*, **79**, 1957, 397–402.
- [46] B.S. White, The stochastic caustic, *SIAM J. Appl. Math.* **44**, 1984, 127–149.
- [47] E. Wigner, On the quantum correction for thermodynamic equilibrium, *Physical Rev.*, **40**, 1932, 749–759.