Weak randomness in evolution problems: an introduction. Lecture notes for the Colorado Spring 2018 school

Lenya Ryzhik*

May 22, 2018

1 Introduction

The basic models

We will discuss in these lecture notes some evolution problems that involve weak random perturbations of very basic deterministic dynamics. For particles, such problems may take the form of an ODE

$$\frac{dX(t)}{dt} = \bar{u}(t, X(t)) + \varepsilon v(t, X(t), \quad X(0) = x, \tag{1.1}$$

with a deterministic background velocity field $\bar{u}(t, x)$ and a random fluctuation $v(t, x)$, or a stochastic differential equation

$$dX(t) = \bar{u}(t, X(t))dt + \varepsilon dB_t, \quad X(0) = x. \tag{1.2}$$

Here, B_t is a standard d-dimensional Brownian motion. In both cases, $\varepsilon \ll 1$ is a small parameter measuring the strength of the fluctuations relative to the background. Such problems are known as passive scalar problems as they are, respectively, related to linear PDEs

$$\frac{\partial \phi}{\partial t} + (\bar{u}(t, x) + \varepsilon v(t, x)) \cdot \nabla \phi = 0, \tag{1.3}$$

and

$$\frac{\partial \phi}{\partial t} + \bar{u}(t, x) \cdot \nabla \phi = \frac{\varepsilon^2}{2} \Delta \phi. \tag{1.4}$$

The word "passive" is used since the flow field in (1.3) and (1.4) does not depend on the "scalar" $\phi(t, x)$ – unlike, say, in the two-dimensional Euler or Navier-Stokes equations that are known as "active" scalar problems.

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

A close relative of the "random velocity" model (1.1) is the stochastic acceleration problem that describes the evolution of a particle driven by a random force:

$$\frac{dX(t)}{dt} = K(t), \quad \frac{dK(t)}{dt} = \varepsilon F(t, X(t)), \quad X(0) = x, \quad K(0) = k. \tag{1.5}$$

Here, $X(t)$ is the particle position, $K(t)$ is its momentum, and $F(t, x)$ is a random force field. The mass of the particle is set to be equal to one.

As far as PDE are concerned, we will discuss just two examples: the random Schrödinger equation

$$i\frac{\partial \phi}{\partial t} + \frac{1}{2}\Delta\phi - \varepsilon V(t, x)\phi = 0, \tag{1.6}$$

and, if time permits, the random heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + \varepsilon V(t, x)u, \tag{1.7}$$

both with a weak random potential. Once again, both in (1.6) and (1.7) we assume that $V(t, x)$ is a random potential field. As a side note, we will not address in these notes the construction of random fields. Note that one can think of (1.6) and (1.7) as infinite-dimensional versions of the particle dynamics in (1.1), with the first term in the right side of (1.6) or (1.7) being the equivalent of an infinite-dimensional mean drift, and the second the weak fluctuation.

A reader may either accept that such objects exist, consult the basic probability textbooks, or just think of the discrete case when a random field is simply a collection of random variables at each lattice point.

A simple example of a diffusive limit

As the random fluctuations are weak in all of these problems, the evolution is expected to follow for a long time the deterministic background dynamics, which is

$$\frac{dX(t)}{dt} = \bar{u}(t, X(t)), \quad X(0) = x, \tag{1.8}$$

for both (1.1) and (1.2), the homogeneous Schrödinger equation

$$i\frac{\partial \phi}{\partial t} + \frac{1}{2}\Delta\phi = 0, \tag{1.9}$$

for (1.6), and the standard heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u \tag{1.10}$$

for (1.7). However, after a sufficiently long time the effect of the fluctuations will build up, and we will no longer be able to neglect them, even in the leading order. To understand when this should happen, let us consider maybe the simplest such situation: we take the background flow in (1.1) to be $\bar{u}(t, x) = 0$ and the random fluctuation is uniform in space. In other words, we look at

$$\frac{dX(t)}{dt} = \varepsilon V(t), \quad X(0) = 0, \tag{1.11}$$

so that

$$X(t) = \varepsilon \int_0^t V(s)ds. \tag{1.12}$$

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 $t_1, t_2, \ldots, t_N$, and a shift h, the joint law of the random variables $V(t_1+h), V(t_2+h), \ldots, V(t_N+h)$ does not depend on h. This means, in particular, that the expected value $\bar{V} = \mathbb{E}[V(t)]$ is independent of 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 (1.12), we see that its average position is

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

where $\bar{V} = \mathbb{E}[V(t)]$ is the mean velocity, which does not depend on t because of the stationarity assumption. Therefore, if $\bar{V} \neq 0$, then the particle moves by a distance $O(1)$ after times of the order $t \sim \varepsilon^{-1}$, which is by no means a surprising result. However, if $\bar{V} \neq 0$, we would simply introduce $Y(t) = X(t) - \bar{V}t$, and $Y(t)$ would satisfy an ODE with a mean-zero drift. Thus, one would still need to face the mean-zero flow situation.

If $\bar{V} = 0$, then the average position is $\bar{X}(t) = 0$ for all $t > 0$, which tells us nothing about the typical size of $X(t)$. The way to find out if the particle performs a non-trivial motion is to look at its variance:

$$\begin{aligned}
&\langle X_i(t)X_j(t)\rangle = \varepsilon^2 \int_0^t ds_1 \int_0^t ds_2 \mathbb{E}(V_i(s_1)V_j(s_2)) = \varepsilon^2 \int_0^t ds_1 \int_0^t ds_2 R_{ij}(s_1 - s_2) \\
&= \varepsilon^2 \int_0^t ds_1 \int_0^{s_1} ds_2 R_{ij}(s_1 - s_2) + \varepsilon^2 \int_0^t ds_1 \int_{s_1}^t ds_2 R_{ij}(s_1 - s_2) \\
&= \varepsilon^2 \int_0^t ds_1 \int_0^{s_1} ds_2 R_{ij}(s_2) + \varepsilon^2 \int_0^t ds_1 \int_0^{t-s_1} ds_2 R_{ij}(-s_2) \\
&= \varepsilon^2 \int_0^t (t - s_2)[R_{ij}(s_2) + R_{ij}(-s_2)]ds_2 = \varepsilon^2[D_{ij}t + O(1)], \quad \text{as } t \to +\infty,
\end{aligned} \tag{1.13}$$

with the diffusivity matrix

$$D_{ij} = \int_{-\infty}^{\infty} R_{ij}(s)ds = \hat{R}_{ij}(0). \tag{1.14}$$

Bochner's theorem says that the matrix $\hat{R}_{ij}(\omega)$ is nonnegative-definite for all $\omega \in \mathbb{R}$. This is because for any rapidly decaying vector-valued function $f(x)$ we have (here and throughout we adopt the convention that repeated indices are summed)

$$\begin{aligned} &\int \hat{R}_{ij}(\omega)\hat{f}_i(\omega)\hat{f}_j(\omega)d\omega = \int R_{ij}(x)(f_i \star f_j)(x)dx = \int R_{ij}(x)f_i(x-y)f_j(y)dxdy \\ &= \int R_{ij}(x-y)f_i(x)f_j(y)dxdy = \sum_{i,j}\int \mathbb{E}[V_i(x)f_i(x)V_j(y)f_j(y)]dxdy \\ &= \mathbb{E}\Big(\sum_i \int V_i(x)f_i(x)dx\Big)^2 \geq 0. \end{aligned} \tag{1.15}$$

This fact is often used to show that the effective parameters in a macroscopic approximation of the microscopic dynamics are positive. In particular, in the present case, the matrix D_{ij} given by (1.14) is non-negative definite, as any respectable diffusion matrix better be.

Expression (1.13) 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 1.1 *Let $V(t)$ be a stationary in time $\mathbb{R}^d$-valued random process with mean zero and correlation function*

$$\mathbb{E}(V_i(s)V_j(t)) = R_{ij}(t-s).$$

Assume that the functions $R_{ij}(t)$ are of the Schwartz class. Then the process

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

converges in law to $Y(t) = B_D(t)$. Here, $B_D(t)$ is the d-dimensional Brownian motion with the diffusion matrix

$$D_{ij} = \int_{-\infty}^{\infty} R_{ij}(s)ds. \tag{1.16}$$

That is, we have the following result: if $X(t)$ solves (1.11) with a mean-zero statistically time homogeneous random field $V(t)$ then the process $X_\varepsilon(t) = X(t/\varepsilon^2)$ converges, as $t \to +\infty$, to a Brownian motion with the covariance matrix D_{ij}. The main observation here is that "mean-zero randomness of size ε has a non-trivial effect on the time scales of the order ε^{-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 $\omega \in \mathbb{R}$.

In lieu of a full proof of Theorem 1.1 – we will prove a more general version later on, we simply explain how to see this when $V(t)$ is a Gaussian field, meaning that for any collection

of times $t_1, t_2, \dots t_N$, the random variables $V(t_1), V(t_2), \dots, V(t_N)$ are jointly Gaussian. Let us look at the advection equation

$$\frac{\partial \phi}{\partial t} + \varepsilon V(t) \cdot \nabla \phi = 0, \quad \phi(0,x) = f(x), \tag{1.17}$$

and rescale the space and time variables: set

$$\phi^\varepsilon(t,x) = \phi\Big(\frac{t}{\varepsilon^2}, x\Big), \tag{1.18}$$

so that (1.17) becomes

$$\frac{\partial \phi^\varepsilon}{\partial t} + \frac{1}{\varepsilon} V\Big(\frac{t}{\varepsilon^2}\Big) \cdot \nabla \phi^\varepsilon = 0, \quad \phi(0,x) = f(x), \tag{1.19}$$

The solution can be written explicitly in terms of the Fourier transform:

$$\phi^\varepsilon(t,x) = \int_{\mathbb{R}^d} \exp\Big\{ -i\varepsilon \int_0^{t/\varepsilon^2} \xi \cdot V(s) ds \Big\} \hat{f}(\xi) \frac{d\xi}{(2\pi)^d}. \tag{1.20}$$

We use here the normalization

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-\xi \cdot x} f(x) dx, \quad f(x) = \int_{\mathbb{R}^d} e^{i\xi \cdot x} \hat{f}(\xi) \frac{d\xi}{(2\pi)^d}. \tag{1.21}$$

As $V(t)$ is a Gaussian vector-valued process, the random variable

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

is also Gaussian with the variance

$$\begin{aligned} \mathbb{E}(I_\varepsilon(t))^2 &= \varepsilon^2 \sum_{i,j=1}^d \xi_i \xi_j \int_0^{t/\varepsilon^2} \int_0^{t/\varepsilon^2} \mathbb{E}(V_i(s_1) V_j(s_2)) ds_1 ds_2 \\ &= \varepsilon^2 \sum_{i,j=1}^d \xi_i \xi_j \int_0^{t/\varepsilon^2} \int_0^{t/\varepsilon^2} R_{ij}(s_1 - s_2) ds_1 ds_2 = \frac{1}{2} D_{ij} \xi_i \xi_j t + O(\varepsilon^2), \quad \text{as } \varepsilon \to 0, \end{aligned} \tag{1.22}$$

as follows from the computation in (1.13). Using this in (1.20) gives

$$\mathbb{E}[\phi^\varepsilon(t,x)] \to \bar{\phi}(t,x) = \int_{\mathbb{R}^d} \exp\Big[-\frac{1}{2} D_{ij} \xi_i \xi_j t \Big] \hat{f}(\xi) \frac{d\xi}{(2\pi)^d}, \quad \text{as } \varepsilon \to 0. \tag{1.23}$$

Note that $\bar{\phi}(t,x)$ is the solution of the diffusion equation

$$\frac{\partial \bar{\phi}}{\partial t} = \frac{1}{2} \nabla \cdot (D \nabla \bar{\phi}), \quad \bar{\phi}(0,x) = f(x). \tag{1.24}$$

The convergence of $\phi^\varepsilon(t,x)$ to $\bar{\phi}(t,x)$ for any sufficiently regular function $f(x)$ means exactly that the process $X^\varepsilon(t)$ converges weakly to the Brownian motion with the diffusivity matrix D_{ij}.

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.

Another toy example: advection equation with a random potential

Let us now consider another extremely simple example where everything can be computed quite explicitly. One of the "real" examples is the Schrödinger equation (1.6):

$$i\phi_t + \Delta\phi - \varepsilon V(x)\phi = 0, \quad \phi(0,x) = \phi_0(x), \tag{1.25}$$

with a small time-independent random potential $\varepsilon V(x)$. The fact that the potential is time-independent makes this problem rather difficult. Indeed, if we consider the problem without any background dynamics:

$$i\phi_t - \varepsilon V(x)\phi = 0, \quad \phi(0,x) = \phi_0(x), \tag{1.26}$$

then we simply have

$$\phi(t,x) = \phi_0(x)\exp[-i\varepsilon V(x)t], \tag{1.27}$$

and there does not seem to be any interesting long time limit, as opposed to the case of a time-dependent potential $V(t)$, where we would get a limit of a solution to (1.26) using Theorem 1.1. Thus, any interesting behavior for the solutions of (1.6) may only result from the interaction between the Laplacian and the random potential in the Schrödinger equation.

To illustrate this point, as a toy model, let us consider a linear advection equation with a similar perturbation:

$$\frac{\partial\phi}{\partial t} + \bar{u}\cdot\nabla\phi + i\varepsilon V(x)\phi = 0, \quad \phi(0,x) = \phi_0(x). \tag{1.28}$$

Here, $\bar{u} \neq 0$ is a constant drift, and $V(x)$ is a spatially homogeneous mean-zero random field:

$$\mathbb{E}V(x) \equiv 0. \tag{1.29}$$

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

$$\mathbb{E}(V(x)V(y)) = R(x-y). \tag{1.30}$$

Recall that this follows from spatial homogeneity of $V(x)$. We assume that the function $R(x)$ is of the Schwartz class for now. Equation (1.28) is easy to solve explicitly:

$$\phi(t,x) = \phi_0(x-\bar{u}t)\exp\Big(-i\varepsilon\int_0^t V(x-s\bar{u})ds\Big). \tag{1.31}$$

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

$$\phi_\varepsilon(t,x) = \phi(\frac{t}{\varepsilon^2}, x), \tag{1.32}$$

so that

$$\phi_\varepsilon(t,x) = \phi_0(x - \frac{t}{\varepsilon^2}\bar{u}) \exp\Big(- i\varepsilon \int_0^{t/\varepsilon^2} V(x - s\bar{u})ds\Big). \tag{1.33}$$

The first term is simply the solution of the "fast" homogeneous problem (1.28) with $V = 0$, and need not have any limit unless $\phi_0(x)$ decays at infinity, while the second factor

$$\zeta_\varepsilon(t,x) = \exp\Big(- i\varepsilon \int_0^{t/\varepsilon^2} V(x - s\bar{u})ds\Big) \tag{1.34}$$

comes from the "slow" random dynamics. Its limit can be computed from Theorem 1.1, which implies that $\zeta_\varepsilon(t,x)$ converges in law to

$$\bar{\zeta}(t,x) = \exp(-i\sqrt{D}B(t)). \tag{1.35}$$

The diffusion coefficient D is

$$D = \int_{-\infty}^{\infty} R(\bar{u}s)ds. \tag{1.36}$$

Let us formulate this result as a theorem.

Theorem 1.2 *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 (1.28), $\phi_\varepsilon(t,x) = \phi(t/\varepsilon^2, x)$, and $\bar{\phi}(t,x)$ be the solution of (1.28) with $V = 0$. Then $\phi_\varepsilon(t,x)$ can be decomposed as*

$$\phi_\varepsilon(t,x) = \bar{\phi}(\frac{t}{\varepsilon^2}, x)\zeta_\varepsilon(t,x). \tag{1.37}$$

The function $\zeta_\varepsilon(t,x)$ converges in law, as $\varepsilon \to 0$ to $\bar{\zeta}(t,x)$ given by (1.35).

This example is very simple but it has some of 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 1.2 fails. Thus, the long time dynamics of the solutions of (1.28) involves a non-trivial interaction of the background dynamics and the random fluctuations. Such interactions can become highly non-trivial as soon as we move away from the most elementary examples.

An example of relaxation enhancement: the Dirichlet eigenvalues

In the previous examples, the background dynamics was very simple. We will stick with this situation in most of the rest of the notes, but to illustrate some of the other possibilities, let us now consider the opposite situation: the background dynamics is "complex" and ask if the random fluctuations can have a "stronger than expected" effect. We will consider an advection-diffusion problem with a time-independent background flow:

$$\frac{\partial \phi}{\partial t} + u(x) \cdot \nabla \phi = \frac{\varepsilon^2}{2}\Delta\phi, \quad x \in \Omega, \tag{1.38}$$

with the Dirichlet boundary condition:

$$\phi(t,x) = 0 \text{ for } x \in \partial\Omega. \tag{1.39}$$

Recall that the underlying stochastic differential equation (1.2) is

$$dX^\varepsilon(t) = u(X^\varepsilon(t))dt + \varepsilon dB_t, \quad X^\varepsilon(0) = x, \tag{1.40}$$

hence randomness is of the strength ε and, according to the previous philosophy, we should rescale the time $t \to t/\varepsilon^2$, setting $\phi^\varepsilon(t,x) = \phi(t/\varepsilon^2, x)$, which gives an initial boundary value problem

$$\begin{aligned} &\frac{\partial \phi^\varepsilon}{\partial t} + \frac{1}{\varepsilon^2} u(x) \cdot \nabla \phi^\varepsilon = \frac{1}{2}\Delta \phi^\varepsilon, \quad x \in \Omega, \\ &\phi^\varepsilon(t,x) = 0 \text{ for } x \in \partial\Omega, \\ &\phi^\varepsilon(0,x) = \phi_0(x), \quad x \in \Omega. \end{aligned} \tag{1.41}$$

Let us assume that $u(x)$ is "complex" in some intuitive sense. We will discuss this "complexity" from the probabilistic and PDE points of view. First, from the probabilistic perspective, recall that the solution of (1.41) has the following probabilistic characterization:

$$\phi(t,x) = \tilde{\phi}_0(X^\varepsilon(t \wedge \tau; x)). \tag{1.42}$$

Here, $X^\varepsilon(t;x)$ is the trajectory generated by the SDE (1.40), τ is the first exit time for $X^\varepsilon(t)$ from Ω:

$$\tau = \sup[t : \ X(t;x) \in \Omega \text{ for all } 0 \le s < t], \tag{1.43}$$

and $t \wedge \tau = \min(t, \tau)$. We have also set $\tilde{\phi}_0(x) = \phi_0(x)$ for $x \in \Omega$ and $\tilde{\phi}_0(x) = 0$ for $x \in \partial\Omega$, so that $\tilde{\phi}_0(X_\tau) = 0$. With this probabilistic interpretation of the solution in mind, we can take "complexity" to mean that the background trajectory

$$\frac{dX(t)}{dt} = u(X(t)), \quad X(0) = x, \tag{1.44}$$

comes close to the boundary $\partial\Omega$ for any starting point $x \in \Omega$. Assuming that the trajectories $X^\varepsilon(t)$ and $X(t)$ stay close until such time, and if the distance between $X^\varepsilon(t)$ and $\partial\Omega$ is of the order ε, then the particle will exit the domain at this moment with a positive probability, due to the diffusion. This would make $\phi^\varepsilon(t,x)$ small for t large as we will have $t < \tau$ with a high probability. In order for this philosophy to be plausible, the flow $u(x)$ better have no sinks where particles can get stuck despite the small diffusion, which is ensured by the incompressibility condition:

$$\nabla \cdot u(x) = 0, \tag{1.45}$$

as then the flow map for (1.44) is Lebesgue measure preserving and sinks can not exist. This heuristics indicates from the probabilistic point of view that if $u(x)$ is "complex" then the solutions $\phi^\varepsilon(t,x)$ to (1.41) should be "very small" if ε is sufficiently small, that is, if the advection term in (1.41) is sufficiently strong.

From the PDE point of view, the argument for the decay of the solutions when $u(x)$ is "complex" is different but also very natural. Let us multiply (1.41) by ϕ^ε and integrate over Ω:

$$\frac{1}{2}\frac{d}{dt}\int_\Omega |\phi^\varepsilon(t,x)|^2 dx + \frac{1}{\varepsilon^2}\int_\Omega \phi^\varepsilon (u \cdot \nabla \phi^\varepsilon) dx = -\frac{1}{2}\int_\Omega |\nabla \phi^\varepsilon(t,x)|^2 dx. \tag{1.46}$$

Assuming that u is divergence free: it satisfies (1.45) and is also tangential at the boundary, so that

$$u \cdot \nu = 0 \text{ on } \partial\Omega, \tag{1.47}$$

the second integral in the left side of (1.46) vanishes:

$$\frac{1}{\varepsilon^2} \int_\Omega \phi^\varepsilon (u \cdot \nabla \phi^\varepsilon) dx = 0, \tag{1.48}$$

and (1.46) becomes

$$\frac{1}{2} \frac{d}{dt} \int_\Omega |\phi^\varepsilon(t,x)|^2 dx = -\frac{1}{2} \int_\Omega |\nabla \phi^\varepsilon(t,x)|^2 dx. \tag{1.49}$$

Thus, the dissipation rate responsible for forcing the solution be small is

$$D = \frac{1}{2} \int_\Omega |\nabla \phi^\varepsilon(t,x)|^2 dx. \tag{1.50}$$

However, a "complex" flow would exactly bring nearby the areas where initially $\phi^\varepsilon(0,x)$ takes different values, creating large gradients in $\phi^\varepsilon(t,x)$. This will make the dissipation rate be large, which, in turn, will lead to the solution being small. Thus, both from the probabilistic and the PDE points of view, one may expect the solution to (1.41) be small. Next, we will see how this phenomenon, known as "relaxation enhancement", can be quantified.

The eigenvalues of the Laplacian

Before we explain how the relaxation enhancement can be formulated and proved, let us first recall some very basic facts about the principal Dirichlet eigenvalues for the Laplacian on a bounded domain [20]. For any smooth bounded domain Ω, there exists an eigenvalue λ_1 (called the principal eigenvalue) that corresponds to a positive eigenfunction $\phi_1 > 0$ in Ω:

$$\begin{aligned} &-\Delta \phi_1 = \lambda_1 \phi_1, \quad x \in \Omega, \\ &\phi_1 = 0 \text{ on } \partial\Omega. \end{aligned} \tag{1.51}$$

Moreover, λ_1 is the smallest of all eigenvalues of the Dirichlet Laplacian on Ω, λ_1 is a simple eigenvalue and all other eigenfunctions of the Laplacian change sign in Ω. For example, if Ω 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 Dirichlet eigenvalue of the Laplacian is given by the variational formula:

$$\lambda_1 = \inf_{\substack{\psi \in H_0^1(\Omega) \\ \|\psi\|_2 = 1}} \int_\Omega |\nabla \psi|^2 dx. \tag{1.52}$$

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

$$\begin{aligned} &\psi_t = \Delta\psi, \quad t > 0,\ x \in \Omega, \\ &\psi(t,x) = 0 \text{ on } \partial\Omega, \\ &\psi(0,x) = g(x). \end{aligned} \tag{1.53}$$

As $\phi_1(x) > 0$ in Ω, 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

$$\psi(t,x) \leq Ce^{-\lambda_1(t-1)}\phi_1(x), \tag{1.54}$$

for $t > 1$, and, similarly,

$$-\psi(t,x) \leq Ce^{-\lambda_1(t-1)}\phi_1(x), \tag{1.55}$$

so that

$$|\psi(t,x)| \leq Ce^{-\lambda_1(t-1)}\phi_1(x), \quad t \geq 1. \tag{1.56}$$

Therefore, all solutions of the Cauchy problem decay at the exponential rate determined by λ_1 as $t \to +\infty$.

The Dirichlet eigenvalues with a drift

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

$$\begin{aligned} &-\Delta\phi + \frac{1}{\varepsilon^2} u \cdot \nabla\phi = \lambda_1(\varepsilon)\phi, \quad \phi(x) > 0 \text{ in } \Omega, \\ &\phi = 0 \text{ on } \partial\Omega. \end{aligned} \tag{1.57}$$

We have dropped the factor $1/2$ in front of the Laplacian, switching from the "probabilistic Laplacian" (with $1/2$) to the "analyst Laplacian" (without $1/2$).

The operator in (1.57) is not self-adjoint (so that its eigenvalues are not necessarily real), and its eigenvalues do not obey an integral variational principle such as (1.52). Nevertheless, the Krein-Rutman theory for positive operators (see Chapter VIII of [16]) 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:

$$\begin{aligned} &\psi_t + \frac{1}{\varepsilon^2} u \cdot \nabla\psi = \Delta\psi, \quad t > 0, x \in \Omega, \\ &\psi(t,x) = 0 \text{ on } \partial\Omega, \\ &\psi(0,x) = g(x), \end{aligned} \tag{1.58}$$

that is,

$$\psi(t,x) \sim e^{-\lambda_1(\varepsilon)t}\phi_1(x), \quad \text{as } t \to +\infty. \tag{1.59}$$

Note that when $u = 0$ the exponential rate of decay for the solutions of (1.58) is simply the principal eigenvalue of the Laplacian. On the other hand, solutions of the Laplacian-less problem

$$\psi_t + \frac{1}{\varepsilon^2} u \cdot \nabla \psi = 0 \tag{1.60}$$

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.

Thus, the Laplacian by itself produces dissipation of order one, and the drift $u \cdot \nabla$ produces no dissipation whatsoever. An interesting phenomenon is that it is possible that solutions of the "combined" Cauchy problem (1.58) can still decay much faster than when $u = 0$ even though the drift can not lead to any decay by itself. To quantify this "enhanced relaxation", let us ask if it is possible that

$$\lambda_1(\varepsilon) \to +\infty \text{ as } \varepsilon \to 0. \tag{1.61}$$

Then (1.59) would show that solutions to (1.58) decay "very fast". Physically, this would mean that a sufficiently mixing flow, together with diffusion, would dramatically increase the cooling of the interior by the boundary. As we have explained above, one would expect this if the flow is sufficiently "complex" or "mixing".

A natural questions is what "mixing" or "complex" flow 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

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

Here, we are asking a PDE question – hence, the first issue is to define what "mixing" means for us. This is quantified by the following beautiful result due to Berestycki, Hamel and Nadirashvili [8]. We denote by $\mathcal{I}_0$ the set of all first integrals of u, solutions of

$$u \cdot \nabla \phi = 0 \text{ a.e. in } \Omega, \tag{1.62}$$

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

Theorem 1.3 *The principal eigenvalue $\lambda_1(\varepsilon)$ of (1.57) tends to $+\infty$ as $\varepsilon \to 0$ 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*

$$\lambda_1(\varepsilon) \to \bar{\lambda} := \min_{w \in \mathcal{I}_0} \frac{\int_\Omega |\nabla w|^2 dx}{\int_\Omega |w|^2 dx} \text{ as } \varepsilon \to 0, \tag{1.63}$$

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 (1.57) 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 (1.57), as irregular first integrals do not prevent strong decay of the solutions of the Cauchy problem. Roughly speaking, if u has an irregular first integral, then the solution

of the Cauchy problem may take its shape first because the drift is strong, but then the diffusion term will start dissipating it very quickly because of the relation

$$\frac{d}{dt}\int_\Omega |\psi(t,x)|^2 dx = -\int_\Omega |\nabla\psi(t,x)|^2 dx \tag{1.64}$$

that holds for solutions of (1.58) since u is incompressible: $\nabla \cdot u = 0$. The irregularity of the first integral would mean that the right side of (1.64) is "huge", leading to a very fast decrease of the L^2-norm of $\psi(t,x)$.

Second, the strong flow essentially forces the eigenfunction to be close to a first integral (if they exist), and then the variational principle (1.53) for the Laplacian operator is replaced by essentially the same expression (1.63) except that the set of allowed test functions is restricted to the first integrals.

Proof of Theorem 1.3

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

$$0 \le \lambda_1(\varepsilon) \le \int_\Omega |\nabla w(x)|^2 dx, \tag{1.65}$$

for any $\varepsilon \in \mathbb{R}$. In order to show that (1.65) holds, we take any first integral $w \in \mathcal{I}_0$, and multiply (1.57) by $w^2/(\phi+\varepsilon)$ with $\varepsilon > 0$ fixed:

$$-\int_\Omega \frac{w^2\Delta\phi}{\phi+\varepsilon}dx + \int_\Omega \frac{w^2}{\phi+\varepsilon}(u\cdot\nabla\phi)dx = \lambda_1(\varepsilon)\int_\Omega \frac{w^2}{\phi+\varepsilon}\phi\, dx. \tag{1.66}$$

Integrating by parts in the first term gives

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

The second term in the left side of (1.66) vanishes because $\nabla\cdot u = 0$ and w is a first integral:

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

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

$$\lambda_1(\varepsilon)\int_\Omega \frac{w^2}{\phi+\varepsilon}\phi dx \le \int_\Omega |\nabla w|^2 dx, \tag{1.67}$$

for any $w \in \mathcal{I}_0$. Passing to the limit $\varepsilon \to 0$ gives (1.65). 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 $\varepsilon_n \to 0$ such that $\lambda_1(\varepsilon_n)$ are bounded, then, again, as u is divergence-free, we have

$$\int_\Omega |\nabla \phi_n(x)|^2 dx = \lambda_1(\varepsilon_n) \int_\Omega |\phi_n(x)|^2 dx = \lambda_1(\varepsilon_n). \tag{1.68}$$

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 ϕ_{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 (1.57) by ε_{n_k} and passing to the limit $k \to +\infty$ gives

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

and

$$\|\bar{w}\|_{L^2(\Omega)} = 1. \tag{1.69}$$

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

$$\lim_{\varepsilon \to 0} \lambda_1(\varepsilon) = +\infty. \tag{1.70}$$

Finally, to show that (1.63) holds, let us assume, once again, that there exists a sequence $\varepsilon_n \to 0$ such that $\lambda_1(\varepsilon_n)$ are bounded. As the convergence of the subsequence ϕ_{n_k} to the first integral $\bar{w}$ is strong in $L^2(\Omega)$ and weak in $H_0^1(\Omega)$, it follows from (1.68), (1.69) and Fatou's lemma that

$$\liminf_{n \to +\infty} \lambda_1(\varepsilon_n) \geq \int_\Omega |\nabla \bar{w}(x)|^2 dx. \tag{1.71}$$

It remains to notice that (1.71) and (1.65) together imply the Rayleigh quotient formula (1.63), finishing the proof of Theorem 1.3.

2 Particles in two-dimensional randomly perturbed flows

The "eigenvalue relaxation enhancement" example shows that the interaction of a strong "complex' flow with a random perturbation may lead to extremely fast mixing. In this section, we consider particles governed by the familar stochastic differential equation

$$dX_t = -u(X)dt + \sqrt{2}\varepsilon dB_t, \tag{2.1}$$

but with a "simple" background flow $u(x)$. As before, we assume that $u(x)$ is an incompressible flow in a simply-connected domain Ω:

$$\nabla \cdot u = 0, \quad \text{for all } x \in \Omega, \tag{2.2}$$

and B_t is the standard Brownian motion. To simplify further, we will consider the two-dimensional case, where, as Ω is simply connected, all incompressible flows are of the form

$$u = \nabla^\perp H = (H_y, -H_x),$$

with some Hamiltonian $H(x,y)$. In particular, $H(x,y)$ is a first integral:

$$u \cdot \nabla H = 0,$$

and the flow u can not be relaxation enhancing. The underlying time-dependent PDE is

$$\frac{\partial \phi}{\partial t} + u \cdot \nabla \phi = \varepsilon^2 \Delta \phi, \tag{2.3}$$

and its time-rescaled version is

$$\frac{\partial \phi}{\partial t} + \frac{1}{\varepsilon^2} u \cdot \nabla \phi = \Delta \phi, \tag{2.4}$$

We will assume that the flow u does not penetrate the boundary of Ω:

$$u \cdot \nu = 0 \text{ on } \partial\Omega. \tag{2.5}$$

Here, ν is the normal to the boundary. We impose the Dirichlet boundary condition at $\partial\Omega$:

$$u = 0 \text{ on } \partial\Omega. \tag{2.6}$$

Let us make one additional observation. Recall that the probabilistic interpretation for the solutions of (2.4) is as follows. Consider the solution of the stochastic differential equation

$$dX_t = -\frac{1}{\varepsilon^2} u(X_t) dt + \sqrt{2} dB_t, \quad X_0 = x. \tag{2.7}$$

Then $\phi(t,x)$ is given by

$$\phi(t,x) = \mathbb{E}_x[\phi_0(X_{\min(t,\tau)}], \tag{2.8}$$

with the convention that $\phi_0(X_\tau) = 0$. Here, τ is the first exit time from the domain Ω for the process X_t starting at $X_0 = x$. Assume now that $\Phi(x)$ is a first integral of the flow $u(x)$:

$$u \cdot \nabla \Phi = 0, \tag{2.9}$$

and consider $Y_t = \Phi(X_t)$. The process Y_t satisfies a stochastic differential equation

$$dY_t = \nabla \Phi(X_t) \cdot dX_t + \Delta \Phi(X_t) dt = \Delta \Phi(X_t) dt + \sqrt{2} \nabla \Phi(X_t) \cdot dB_t. \tag{2.10}$$

In particular, there is no large term in (2.10) – 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, because X_t moves very fast on the level sets of the Hamiltonian but not across the level sets.

Oscillation on streamlines

Let us first show that the solution becomes nearly uniform on the streamlines of the flow. To keep the presentation simple, we will consider a steady version of this problem:

$$\begin{aligned} &-\Delta \phi^\varepsilon + \frac{1}{\varepsilon^2} u \cdot \nabla \phi^\varepsilon = g(x), \\ &\phi^\varepsilon(x) = 0, \quad x \in \partial\Omega. \end{aligned} \tag{2.11}$$

Here, Ω is a bounded smooth domain. Multiplying by $\phi^\varepsilon(x)$ and integrating by parts gives

$$\int_\Omega |\nabla \phi^\varepsilon|^2 dx = \int_\Omega f\phi^\varepsilon dx \le \|g\|_{L^2}\|\phi^\varepsilon\|_{L^2}. \tag{2.12}$$

The Poincaré inequality implies that

$$\|\phi^\varepsilon\|_{L^2} \le C_p \|\nabla \phi^\varepsilon\|_{L^2}. \tag{2.13}$$

It follows from (2.12) and (2.13) that

$$\|\phi^\varepsilon\|_{L^2} \le C\|\nabla \phi^\varepsilon\|_{L^2} \le C\|g\|_{L^2}. \tag{2.14}$$

Next, we multiply (2.11) by $u \cdot \nabla \phi^\varepsilon$ and integrate to get

$$\begin{aligned} \int_\Omega |u\cdot\nabla\phi^\varepsilon|^2 dx &= \varepsilon^2 \int_\Omega (u\cdot\nabla\phi^\varepsilon)\Delta\phi^\varepsilon dx + \varepsilon^2 \int_\Omega g(x)(u\cdot\nabla\phi) \\ &\le -\varepsilon^2 \int_\Omega \nabla(u\cdot\nabla\phi^\varepsilon)\cdot\nabla\phi^\varepsilon dx + \frac{\varepsilon^2}{2}\|g\|_{L^2}^2 + \frac{\varepsilon^2}{2}\int_\Omega |u\cdot\nabla\phi^\varepsilon|^2 dx. \end{aligned} \tag{2.15}$$

We rewrite the integrand in the second line above above as:

$$\nabla(u\cdot\nabla\phi^\varepsilon)\cdot\nabla\phi^\varepsilon = \frac{\partial u_k}{\partial x_i}\frac{\partial\phi^\varepsilon}{\partial x_k}\frac{\partial\phi^\varepsilon}{\partial x_i} + u_k\frac{\partial^2\phi^\varepsilon}{\partial x_k\partial x_i}\frac{\partial\phi^\varepsilon}{\partial x_i} = \frac{\partial u_k}{\partial x_i}\frac{\partial\phi^\varepsilon}{\partial x_k}\frac{\partial\phi^\varepsilon}{\partial x_i} + \frac{1}{2}u\cdot\nabla(|\nabla\phi^\varepsilon|^2). \tag{2.16}$$

Once again using incompressibility of u, we obtain from the above

$$\begin{aligned} -\int_\Omega \nabla(u\cdot\nabla\phi^\varepsilon)\cdot\nabla\phi^\varepsilon dx &= \frac{1}{2}\int_\Omega (u\cdot\nabla\left(|\nabla\phi^\varepsilon|^2\right))dx - \int_\Omega \frac{\partial u_n}{\partial x_m}\frac{\partial\phi^\varepsilon}{\partial x_m}\frac{\partial\phi^\varepsilon}{\partial x_n}dx \\ &\le M\varepsilon^2\int_\Omega |\nabla\phi^\varepsilon|^2 dx \le CM\varepsilon^2\|g\|_{L^2}, \end{aligned} \tag{2.17}$$

where $M = \|\nabla u\|_{L^\infty(\Omega)}$. We deduce that

$$\int_\Omega |u\cdot\nabla\phi^\varepsilon|^2 dx \le C\varepsilon^2\|g\|_{L^2}^2. \tag{2.18}$$

Informally, this estimate means that the oscillation of ϕ^ε along the stream lines of u is small. There are, of course, ways to make this more precise but it says, roughly, that as $\varepsilon \to 0$ 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

For a time-dependent problem, let us first consider a special situation when a radially symmetric 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 (2.1) becomes

$$\begin{aligned} dX_t &= Y_t dt + \varepsilon dB_t^{(1)}, \\ dY_t &= -X_t dt + \varepsilon dB_t^{(2)}. \end{aligned} \tag{2.19}$$

The time rescaling $t \to t/\varepsilon^2$ leads to

$$\begin{aligned} dX_t &= \frac{1}{\varepsilon^2} Y_t dt + dB_t^{(1)}, \quad X_0 = x, \\ dY_t &= -\frac{1}{\varepsilon^2} X_t dt + dB_t^{(2)}, \quad Y_0 = y. \end{aligned} \tag{2.20}$$

The corresponding PDE is

$$\begin{aligned} &\frac{\partial v}{\partial t} = \frac{y}{\varepsilon^2}\frac{\partial v}{\partial x} - \frac{x}{\varepsilon^2}\frac{\partial v}{\partial y} + \frac{1}{2}\varepsilon v, \\ &v(0,x) = v_0(x), \end{aligned} \tag{2.21}$$

in the sense that

$$v(t,x) = \mathbb{E}_{x,y}[v_0(X_t, Y_t)]. \tag{2.22}$$

Switching to the polar coordinates $x = r\cos\theta$, $y = r\sin\theta$ gives

$$\begin{aligned} &\frac{\partial v}{\partial t} = -\frac{1}{\varepsilon^2}\frac{\partial v}{\partial \theta} + \frac{1}{2}\Big[\frac{\partial^2 v}{\partial^2 r} + \frac{1}{r}\frac{\partial v}{\partial r} + \frac{1}{r^2}\frac{\partial^2 v}{\partial \theta^2}\Big], \\ &v(0,r,\theta) = v_0(r,\theta). \end{aligned} \tag{2.23}$$

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

$$\begin{aligned} &\frac{\partial \bar{v}}{\partial t} = \frac{1}{2}\Big[\frac{\partial^2 \bar{v}}{\partial^2 r} + \frac{1}{r}\frac{\partial \bar{v}}{\partial r}\Big], \\ &\bar{v}(0,r) = \bar{v}_0(r). \end{aligned} \tag{2.24}$$

The corresponding process $R_t = (X_t^2 + Y_t^2)^{1/2}$ converges to a diffusion with the generator as in (2.24). There is an additional feature here, specific to the quadratic Hamiltonian: we may write

$$v(t,r,\theta) = w(t,r,\theta - \frac{t}{\varepsilon^2}). \tag{2.25}$$

The function w satisfies

$$\begin{aligned} &\frac{\partial w}{\partial t} = \frac{1}{2}\Big[\frac{\partial^2 w}{\partial^2 r} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{1}{r^2}\frac{\partial^2 w}{\partial \theta^2}\Big], \\ &w(0,r,\theta) = v_0(r,\theta). \end{aligned} \tag{2.26}$$

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 $\varepsilon \to 0$.

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 (2.25) is

$$\bar{v}(t, r) = \frac{1}{2\pi} \int_0^{2\pi} w(t, r, \theta) d\theta, \tag{2.27}$$

and is independent of θ.

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 Ω means that $\partial\Omega$ is a level set of $H(x, y)$. Solutions of (2.11)

$$\begin{aligned} &-\Delta\phi^\varepsilon + \frac{1}{\varepsilon^2} u \cdot \nabla\phi^\varepsilon = g(x), \\ &\phi^\varepsilon(x) = 0, \quad x \in \partial\Omega. \end{aligned} \tag{2.28}$$

are uniformly bounded in $L^\infty(\Omega) \cap H_0^1(\Omega)$:

$$0 \le \phi^\varepsilon \le C, \quad \int |\nabla\phi^\varepsilon|^2 dx \le C, \tag{2.29}$$

with the constant $C > 0$ independent of $\varepsilon > 0$. The L^2-bound on the gradient in (2.29) is obtained by a simple multiplication of (2.28) by ϕ^ε and integration by parts. The L^∞-bound independent of ε in (2.29) is slightly trickier to prove and is a good exercise for the reader. Hence, the family ϕ_ε 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 (2.28) by ε^2 and let $\varepsilon \to +\infty$, we get

$$u \cdot \nabla\bar{\phi} = 0 \tag{2.30}$$

in the sense of distributions. In order to get an equation characterizing the limit $\bar{\phi}(h)$, it is convenient to introduce the curvilinear coordinates (h, θ). 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 θ so that $0 \le \theta \le 2\pi$. 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 (2.30) implies that $\bar{\phi}$ depends only on the variable h. The L^∞-bound in (2.29) implies that

$$0 \le \bar{\phi}(h) \le C.$$

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

$$\int_0^{2\pi} \frac{|\nabla H|^2}{J} d\theta = \oint_{H(x,y)=h} |\nabla H| dl := a(h), \tag{2.31}$$

and thus we have a weighted H^1-bound

$$\int_0^{H_0} a(h) |\bar{\phi}_h|^2 dh < +\infty,$$

which follows from (2.29), 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 (2.28) in the curvilinear coordinates:

$$\begin{aligned} &-\frac{|\nabla H|^2}{J} \frac{\partial^2 \phi^\varepsilon}{\partial h^2} - \frac{|\nabla\Theta|^2}{J} \frac{\partial^2 \phi^\varepsilon}{\partial \theta^2} - \frac{(\varepsilon H)}{J} \frac{\partial \phi^\varepsilon}{\partial h} - \frac{(\varepsilon\Theta)}{J} \frac{\partial \phi^\varepsilon}{\partial \theta} + \frac{1}{\varepsilon^2} \frac{\partial \phi^\varepsilon}{\partial \theta} = \frac{1}{J} g(h,\theta), \\ &\phi^\varepsilon(0,\theta) = 0, \ \phi^\varepsilon(h,\theta) \text{ is bounded for } 0 \le h \le H_0. \end{aligned} \tag{2.32}$$

Integrating this equation in θ and passing to the limit $\varepsilon \to 0$ we obtain the limit problem for the function $\bar{\phi}$:

$$\begin{aligned} &-a(h)\bar{\phi}''(h) - b(h)\bar{\phi}'(h) = \bar{g}(h), \\ &\bar{\phi}(0) = 0, \quad \bar{\phi}(h) \text{ is bounded for } 0 \le h \le H_0, \end{aligned} \tag{2.33}$$

with $a(h)$ as in (2.31), 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 (2.33) 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 $G_h = \{h \le H(x,y) \le H_0\}$ is the interior of the streamline $\{H(x,y) = h\}$. Hence, the Freidlin problem (2.33) can be re-written in a self-adjoint form as

$$\begin{aligned} &-\frac{d}{dh}\left(a(h) \frac{d\bar{\phi}(h)}{dh} \right) = \bar{g}(h), \\ &\bar{\phi}(0) = 0, \quad \bar{\phi}(h) \text{ is bounded for } 0 \le h \le H_0. \end{aligned} \tag{2.34}$$

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 (2.34) along each edge, with the coefficients $a(h)$ computed as in (2.31), 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 (2.28)

$$-\Delta\phi^\varepsilon+\frac{1}{\varepsilon^2}u\cdot\nabla\phi^\varepsilon=g(x), \tag{2.35}$$

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

$$\oint_{C_{out}}\frac{\partial\phi}{\partial n}dl=\oint_{C_{k1}}\frac{\partial\phi}{\partial n}dl+\oint_{C_{k2}}\frac{\partial\phi}{\partial n}dl+\text{l.o.t.} \tag{2.36}$$

Note that, since all contours in (2.36) are level sets, we have

$$\frac{\partial\phi}{\partial n}\approx\frac{\partial\bar{\phi}}{\partial h}|\nabla H|.$$

Using this in (2.36) leads to

$$\frac{\partial\bar{\phi}_{out}}{\partial h}\oint_{C_{out}}|\nabla H|dl=\frac{\partial\bar{\phi}_{k,1}}{\partial h}\oint_{C_{k1}}|\nabla H|dl+\frac{\partial\bar{\phi}_{k,2}}{\partial h}\oint_{C_{k2}}|\nabla H|dl. \tag{2.37}$$

In other words, the gluing condition at vertex k corresponding to the level set $H(x,y)=H_k$, is:

$$a_{k,out}\frac{\partial\bar{\phi}_{out}}{\partial h}(H_k)=a_{k,in1}\frac{\partial\bar{\phi}_{k,1}}{\partial h}(H_k)+a_{k,in1}\frac{\partial\bar{\phi}_{k,1}}{\partial h}(H_k). \tag{2.38}$$

These gluing conditions together with the Poisson equations (2.34) completely describe the limit problem.

3 Mixing in strong shear flows

In the previous section, we have considered a strong flow with closed streamlines, perturbed by a small diffusion. Let us now consider what happens when the streamlines are open. A simple example of such dynamics is advection by a shear flow in a channel

$$D=\{(x,y):\ x\in\mathbb{R},\ y\in\Omega\subset\mathbb{R}^d\}\subset\mathbb{R}^{d+1}. \tag{3.1}$$

Here, Ω is a smooth bounded domain – the channel cross-section. The flow trajectories are straight lines along the channel:

$$\frac{dX}{dt} = u(Y(t)), \quad \frac{dY}{dt} = 0, \quad X(0) = x, \ Y(0) = y. \tag{3.2}$$

This, of course, has an explicit solution

$$X(t) = x + u(y)t, \quad Y(t) = y. \tag{3.3}$$

We will now add a diffusive perturbation to the flow and consider a system of two stochastic differential equations

$$\begin{aligned} dX_t &= -u(Y_t) + \sqrt{2}\varepsilon dB_t^{(1)}, \\ dY_t &= \sqrt{2}\varepsilon dB_t^{(2)}. \end{aligned} \tag{3.4}$$

The Brownian motion $B_t^{(1)}$ is one-dimensional, while $B_t^{(2)}$ is d-dimensional. The corresponding Kolmogorov equation is

$$\frac{\partial \phi}{\partial t} + u(y)\frac{\partial \phi}{\partial x} = \varepsilon^2 \Delta \phi. \tag{3.5}$$

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

$$\frac{\partial \phi}{\partial t} + \frac{1}{\varepsilon^2}u(y)\frac{\partial \phi}{\partial x} = \Delta \phi. \tag{3.6}$$

This problem is posed in the channel D in (3.1), with the Neumann boundary condition at the boundary:

$$\frac{\partial \phi}{\partial \nu} = 0 \text{ at } \partial D = \mathbb{R} \times \partial \Omega. \tag{3.7}$$

Note that if $u(y) \equiv \bar{u} = \text{const}$, then solution of (3.7) is simply a translate of the solution of the heat equation:

$$\phi(t, x, y) = \bar{\phi}(t, x - \bar{u}\frac{t}{\varepsilon^2}, y), \tag{3.8}$$

where $\bar{\phi}(t, x, y)$ is the solution of the standard heat equation

$$\frac{\partial \bar{\phi}}{\partial t} = \Delta \bar{\phi}, \tag{3.9}$$

with the Neumann boundary conditions (3.7). Therefore, if $u(y)$ is a uniform flow, then the solution of (3.6) 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. Thus, we will ask the following question:

consider the solution of

$$\frac{\partial \phi}{\partial t} + \frac{1}{\varepsilon^2} u(y) \frac{\partial \phi}{\partial x} = \Delta \phi. \tag{3.10}$$

$$\frac{\partial \phi}{\partial \nu} = 0 \text{ at } \partial D = \mathbb{R} \times \partial \Omega, \tag{3.11}$$

$$\phi(0, x, y) = \phi_0(x, y),$$

with a rapidly decaying initial condition $\phi_0(x, y)$. When is it true that for any time $\tau > 0$ and $\varepsilon > 0$ we have

$$\|\phi(\tau, \cdot)\|_{L^\infty(D)} \leq \varepsilon, \tag{3.12}$$

provided that $\varepsilon < \varepsilon_0(\tau, \varepsilon)$?

Definition 3.1 *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 ε_0 such that the solution of (3.10) satisfies (3.12) 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 3.2 *We say that the profile $u(y) \in C^\infty(\Omega)$ satisfies the H-condition if*

$$\text{there is no point } y \in \Omega, \text{ where all derivatives of } u(y) \text{ vanish.} \tag{3.13}$$

The H-condition guarantees that the operator

$$\frac{\partial}{\partial t} + u(y) \frac{\partial}{\partial x} - \varepsilon_y \tag{3.14}$$

is hypoelliptic [32]. The study of existence of smooth fundamental solutions for such operators was initiated by Kolmogorov [41]. 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 [32]. The hypoellipticity of the operator (3.14) plays a key role in some of our considerations. The next result shows that the H-condition implies quenching.

Theorem 3.3 *Let $u \in C^\infty(\Omega)$ satisfy the H-condition. Then $u(y)$ is quenching. That is, for any $\varepsilon > 0$ and any $\tau > 0$ there exists a constant $C(u, \Omega, \tau, \varepsilon) > 0$ that is independent of $\varepsilon \in (0, 1)$ such that*

$$\|\phi(\tau, \cdot)\|_{L^\infty(D)} \leq \varepsilon \tag{3.15}$$

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 3.3 can be found in [12, 40].

We now prove Theorem 3.3. Let $\phi(t, x, y)$ be the solution of

$$\phi_t + \frac{1}{\varepsilon^2} u(y) \phi_x = \Delta \phi \tag{3.16}$$

$$\phi(0, x, y) = \phi_0(x, y)$$

$$\frac{\partial \phi}{\partial n} = 0 \text{ on } \partial D.$$

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

$$\begin{aligned} &\Psi_t + \frac{1}{\varepsilon^2}u(y)\Psi_x = \Delta_y \Psi \\ &\Psi(0,x,y) = \phi_0(x,y) \\ &\frac{\partial \Psi}{\partial n} = 0 \ \text{ on } \partial D. \end{aligned} \tag{3.17}$$

Here, $G(t,x)$ is the standard heat kernel

$$G(t,x) = \frac{1}{\sqrt{4\pi t}} \exp\left(-\frac{x^2}{4t}\right).$$

If $u(y)$ satisfies the H-condition (3.13) then the diffusion process defined by (3.17) has a unique smooth transition probability density. Indeed, the Lie algebra generated by the operators ∇_y and $\partial_t + \varepsilon^{-2}u(y)\partial_x$ consists of vector fields of the form

$$\nabla_y, \frac{\partial}{\partial t} + u(y)\frac{\partial}{\partial x}, \frac{\partial u(y)}{\partial y_k}\frac{\partial}{\partial x}, \frac{\partial^2 u(y)}{\partial y_i \partial y_j}\frac{\partial}{\partial x}, \ldots, u^{(n)}(y)\frac{\partial}{\partial x}, \ldots$$

which span $\mathbb{R}^2$ if $u(y)$ satisfies (3.13). Then the theory of Hörmander [32], and the results of Ichihara and Kunita [31] imply that there exists a smooth transition probability density $p_\varepsilon(t,x,y,y')$ such that

$$\Psi(t,x,y) = \int\limits_{\mathbb{R}} dx' \int\limits_{\Omega} dy' p_\varepsilon\left(t, x-x', y, y'\right) \phi_0(x',y').$$

In particular, the function $p_\varepsilon(t)$ is uniformly bounded from above for any $t > 0$ [31]. Then we have

$$\|\phi(t)\|_{L^\infty_D} \le \|p_\varepsilon(t)\|_{L^\infty(D)} \|\phi_0\|_{L^1(D)}.$$

It is straightforward to observe that

$$p_\varepsilon(t,x,y,y') = \varepsilon^2 p_0(t, \varepsilon^2 x, y, y')$$

with p_0 being the transition probability density for (3.17) 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

$$\phi(t,x,y) \le \varepsilon^2 \|p_0(t)\|_{L^\infty(D)} \|\phi_0\|_{L^1(D)},$$

and the conclusion of Theorem 3.3 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 3.3 are natural.

Theorem 3.4 *There exists a universal constant $C_0 > 0$, such that, if $u(y) = \bar{u} = const$ 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*

$$\|\phi(t=1,\cdot)\|_{L^\infty(D)} \geq C_0, \tag{3.18}$$

for all $\varepsilon \in (0,1)$.

The proof is quite simple: solution of (3.16) is above the solution of the Dirichlet problem in the smaller channel $D' = \{(x,y) : \ x \in \mathbb{R}, \ y \in B(a,h)\}$:

$$\begin{aligned} &\phi_t + \frac{1}{\varepsilon} u(y)\phi_x = \Delta\phi \text{ in } D', \\ &\phi(0,x,y) = \phi_0(x,y) \\ &\phi = 0 \ \text{ on } \partial D'. \end{aligned} \tag{3.19}$$

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

$$\begin{aligned} &\psi_t = \Delta\phi \text{ in } D', \\ &\phi(0,x,y) = \phi_0(x,y) \\ &\psi = 0 \ \text{ on } \partial D'. \end{aligned} \tag{3.20}$$

The conclusion of Theorem 3.4 follows simply from the fact that the function ψ does not depend on ε.

4 A limit theorem for a particle in a random flow

Let us now start adding perturbations that are not just white in time, which is what we have done so far in our considerations of the evolution of particles in incompressible flows perturbed by a diffusion. We will go just one level up in difficulty compared to the toy problem

$$\frac{dX}{dt} = \varepsilon V(t), \tag{4.1}$$

considered in the introduction. To illustrate how the background dynamics improves mixing in a very simple setting, we look at a particle moving in a time-independent random flow with a large mean:

$$\frac{dX}{dt} = \bar{u} + \varepsilon v(X). \tag{4.2}$$

An important remark is that, again, as in another example in the introduction, of an advection equation with a potential, a non-zero background flow $\bar{u} \neq 0$ is absolutely essential for the results of this section: very little is known about the solutions of the ODE

$$\dot{X}(t) = \varepsilon v(X), \tag{4.3}$$

with a random velocity field $v(X)$. In particular, $\varepsilon > 0$ here is superfluous: one can remove this factor by a simple rescaling of time $t \to t/\varepsilon$, to get

$$\dot{X}(t) = v(X). \tag{4.4}$$

The first order PDE corresponding to (4.2) is

$$\phi_t + (\bar{u} + \varepsilon v(x)) \cdot \nabla \phi = 0. \tag{4.5}$$

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

$$\phi_t + \frac{1}{\varepsilon^2}(\bar{u} + \varepsilon v(x)) \cdot \nabla \phi = 0. \tag{4.6}$$

The background dynamics is very simple:

$$\bar{\phi}_t + \frac{1}{\varepsilon^2} \bar{u} \cdot \nabla \bar{\phi} = 0, \tag{4.7}$$

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 (1.33):

$$\phi(t, x) = \psi(t, x - \bar{u}\frac{t}{\varepsilon^2}). \tag{4.8}$$

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

$$\psi_t + \frac{1}{\varepsilon} v(x + \bar{u}\frac{t}{\varepsilon^2}) \cdot \nabla \psi = 0. \tag{4.9}$$

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:

$$\dot{X} = \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, X\right), \quad X(0) = x, \tag{4.10}$$

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

$$X(t) = \frac{1}{\varepsilon} \int_0^t V(\frac{s}{\varepsilon^2}) ds = \varepsilon \int_0^{t/\varepsilon^2} V(s) ds, \tag{4.11}$$

and $X(t)$ converges in law to a Brownian motion, according to Theorem 1.1. In the general case, when $V(t, x)$ is not spatially uniform, this question goes back to the papers by Khasminskii [38] from the 60's with subsequent contributions by various authors: without any

attempt at completeness we mention the work of Papanicolaou and Kohler [50], and Kesten and Papanicolaou [36]. We present a version of the limit theorem due to T. Komorowski [42].

Let us explain where the scaling in (4.10) 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 $X = Y/x_0$, $s = T/t_0$:

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

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 (4.10). The limit $\varepsilon \to 0$ 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 (4.10) is

$$\frac{\partial \phi}{\partial t} - \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, x\right) \cdot \nabla \phi = 0, \quad \phi(0,x) = \phi_0(x). \tag{4.12}$$

Its solution is $\phi(t,x) = \phi_0(X(t;x))$, where $X(t;x)$ is the solution of (4.10).

When does one expect the trajectories of (4.10) 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 $t_1, t_2, \ldots, t_m \in \mathbb{R}$, $x_1, \ldots, x_m \in \mathbb{R}^n$, and each $h \in \mathbb{R}$ and $y \in \mathbb{R}^n$ the joint distribution of

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

is the same as that of

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

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

$$R_{nm}(t,x) = \mathbb{E}\left\{V_n(s,y) V_m(t+s, y+x)\right\}. \tag{4.13}$$

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 σ-algebras $\tilde{\mathcal{V}}_a^b$ generated by the sets of the form

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

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

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

The mixing assumption would be that or any $m\ge 0$ the mixing coefficient satisfies

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

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

$$V(t,x)=v(x+\bar{u}t), \tag{4.14}$$

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

$$V(t,x)=V(t+h,y-h\bar{u}), \tag{4.15}$$

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 (4.14). On the formal level, this is reflected in the fact that all σ-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 (4.14), 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 (4.15) 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 $I_{ab}=a\le t\le b$, consider the sets

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

We denote by $\mathcal{V}_a^b$ the σ-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

$$\beta(h)=\sup_{0\le t\le 1+h^{3/2}}\sup_{A\in\mathcal{V}_{t+h}^{\infty},B\in\mathcal{V}_0^t}\frac{|P(A\cap B)-P(A)P(B)|}{P(B)}, \tag{4.16}$$

and our mixing assumption is that it satisfies

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

Let us see why this mixing condition is reasonable for velocity fields of the form (4.14), and why we have the restriction

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

in the supremum in (4.16). Consider two space-time points $(s_1, x_1) \in S_0^t$, and $(s_2, x_2) \in S_{t+h}^\infty$, with $0 \le t \le 1 + h^{3/2}$, then

$$V(s_1, x_1) = v(x_1 + s_1\bar{u}), \quad V(s_2, x_2) = v(x_2 + s_2\bar{u}), \tag{4.17}$$

and

$$\begin{aligned} d &= |x_2 + s_2\bar{u} - (x_1 + s_1\bar{u})| \ge (s_2 - s_1)|\bar{u}| - |x_2| - |x_1| \\ &\ge (s_2 - s_1)|\bar{u}| - C(1 + \sqrt{s_2}) - C(1 + \sqrt{s_1}) \ge (s_2 - s_1)|\bar{u}| - 2C(1 + \sqrt{s_2}). \end{aligned} \tag{4.18}$$

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

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

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

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

Thus, the distance between the points entering $V(s_1, x_1)$ and $V(s_2, x_2)$ in (4.17) is large, and spatial decorrlation 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 [36] 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

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

for all $1 \le i, j, l \le n$. 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

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

and its symmetric non-negative definite square-root matrix σ: $\sigma^2 = a$. 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*

$$\dot{X}_\varepsilon = \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, X_\varepsilon\right), \quad X_\varepsilon(0) = x, \tag{4.19}$$

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

The main result of [42] 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 [10] 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 Ω converges weakly to a Borel measure P on Ω if for every bounded continuous real function f we have

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

Equivalently, for every set A with $P(\partial A) = 0$ we have $P_n(A) \to P(A)$. A family F of (Borel) probability measures on Ω 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 distirbutions 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 $P_n = \varepsilon_{z_n}$ and let $P = \varepsilon_0$, 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 $t_1, \ldots, t_k$ 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 $i = 1, \ldots, k$ such that $t_i > 0$ we have $P_n(A) = P(A)$ simply because $z_n(t_j) = z(t_j)$ for all $j = 1, \ldots, k$ (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 advanatage of tightness is that it is a verifiable notion by means of various moduli of continuity. The usual modulus of coninuity 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

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

The following theorem (Theorem 7.3 in [10]) 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*

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

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

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

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:

$$\begin{aligned} &\text{(i) For } 0 \leq t < 1 \text{ the right limit } x(t^+) = \lim_{s \to t^+} x(s) \text{ exists and } x(t) = x(t^+). \\ &\text{(ii) For } 0 < t \leq 1 \text{ the left limit } x(t^-) = \lim_{s \to t^-} x(s). \end{aligned} \tag{5.3}$$

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 $0 = t_0 < t_1 < \cdots < t_n = 1$ 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

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

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

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

This metric defines the Skorohod topology.

Let Λ 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

$$|x_n(t) - x(t)| \le |x_n(t) - x(\lambda_n(t))| + |x(\lambda_n(t)) - x(t)|, \tag{5.6}$$

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 $0 \le t \le 1/2^n$ 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 λ_n maps $[0, 1/2^n]$ onto $[0, 1/2^{n+1}]$. Then $x_{n+1}(\lambda_n(t)) = x_n(t)$ and $|\lambda_n(t) - t| \le 1/2^{n+1}$. This means that $d(x_n, x_{n+1}) \le 1/2^{n+1}$ 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 λ 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 $\varepsilon \ge 0$ 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 λ is close to identity in the uniform norm but the slopes of λ are all close to one. In particular, the above example of a non-converging Cauchy sequence involves λ_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 $d(x_n, x) \to 0$ if and only if $d_0(x_n, x) \to 0$. 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 $w_x(\varepsilon)$ does not vanish in the limit $\varepsilon \to 0$. 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

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

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

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

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

$$\lim_{\varepsilon \to 0} w'_x(\varepsilon) = 0,$$

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

$$w'_x(\varepsilon) \le w_x(2\varepsilon).$$

There can be no inequality in the opposite direction because the usual modulus of continuity $w_x(\varepsilon)$ does not go to zero as $\varepsilon \to 0$ for a discontinuous function from D. However, for a continuous function $x(t)$ we do have an inequality $w_x(\varepsilon) \le 2w'_x(\varepsilon)$ 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 $\lim_{\varepsilon \to 0} \sup_{x \in A} w'_x(\varepsilon) = 0$.*

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*

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

and

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

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

$$w''_x(\varepsilon) = \sup_{0\le u-s\le\varepsilon}\left[\sup_{s\le t\le u}\left(\min\left[|x(u)-x(t)|, |x(t)-x(s)|\right]\right)\right].$$

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

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

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

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

Nevertheless, this is the only obstacle for a compactness criterion in terms of $w''_x(\varepsilon)$ 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* $\sup_{x\in A}\|x\| < \infty$, $\lim_{\varepsilon\to 0}\sup_{x\in A} w''_x(\varepsilon) = 0$ *and*

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

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*

$$(i) \quad \lim_{a\to\infty}\limsup_n P_n\left[x: \ \|x\| \ge a\right] = 0,$$

and

$$(ii.1) \quad \lim_{\varepsilon\to 0}\limsup_n P_n\left[x: \ w''_x(\varepsilon) \ge \varepsilon\right] = 0 \text{ for all } \varepsilon > 0,$$

and

$$(ii.2) \quad \begin{cases} \lim_{\varepsilon\to 0}\limsup_n P_n\left[x: \ |x(\varepsilon)-x(0)| \ge \varepsilon\right] = 0 \\ \lim_{\varepsilon\to 0}\limsup_n P_n\left[x: \ |x(1^-)-x(1-\varepsilon)| \ge \varepsilon\right] = 0. \end{cases}$$

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

$$J_t = \{x \in D: \ x(t) \ne x(t^-)\}$$

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 distirbutions $(X^n_{t_1}, \dots, X^n_{t_k})$ of random variables X^n defined on D converge weakly as $n \to \infty$ to $(X_{t_1}, \dots, X_{t_k})$ whenever all t_i lie in T_X, and $X_1 - X_{1-\varepsilon}$ goes weakly to zero as $\varepsilon \to 0$. 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*

$$P\left[\min\{|X^n_s - X^n_r|, |X^n_t - X^n_s|\} \geq \lambda\right] \leq \frac{C}{\lambda^{4\beta}} |F(t) - F(s)|^{2\alpha}, \tag{5.7}$$

where F is a non-decreasing continuous function on $[0,1]$. Then X_n converge weakly to X as $n \to \infty$.

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, α and β so that

$$P\left[w''_{X^n}(\varepsilon) \geq \varepsilon\right] \leq \frac{K}{\varepsilon^{4\beta}} (F(1) - F(0)) [w_F(2\varepsilon)]^{2\alpha - 1}, \tag{5.8}$$

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

$$\mathbb{E}\left\{|X^n_s - X^n_r|^{2\beta} |X^n_t - X^n_s|^{2\beta}\right\} \leq C|t - r|^{2\alpha} \tag{5.9}$$

for all n. Then we may take $F(t) = t$ and (5.8) becomes

$$P\left[w''_{X^n}(\varepsilon) \geq \varepsilon\right] \leq \frac{K}{\varepsilon^{4\beta}} \varepsilon^{2\alpha - 1}. \tag{5.10}$$

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

$$\mathbb{E}\left\{|X_n(u) - X_n(t)|^2 |X_n(t) - X_n(s)|^\nu\right\} \leq C(T, \nu)(u - t)\mathbb{E}\left\{|X_n(t) - X_n(s)|^\nu\right\}. \tag{5.11}$$

Indeed, when $\nu = 0$ in (5.11) we have

$$\mathbb{E}\left\{|X_n(u) - X_n(t)|^2\right\} \leq C(T, \nu)(u - t) \text{ for all } n \text{ and all } 0 \leq t \leq u \leq T.$$

Taking $\nu = 2$ in (5.11) we get, using the above:

$$\begin{aligned} &\mathbb{E}\left\{|X_n(u) - X_n(t)|^2 |X_n(t) - X_n(s)|^2\right\} \leq C(T, \nu)(u - t)\mathbb{E}\left\{|X_n(t) - X_n(s)|^2\right\} \\ &\leq C(T, \nu)(u - t)(t - s) \leq C(T, \nu)(u - s)^2 \end{aligned} \tag{5.12}$$

and thus (5.9) indeed holds. A somehwat more general estimate than (5.11) is a reformulation in terms of the conditional expectation

$$\mathbb{E}\left\{|X_n(t) - X_n(s)|^2 \middle| \mathcal{F}_s\right\} \leq C(T)(t - s). \tag{5.13}$$

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

$$\dot{X}_\varepsilon = \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, X_\varepsilon\right), \quad X_\varepsilon(0) = x. \tag{6.1}$$

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 ε^{-1} in front of the velocity field. Understanding 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 main step is to establish tightness of the processes $X_\varepsilon(t)$, so that a limit in law along a subsequence $\varepsilon_k \to 0$ 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 σ. This uses the martingale characterization of the Brownian motion.

Tightness of X_ε is a consequence of the following.

Proposition 6.1 *There exist $C > 0$ and $\nu > 0$ so that*

$$\mathbb{E}\left\{|X_\varepsilon(t) - X_\varepsilon(s)|^2 |X_\varepsilon(s) - X_\varepsilon(u)|^2\right\} \leq C(t-u)^{1+\nu}, \tag{6.2}$$

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

$$\mathbb{E}\left\{|X_\varepsilon(t) - X_\varepsilon(s)|^2 \,\middle|\, \mathcal{F}_s\right\} \leq C(t-s) \text{ for } t - s > 10\varepsilon^\gamma. \tag{6.3}$$

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

$$|X_\varepsilon(t) - X_\varepsilon(s)| \le \frac{C(t-s)}{\varepsilon}.$$

Therefore, the following estimate holds for such times t, s and u:

$$\begin{aligned}&\mathbb{E}\left\{|X_\varepsilon(t) - X_\varepsilon(s)|^2 |X_\varepsilon(s) - X_\varepsilon(u)|^2\right\} \le \frac{C}{\varepsilon^2}(t-s)^2(s-u)\\ &\le C\varepsilon^{7\gamma/4-2}(t-u)^{5/4} \le C(t-u)^{5/4},\end{aligned}$$

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 $T \ge 0$ and let $0 \le u \le s \le T$. 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 $0 \le u \le s \le s_2 \le s_1 \le T$ and $0 < \varepsilon < \varepsilon_0$ we have*

$$\left|\mathbb{E}\left\{V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right) Y\left(\frac{s}{\varepsilon^2}\right)\right\}\right| \le C\beta(s_1 - s)\mathbb{E}\left|Y\left(\frac{s}{\varepsilon^2}\right)\right|, \tag{6.4}$$

$$\left|\mathbb{E}\left\{\frac{\partial}{\partial x_k}\left[V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right)\right] Y\left(\frac{s}{\varepsilon^2}\right)\right\}\right| \le C\beta(s_1 - s)\mathbb{E}\left|Y\left(\frac{s}{\varepsilon^2}\right)\right|, \tag{6.5}$$

and

$$\left|\mathbb{E}\left\{G_1\left(\frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u)\right) Y\left(\frac{s}{\varepsilon^2}\right)\right\}\right| \le C\beta^{1/2}(s_1 - s_2)\beta^{1/2}(s_2 - s)\mathbb{E}\left|Y\left(\frac{s}{\varepsilon^2}\right)\right|, \tag{6.6}$$

$$\left|\mathbb{E}\left\{\frac{\partial}{\partial x_k}\left[G_1\left(\frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u)\right)\right] Y\left(\frac{s}{\varepsilon^2}\right)\right\}\right| \le C\beta^{1/2}(s_1 - s_2)\beta^{1/2}(s_2 - s)\mathbb{E}\left|Y\left(\frac{s}{\varepsilon^2}\right)\right|, \tag{6.7}$$

for all $1 \le k \le n$.

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 $X_\varepsilon(t)$, $0 \le t \le u \le T$ 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:

$$|X_\varepsilon(t)| \le \frac{1}{\varepsilon}\int_0^u \left|V\left(\frac{s}{\varepsilon^2}, X_\varepsilon(s)\right)\right| ds \le \frac{Cu}{\varepsilon} \le C\left(1 + \frac{u^\rho}{\varepsilon^{2\rho}}\right)$$

for all $0 \le t \le u$.

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

$$A(l) = \left[\omega : \ \frac{l_j}{M} \le X_j^\varepsilon(u) < \frac{l_j+1}{M}, \ \ j = 1, \dots, n\right], \ \ l = (l_1, \dots, l_n).$$

The event $A(l)$ is $\mathcal{V}_0^{s/\varepsilon^2}$ measurable since $u \le s$. 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:

$$\left|\mathbb{E}\left\{\frac{\partial}{\partial x_k}\left[V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right)\right] Y\left(\frac{s}{\varepsilon^2}\right)\right\}\right| = \left|\sum_l \mathbb{E}\left\{\frac{\partial}{\partial x_k}\left[V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right)\right] Y\left(\frac{s}{\varepsilon^2}\right)\chi_{A(l)}\right\}\right|$$

$$\le \left|\sum_l \mathbb{E}\left\{\left[\frac{\partial}{\partial x_k}\left[V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right)\right] - \frac{\partial}{\partial x_k}\left[V\left(\frac{s_1}{\varepsilon^2}, \frac{l}{M}\right)\right]\right] Y\left(\frac{s}{\varepsilon^2}\right)\chi_{A(l)}\right\}\right|$$

$$+ \left|\sum_l \mathbb{E}\left\{\frac{\partial}{\partial x_k}\left[V\left(\frac{s_1}{\varepsilon^2}, \frac{l}{M}\right)\right] Y\left(\frac{s}{\varepsilon^2}\right)\chi_{A(l)}\right\}\right| = I + II.$$

As the points l/M are deterministic, the second term above may be now estimated using the mixing property (4.16) and the fact that $\mathbb{E}\{\partial V/\partial x_k\} = 0$ by

$$II \le 2K\beta\left(\frac{s_1 - s}{\varepsilon^2}\right)\sum_l \mathbb{E}\left\{\left|Y\left(\frac{s}{\varepsilon^2}\right)\right|\chi_{A(l)}\right\} = 2K\beta\left(\frac{s_1 - s}{\varepsilon^2}\right)\mathbb{E}\left\{\left|Y\left(\frac{s}{\varepsilon^2}\right)\right|\right\},$$

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

$$I \le \frac{C}{M}\sum_l \mathbb{E}\left\{\left|Y\left(\frac{s}{\varepsilon^2}\right)\right|\chi_{A(l)}\right\} = \frac{C}{M}\sum_l \mathbb{E}\left\{\left|Y\left(\frac{s}{\varepsilon^2}\right)\right|\right\} \to 0 \text{ as } M \to +\infty,$$

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:

$$\left|\mathbb{E}\left\{\frac{\partial^2}{\partial x_k \partial x_m}\left[V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right)\right] Y\left(\frac{s}{\varepsilon^2}\right)\right\}\right| \le C\beta(s_1 - s)\mathbb{E}\left|Y\left(\frac{s}{\varepsilon^2}\right)\right|. \tag{6.8}$$

We now prove (6.7) – the proof of (6.6) is identical. Let us first write out the expression for G_1 :

$$\left|\mathbb{E}\left\{\frac{\partial}{\partial x_k}\left[G_1\left(\frac{s_1}{\varepsilon^2},\frac{s_2}{\varepsilon^2},X_\varepsilon(u)\right)\right]Y\left(\frac{s}{\varepsilon^2}\right)\right\}\right|$$
$$\leq\sum_{p=1}^{n}\left|\mathbb{E}\left\{\frac{\partial}{\partial x_k}\left[V_p\left(\frac{s_2}{\varepsilon^2},X_\varepsilon(u)\right)\frac{\partial}{\partial x_p}\left(V\left(\frac{s_1}{\varepsilon^2},X_\varepsilon(u)\right)\right)\right]Y\left(\frac{s}{\varepsilon^2}\right)\right\}\right|$$

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 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 conculde that (6.7) holds. This finishes the proof of Lemma 6.3. □

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 $s=t_0<t_1<\cdots<t_{M+1}=t$ of the interval $[s,t]$ into subintervals of the length

$$\varepsilon t=l_\varepsilon=(t-s)\left(\left[\frac{t-s}{\varepsilon^\gamma}\right]\right)^{-1},$$

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 εt is much larger than the velocity correlation time ε^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 $O(\varepsilon t)$. 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 ε of X_ε we write for $t>s$:

$$X(t)-X(s)=\frac{1}{\varepsilon}\int_s^t V\left(\frac{u}{\varepsilon^2},X(u)\right)du=\frac{1}{\varepsilon}\sum_{i=0}^{M}\int_{t_i}^{t_{i+1}}V\left(\frac{u}{\varepsilon^2},X(u)\right)du \tag{6.9}$$

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 $t_i \le u \le t_{i+1}$ 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

$$G_2(u, u_1, u_2, x) = \sum_{q=1}^{n} \frac{\partial}{\partial x_q}\left[G_1\left(u, u_1, x\right)\right] V_q\left(u_2, x\right).$$

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 γ: the time interval in each integration is smaller than ε^γ and the total number of terms is at most $2(t-s)/\varepsilon^\gamma$ as we have assumed that $t - s \ge 10\varepsilon^\gamma$. Therefore, the last integral is bounded by

$$\frac{1}{\varepsilon^3}\left|\sum_{i=0}^{M}\int_{t_i}^{t_{i+1}}\left[\int_{t_{i-1}}^{u}\left[\int_{t_{i-1}}^{u_1} G_2\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, \frac{u_2}{\varepsilon^2}, X(u_2)\right) du_2\right] du_1\right] du\right| \le C\varepsilon^{2\gamma-3}(t-s)$$

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 $t \geq s + 10\varepsilon^{\gamma}$:

$$\mathbb{E}\left\{|X(t) - X(s)|^2 Y\right\} \leq C(T)(t-s)\mathbb{E}\left\{Y\right\}.$$

Our estimates in Step 1 show that it is actually enough to verify that

$$\mathbb{E}\left\{(L_j(s,t))^2 Y\right\} \leq C(t-s)\mathbb{E}\{Y\}, \quad j = 1, 2.$$

An estimate for L_1. We first look at the term corresponding to L_1: it is equal to

$$\mathbb{E}\left\{(L_1(s,t))^2 Y\right\} = \frac{2}{\varepsilon^2}\sum_{i<j}\sum_{p=1}^{n}\int_{t_i}^{t_{i+1}}\int_{t_j}^{t_{j+1}} \mathbb{E}\left\{V_p\left(\frac{u}{\varepsilon^2}, X(t_{i-1})\right) V_p\left(\frac{u'}{\varepsilon^2}, X(t_{j-1})\right) Y\right\} du'du$$

$$+\frac{1}{\varepsilon^2}\sum_{j}\sum_{p=1}^{n}\int_{t_j}^{t_{j+1}}\int_{t_j}^{t_{j+1}} \mathbb{E}\left\{V_p\left(\frac{u}{\varepsilon^2}, X(t_{j-1})\right) V_p\left(\frac{u'}{\varepsilon^2}, X(t_{j-1})\right) Y\right\} du'du = \sum_{i\leq j} I_{ij}.$$

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 ε^2:

$$\sum_{j=0}^{M}\sum_{i\leq j-2}|I_{ij}| \leq \frac{C}{\varepsilon^2}\sum_{j=0}^{M}\sum_{i\leq j-2}\int_{t_i}^{t_{i+1}}\int_{t_j}^{t_{j+1}} \beta\left(\frac{u'-t_{j-1}}{\varepsilon^2}\right)\mathbb{E}\left\{Y\right\} du'du$$

$$\leq \frac{C}{\varepsilon^2}\beta\left(\varepsilon^{-2+\gamma}\right)(t-s)^2 E\left\{Y\right\} \leq C\varepsilon^p(t-s)E\left\{Y\right\}$$

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':

$$\sum_{j=0}^{M}|I_{jj}|\le\frac{C}{\varepsilon^2}\sum_{j=0}^{M}\int\limits_{t_j}^{t_{j+1}}\int\limits_{t_j}^{t_{j+1}}\mathbb{E}\left\{V_p\left(\frac{u}{\varepsilon^2},X(t_{j-1})\right)V_p\left(\frac{u'}{\varepsilon^2},X(t_{j-1})\right)Y\right\}du'du \tag{6.10}$$

$$\le\frac{2C}{\varepsilon^2}\sum_{j\in I}\int\limits_{t_j}^{t_{j+1}}\int\limits_{u'}^{t_{j+1}}\beta\left(\frac{u-u'}{\varepsilon^2}\right)dudu'E\left\{Y\right\}\le C(t-s)E\left\{Y\right\}\int_0^\infty\beta(u)du.$$

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 $1\le j,m\le n$. 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

$$|I_1|\le C\varepsilon^\alpha(t-s)E\{Y\},$$

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:

$$\sum_{j\in I}|I_{jj}|=\frac{1}{\varepsilon^2}\sum_{j=0}^{M}\int\limits_{t_j}^{t_{j+1}}\int\limits_{t_j}^{t_{j+1}}\mathbb{E}\left\{V_j\left(\frac{u}{\varepsilon^2},X(t_{j-1})\right)V_m\left(\frac{u'}{\varepsilon^2},X(t_{j-1})\right)Y\right\}du'du \tag{6.11}$$

$$=\frac{1}{\varepsilon^2}\sum_{j=0}^{M}\int\limits_{t_j}^{t_{j+1}}\int\limits_{t_j}^{t_{j+1}}R_{jm}\left(\frac{u-u'}{\varepsilon^2},0\right)dudu'E\left\{Y\right\}+o(1)(t-s)E\left\{Y\right\}$$

$$=\left[\int_{-\infty}^{\infty}R_{jm}(\tau,0)d\tau+o(1)\right](t-s)E\left\{Y\right\}.$$

Finally, I_2 corresponding to $i=j-1$ is computed as

$$\sum_{j\in I}|I_{j-1,j}|=\frac{1}{\varepsilon^2}\sum_{j=0}^{M}\int\limits_{t_j}^{t_{j+1}}\int\limits_{t_{j-1}}^{t_j}\mathbb{E}\left\{V_j\left(\frac{u}{\varepsilon^2},X(t_{j-1})\right)V_m\left(\frac{u'}{\varepsilon^2},X(t_{j-2})\right)Y\right\}du'du \tag{6.12}$$

$$=\frac{1}{\varepsilon^2}\sum_{j\in I}\int\limits_{t_j}^{t_{j+1}}\int\limits_{t_{j-1}}^{t_j}R_{jm}\left(\frac{u-u'}{\varepsilon^2},0\right)dudu'E\left\{Y\right\}+o(1)(t-s)E\left\{Y\right\}=o(1)(t-s)E\left\{Y\right\}.$$

because $t_{j+1}-t_j=\varepsilon^\gamma\gg\varepsilon^2$. Therefore we actually have a more precise estimate

$$\mathbb{E}\left\{(L_{1,j}(s,t)L_{1,m}(s,t))Y\right\}=\left[\int_{-\infty}^{\infty}R_{jm}(\tau,0)d\tau+o(1)\right](t-s)E\left\{Y\right\}. \tag{6.13}$$

An estimate for L_2. Following the above steps one also establishes the required estimate for L_2:

$$\mathbb{E}\left\{(L_2(s,t))^2Y\right\}\le C(t-s)E\{Y\}. \tag{6.14}$$

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 γ:

$$\mathbb{E}\left\{(L_2(s,t))^2 Y\right\} \le C\varepsilon^{\alpha}(t-s)E\{Y\} \tag{6.15}$$

with $\alpha > 0$. We will need (6.15) in the identification of the limit, thus we will show it now:

$$\begin{aligned}&\mathbb{E}\left\{(L_2(s,t))^2 Y\right\}\\ &= \frac{1}{\varepsilon^4}\sum_{i,j}\int_{t_i}^{t_{i+1}} du \int_{t_j}^{t_{j+1}} du' \int_{t_{i-1}}^{u} du_1 \int_{t_{j-1}}^{u'} du_1' \mathbb{E}\left\{G_1\left(\frac{u}{\varepsilon^2},\frac{u_1}{\varepsilon^2},X(t_{i-1})\right) G_1\left(\frac{u'}{\varepsilon^2},\frac{u_1'}{\varepsilon^2},X(t_{j-1})\right) Y\right\}.\end{aligned}$$

Once again, you split the sum above into terms with $i \le j-2$, $i = j-1$ and $i = j$: those with $i \le j-2$ add up to

$$\begin{aligned}&\frac{1}{\varepsilon^4}\sum_{i\le j-2}\int_{t_i}^{t_{i+1}} du \int_{t_j}^{t_{j+1}} du' \int_{t_{i-1}}^{u} du_1 \int_{t_{j-1}}^{u'} du_1' \mathbb{E}\left\{G_1\left(\frac{u}{\varepsilon^2},\frac{u_1}{\varepsilon^2},X(t_{i-1})\right) G_1\left(\frac{u'}{\varepsilon^2},\frac{u_1'}{\varepsilon^2},X(t_{j-1})\right) Y\right\}\\ &\le C\varepsilon^{2\gamma-4}\beta\left(\varepsilon^{\gamma-2}\right)(t-s)^2\mathbb{E}\{Y\}.\end{aligned}$$

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 ε^γ 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:

$$\begin{aligned}&\frac{1}{\varepsilon^4}\sum_{i}\int_{t_i}^{t_{i+1}} du \int_{t_i}^{t_{i+1}} du' \int_{t_{i-1}}^{u} du_1 \int_{t_{i-1}}^{u'} du_1' \mathbb{E}\left\{G_1\left(\frac{u}{\varepsilon^2},\frac{u_1}{\varepsilon^2},X(t_{i-1})\right) G_1\left(\frac{u'}{\varepsilon^2},\frac{u_1'}{\varepsilon^2},X(t_{i-1})\right) Y\right\}\\ &\qquad\le C\varepsilon^{3\gamma-4}(t-s)\mathbb{E}\{Y\}\end{aligned}$$

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

$$\mathbb{E}\left\{(X_m(t)-X_m(s))(X_n(t)-X_n(s))Y\right\} = \left[\int_{-\infty}^{\infty} R_{mn}(\tau,0)d\tau + o(1)\right](t-s)E\left\{Y\right\} \tag{6.16}$$

for all $t > s$ with $t - s \ge 10\varepsilon^\gamma$. 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,

$$\lim_{\varepsilon\to 0}\mathbb{E}\left\{\left[(X_j^\varepsilon(t)-X_j^\varepsilon(s))(X_m^\varepsilon(t)-X_m^\varepsilon(s)) - a_{jm}(t-s)\right]\Psi\right\} = 0$$

for all bounded non-negative continuous functions

$$\Psi = \Psi(X_\varepsilon(t_1), \dots, X_\varepsilon(t_n)),$$

with $0 \le t_1 \le t_2 \le \cdots \le t_n \le s < t \le T$. Second, we need

$$\limsup_{\varepsilon \to 0} \mathbb{E}\left\{(X_j^\varepsilon(t))^4\right\} < +\infty$$

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 The random Schrödinger equation: geometric optics

We will now consider the Schrödinger equation with a weak random potential:

$$i\psi_t + \frac{1}{2}\Delta\psi - \varepsilon V(t,x)\psi = 0. \tag{7.1}$$

The random potential may be either time-dependent or time-independent, so that $V = V(x)$. We will always assume that V is a mean-zero random field, that is either space-time stationary if $V = V(t,x)$ or space stationary if $V = V(x)$. Typically, the problems with a time-dependent potential are easier since one can use mixing in time, in addition to the spatial mixing. As usual, we will denote by $R(t,x)$ (or $R(x)$ for time-independent potentials) the two-point correlation function of $V(t,x)$ (or $V(x)$).

The Schrödinger equation (7.1) needs to be supplemented by an initial condition, which is important both for the physics and mathematics of the problem. If a typical scale of variations in the initial condition is l_{in} and the correlation function of the random potential has support of size l_{cor}, then we should distinguish three regimes: $l_{in} \ll l_{cor}$ – the initial condition varies on the scales much faster than the random potential. This is known as the random geometric optics regime. One can first pass to the ray description of wave propagation, and then analyze the resulting random classical system. If $l_{in} \sim l_{cor}$ then wave has full interactions with the random inhomogeneities, and energy scattering is described by a kinetic equation. Accordingly, this is known as the kinetic regime. Finally, if $l_{in} \gg l_{cor}$ so that the scale of the random heterogeneities is much smaller than that of the initial condition, then the heterogeneities produce an effective potential, and this situation is known as the homogenization regime. We will discuss these regimes separately.

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

$$i\psi_t + \frac{1}{2}\Delta\psi - V(\varepsilon x)\psi = 0. \tag{7.2}$$

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

$$i\varepsilon\psi_t + \frac{\varepsilon^2}{2}\Delta\psi - V(x)\psi = 0. \tag{7.3}$$

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 ε) but, for $\varepsilon \ll 1$, the function $\psi(t,x)$ oscillates on the scale ε, which is the original microscopic scale in (7.2). Inserting this ansatz for $\psi(t,x)$ into (7.3), we get, in the leading order the eikonal equation

$$S_t + \frac{|\nabla S|^2}{2} + V(x) = 0. \tag{7.4}$$

The next order of powers in ε gives the amplitude equation

$$(|A|^2)_t + \nabla \cdot (|A|^2 \nabla S) = 0. \tag{7.5}$$

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 (7.4), and the amplitude equation (7.5). Then, a direct computation shows that $W(t,x,\xi)$ satisfies the Liouville equation of the classical mechanics

$$\frac{\partial W}{\partial t} + \xi \cdot \nabla_x W - \nabla_x V \cdot \nabla_\xi W = 0. \tag{7.6}$$

The corresponding characteristics are

$$\frac{dX}{dt} = K, \quad \frac{dK}{dt} = -\nabla V(X), \tag{7.7}$$

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 (7.5) 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 (7.6) is linear, and its solutions should not break down. Let us see what happens: as $V = 0$, the Liouville equation is

$$W_t + k \cdot \nabla_x W = 0, \quad W(0,x,k) = W_0(x,k). \tag{7.8}$$

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 (7.8) 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 options 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. As we have seen, typically, a mean zero random fluctuation of size ε will produce a non-trivial effect on a time scale $T_\varepsilon \sim \varepsilon^{-2}$. 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

$$dZ = -Z^2 dt + dB_t. \tag{7.9}$$

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, λ – 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 λ 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 λ 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 ε – 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 λ, 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 λ, 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

$$\frac{\partial W(t,x,\xi)}{\partial t} + \nabla_\xi \omega(\xi) \cdot \nabla_x W(t,x,\xi) = \mathcal{L}_{sc} W(t,x,\xi). \tag{7.10}$$

Here, $\omega(\xi)$ is the dispersion relation of the wave and depends on the particular type of the wave. The left side of (7.10) has nothing to do with the inhomogeneities of the medium[1] 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 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.

Particles in weakly random Hamiltonian flows

As we have seen, random heometric optics for the Schrödinegr equation reduces to the motion of a Newtonian particle (7.7). If the random potential is weak, we are faced with the problem

$$\frac{dX(t)}{dt} = K(t), \quad \frac{dK(t)}{dt} = -\varepsilon \nabla V(X(t)), \tag{7.11}$$

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

$$H_\varepsilon(x,k) = \frac{k^2}{2} + \varepsilon V(x). \tag{7.12}$$

We will assume that $V(x)$ is a spatially homogeneous random process with mean zero and a smooth and rapidly decaying correlation function $R(x)$:

$$\mathbb{E}[V(x)] = 0, \quad R(x) = \mathbb{E}[V(y)V(x+y)]. \tag{7.13}$$

[1]Strictly speaking, this statement assumes that the fluctuations are sufficiently weak so that they do not modify the wave dispersion relation.

One may consider more general Hamiltonians of the form

$$H_\varepsilon(x,k) = H_0(k) + \varepsilon H_1(x,k), \tag{7.14}$$

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

We have previously considered in Section 2 white in time perturbations of a two-dimensional Hamiltonian system that led to a slow diffusion of the particle across the level sets of the Hamiltonian. The present situation is different. If $\varepsilon = 0$ in (7.11), then 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 (7.11) we actually have $K(t) = K(0)$ but that is beside the point. However, unlike for small white-in-time perturbations, here the fluctuation of the Hamiltonian $H_0(k)$ in (7.14) 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 $H_0(k)$. Rather, as the process $K(t)$ stays on the level sets of $H_\varepsilon(x,k)$, in the limit $\varepsilon \to 0$, 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 (7.12), this is simply the sphere $|K(t)| = |K(0)|$.

Thus, we will consider the process

$$K_\varepsilon(t) = K\Big(\frac{t}{\varepsilon^2}\Big), \tag{7.15}$$

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 $\dot{X}(t) = K(t)$ 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

$$X_\varepsilon(t) = \varepsilon^2 X\Big(\frac{t}{\varepsilon^2}\Big) \tag{7.16}$$

converges to

$$\bar{X}(t) = \int_0^t \bar{K}(s)ds. \tag{7.17}$$

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

$$\phi_t + k\cdot\nabla_x\phi - \varepsilon\nabla V(x)\cdot\nabla_k\phi = 0. \tag{7.18}$$

The above result says that

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

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

$$\bar{\phi}_t + k\cdot\nabla_x\bar{\phi} - \frac{\partial}{\partial k_j}\Big(D_{jm}(k)\frac{\partial\bar{\phi}}{\partial k_m}\Big) = 0. \tag{7.19}$$

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. Essentially the role of the background Hamiltonian is ensure that the particles "goes around" and always sees a new randomness, which creates mixing needed for the long time diffusive behavior.

Particles in a random force field: "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 [35]. The characteristics for the Liouville equation (7.18) are

$$\frac{dX}{dt} = -K(t), \quad \frac{dK}{dt} = \varepsilon \nabla V(X(t)), \quad X(0) = x, \quad K(0) = k. \tag{7.20}$$

Let us seek the trajectories $X(t), K(t)$ as a formal perturbation expansion

$$X(t) = X_0(t) + \varepsilon X_1(t) + \varepsilon^2 X_2(t) + \dots, \quad K(t) = K_0(t) + \varepsilon K_1(t) + \varepsilon^2 K_2(t) + \dots.$$

We insert this expansion into the characteristics (7.20), 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 ε is

$$K_1(t) = \int_0^t \nabla V(X_0(s))ds = \int_0^t \nabla V(x - ks)ds, \tag{7.21}$$

and

$$X_1(t) = \int_0^t K_1(s)ds = \int_0^t (t-s)\nabla V(x - ks)ds. \tag{7.22}$$

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 $\varepsilon T^2 \ll 1$, or $T \ll \varepsilon^{-1/2}$ 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 (1.13):

$$\begin{aligned}\langle K_1^2(t)\rangle &= \int_0^t \int_0^t \langle \nabla V(x - ks) \cdot \nabla V(x - ks')\rangle ds ds' \\ &= -\int_0^t \int_0^t \Delta R(k(s-s'))ds ds' = Dt + O(1), \quad \text{as } t \to +\infty,\end{aligned}$$

with the diffusion coefficient

$$D = -\int_{-\infty}^{\infty} \Delta R(ks) ds. \tag{7.23}$$

We used here the correlation matrix for ∇V:

$$\mathbb{E}\Big[\frac{\partial V(y)}{\partial y_i}\frac{\partial V(x+y)}{\partial y_j}\Big] = -\frac{\partial^2 R(x)}{\partial x_i \partial x_j}. \tag{7.24}$$

As in Theorem 1.1, one can show that an appropriate rescaling $K_\delta(t) = \delta K_1(t/\delta^2)$ converges, as $\delta \to 0$, to a Brownian motion with the diffusion matrix

$$D_{ij} = -\int_{-\infty}^{\infty} \frac{\partial^2 R(ks)}{\partial x_i \partial x_j} ds. \tag{7.25}$$

The variance of $X_1(t)$ can also be computed explicitly:

$$\begin{aligned}\langle X_1^2(t)\rangle &= \int_0^t \int_0^t (t-s)(t-s')\langle \nabla V(x-ks)\cdot \nabla V(x-ks')\rangle ds ds' \\ &= -\int_0^t \int_0^t (t-s)(t-s')\Delta R(k(s-s'))ds ds' = \frac{Dt^3}{3} + O(1), \quad \text{as } t \to +\infty,\end{aligned}$$

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

$$\varepsilon^2 T^3 \sim O(1),$$

that is, for times of the order $T \sim \varepsilon^{-2/3}$, which is much longer than the "deterministic prediction" $T \sim \varepsilon^{-1/2}$. One should also note that this time scale is exactly when the stochastic caustic happens according to the analysis of White and Kulkarny that led to (7.9).

Formally, this means that for large times (but much smaller than $\varepsilon^{-2/3}$), the expected value of the solutions of the Liouville equation

$$W_t + k\cdot \nabla_x W - \varepsilon \nabla V(x)\cdot \nabla_k W = 0, \tag{7.26}$$

is well-approximated by the solutions of the Fokker-Planck kinetic equation

$$\frac{\partial \bar{W}}{\partial t} + k\cdot \nabla_x \bar{W} = \varepsilon^2 \sum_{i,j=1}^{n} D_{ij}\frac{\partial^2 \bar{W}}{\partial k_i \partial k_j}, \tag{7.27}$$

that is, $\mathbb{E}[W(t,x,k)] \approx \bar{W}(t,x,k)$. 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 $t \ll \varepsilon^{-2/3}$, 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 (7.20) 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.

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:

$$\frac{dX^\varepsilon}{dt} = \nabla_k H_\varepsilon, \quad \frac{dK^\varepsilon}{dt} = -\nabla_x H_\varepsilon, \quad X^\varepsilon(0) = 0, \quad K^\varepsilon(0) = k_0, \tag{7.28}$$

with a random Hamiltonian of the form $H_\varepsilon(x,k) = H_0(k) + \varepsilon H_1(x,k)$. Here $H_0(k)$ is the background Hamiltonian and $H_1(x,k)$ is a random perturbation, while the small parameter $\varepsilon \ll 1$ measures the relative strength of random fluctuations. This was done in [3] and [43]. 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

$$\frac{\partial \phi}{\partial t} + k \cdot \nabla_x \phi - \varepsilon \nabla V(x) \cdot \nabla_k \phi = 0, \tag{7.29}$$

and the corresponding characteristics

$$\frac{dX}{dt} = K, \quad \frac{dK}{dt} = -\varepsilon \nabla_x V(X), \quad X(0) = 0, \quad K(0) = k_0, \tag{7.30}$$

on the time scale $t \sim \varepsilon^{-2}$. As usual, we will assume that the random potential $V(x)$ is a man-zero statistically homogeneous random field, with a rapidly decaying correlation function $R(x)$:

$$\mathbb{E}[V(x)] = 0, \quad \mathbb{E}[V(y)V(x+y)] = R(x). \tag{7.31}$$

We have already seen that at relatively short times $t \ll \varepsilon^{-2/3}$ the "boosted" deviation $(K(t) - k_0)/\varepsilon$ 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 (7.30) preserves the Hamiltonian

$$H(x,k) = \frac{k^2}{2} + \varepsilon V(x). \tag{7.32}$$

Hence, the law of any possible limit for the process $K_\varepsilon(t) = K(t/\varepsilon^2)$, as $\varepsilon \to 0$, 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 $K_\varepsilon(t)$ tends to a uniform distribution on the sphere $\{|k| = |k_0|\}$ – and this is, indeed, what happens at times $t \gg \varepsilon^{-2}$. However, at an intermediate stage, at times of the order ε^{-2}, the process $K_\varepsilon(t)$ 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 [37] in dimensions higher than two, and later extended to two dimensions with the Poisson distribution of scatterers in [17], and in a general two-dimensional setting in [44]. The rescaled spatial component $X^\varepsilon(t) = \varepsilon^2 X(t/\varepsilon^2)$ 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 ε^{-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

$$H_\varepsilon(x,k) = (c_0 + \varepsilon c_1(x))|k|, \tag{7.33}$$

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 (7.32) can be generalized in a relatively straightforward way – see [43] for details. We stick here with (7.32) solely for the sake of simplicity of presentation.

The Fokker-Planck limit

Let the function $\phi_\varepsilon(t,x,k)$ satisfy the Liouville equation

$$\begin{aligned} &\frac{\partial \phi^\varepsilon}{\partial t} + k \cdot \nabla_x \phi^\varepsilon - \varepsilon \nabla V(x) \cdot \nabla_k \phi^\varepsilon = 0, \\ &\phi^\varepsilon(0,x,k) = \phi_0(\varepsilon^2 x, k). \end{aligned} \tag{7.34}$$

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 $1/\varepsilon^2$ relative to the scale of the variations of the potential. In the terminology of the introduction, this means that $l_c/L = \varepsilon^2$ – or, we choose the particular observation scale $L = l_c/\varepsilon^2$. 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

$$D_{ml}(k) = -\frac{1}{|k|} \int_{-\infty}^{\infty} \frac{\partial^2 R(s\hat{k})}{\partial x_n \partial x_m} ds, \quad m,l = 1,\dots,n. \tag{7.35}$$

Note that if the correlation function is isotropic: $R = R(|x|)$, then D_{mn} has a particularly simple form:

$$D_{ml}(k) = D(\delta_{mn} - \hat{k}_l \hat{k}_m), \quad D = -\frac{2}{|k|} \int_0^\infty \frac{R'(r)}{r} dr, \quad m,l = 1,\dots,n. \tag{7.36}$$

We have the following result.

Theorem 7.1 *Let ϕ^ε be the solution of (7.34), 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*

$$\begin{aligned} &\frac{\partial \bar\phi}{\partial t} + k \cdot \nabla_x \bar\phi = \sum_{m,n=1}^d \frac{\partial}{\partial k_m}\left(D_{mn}(k) \frac{\partial \bar\phi}{\partial k_n}\right) \\ &\bar\phi(0,x,k) = \phi_0(x,k). \end{aligned} \tag{7.37}$$

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 $T \geq T_0$

$$\sup_{(t,x,k)\in[0,T]\times K} \left| \mathbb{E}\phi^\varepsilon\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon^2}, k\right) - \bar\phi(t,x,k)\right| \leq CT(1+\|\phi_0\|_{1,4})\varepsilon^{\alpha_0} \tag{7.38}$$

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 (7.37) is indeed a diffusion process on a sphere $|k| =$ const, or, equivalently, equations (7.37) 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 ε^{-2}.

Note that the microscopic dynamics is reflected in the macroscopic limit in the dependence of the diffusion matrix $D_{mn}(k)$ on k – this, in particular, ensures that the limit diffusion stays on the sphere, and, in a sense, is similar to our previous observations about the non-trivial interactions between the background dynamics and fluctuations that lead to a macroscopic limit.

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 ε^{-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 (7.37) 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 (7.37) 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:

$$\begin{aligned} &\frac{\partial w}{\partial t} = \sum_{m,n=1}^d a_{mn}(|k|)\frac{\partial^2 w}{\partial x_n \partial x_m}, \\ &w(0,x,|k|) = \bar\phi_0(x,|k|) \end{aligned} \tag{7.39}$$

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 Γ_n is the area of an n-dimensional sphere. The diffusion matrix $A := [a_{nm}]$ in (7.39) is given explicitly as

$$a_{ij}(k) = \frac{|k|^2}{\Gamma_{n-1}} \int_{\mathbb{S}^{n-1}} \hat{k}_i \chi_j(k) d\Omega(\hat{k}). \tag{7.40}$$

The functions χ_j appearing above are the mean-zero solutions of

$$\sum_{m,i=1}^{d} \frac{\partial}{\partial k_m} \left(D_{mi}(k) \frac{\partial \chi_j}{\partial k_i} \right) = -\hat{k}_j, \tag{7.41}$$

and when the correlation function $R(x)$ is isotropic, so that D_{mi} is given by (7.36), 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 (7.39) becomes the standard diffusion equation

$$\frac{\partial w}{\partial t} = \bar{a}(|k|)\varepsilon_x w, \tag{7.42}$$

with an appropriate diffusion constant $\bar{a}$.

Theorem 7.2 *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 (7.37) converges as $\gamma \to 0$ 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*

$$\|w(t,\cdot) - \bar{\phi}_\gamma(t,\cdot)\|_{L^\infty} \leq C_0 \left(\gamma T + \sqrt{\gamma}\right) \|\phi_0\|_{C^1}, \tag{7.43}$$

for all $T_* \leq t \leq T$.

The proof of Theorem 7.2 is based on classical asymptotic expansions and is quite straightforward. As an immediate corollary of Theorems 7.1 and 7.2, we obtain the following result.

Theorem 7.3 *Let ϕ_ε be solution of (7.34) with the initial data $\phi_\varepsilon(0,x,k) = \phi_0(\varepsilon^{2+\alpha}x, k)$ and let $\bar{w}(t,x)$ be the solution of the diffusion equation (7.39) 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 $K \subset \mathcal{A}(M)$:*

$$\sup_{(t,x,k)\in[T_*,T]\times K} \left| w(t,x,k) - \mathbb{E}\bar{\phi}_\varepsilon(t,x,k) \right| \leq CT\varepsilon^{\alpha_0-\alpha}, \tag{7.44}$$

where $\bar{\phi}_\varepsilon(t,x,k) := \phi_\varepsilon\left(t/\varepsilon^{2+2\alpha}, x/\varepsilon^{2+\alpha}, k\right)$.

Theorem 7.3 shows that if the initial data varies on a scale slightly larger than ε^{-2} then we observe spatial diffusion for the solution (and uniform distribution in k) on the appropriate

time scale. The requirement that α 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

$$\frac{\partial \phi}{\partial t} + k \cdot \nabla_x \phi - \varepsilon V(x) \cdot \nabla_x \phi = 0, \tag{7.45}$$

varies on the scale ε^{-2}: $\phi(0,x) = \phi_0(\varepsilon^2 x, k)$, then on the time scale $t \sim \varepsilon^{-2}$ the expectation of the rescaled solution $\phi_\varepsilon(t,x,k) = \phi(t/\varepsilon^2, x/\varepsilon^2, k)$ converges to the solution of the Fokker-Planck equation. On the other hand, if the initial data varies on an even larger scale: $\phi(0,x,k) = \phi(\varepsilon^{2+\alpha}x, k)$ then on the time scale $t \sim \varepsilon^{-2-2\alpha}$ the expectation of the rescaled field $\phi_\varepsilon(t,x,k) = \phi(t/\varepsilon^{2+2\alpha}, x/\varepsilon^{2+\alpha}, k)$ 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 (7.37) can be derived in a quick formal way. We represent the solution of (7.34) as $\phi^\varepsilon(t,x,k) = \psi^\varepsilon(\varepsilon^2 t, \varepsilon^2 x, k)$ and write an asymptotic multiple scale expansion for ψ^ε

$$\psi^\varepsilon(t,x,k) = \bar{\phi}(t,x,k) + \varepsilon \phi_1\left(t,x,\frac{x}{\varepsilon^2},k\right) + \varepsilon^2 \phi_2\left(t,x,\frac{x}{\varepsilon^2},k\right) + \dots \tag{7.46}$$

We assume formally that the leading order term $\bar{\phi}$ is deterministic and independent of the fast variable $z = x/\varepsilon^2$. We insert this expansion into (7.34) and obtain in the order $O(\varepsilon^{-1})$:

$$\nabla V(z) \cdot \nabla_k \bar{\phi} - k \cdot \nabla_z \phi_1 = 0. \tag{7.47}$$

Let $\theta \ll 1$ be a small positive regularization parameter that will be later sent to zero, and consider a regularized version of (7.47):

$$\frac{1}{|k|} \nabla V(z) \cdot \nabla_k \bar{\phi} - \hat{k} \cdot \nabla_z \phi_1 + \theta \phi_1 = 0,$$

Its solution is

$$\phi_1(z,k) = -\frac{1}{|k|} \int_0^\infty \sum_{m=1}^d \frac{\partial V(z + s\hat{k})}{\partial z_m} \frac{\partial \bar{\phi}(t,x,k)}{\partial k_m} e^{-\theta s} ds, \tag{7.48}$$

and the role of $\theta > 0$ is to ensure that the integral in the right side converges. The next order equation becomes upon averaging

$$\frac{\partial \bar{\phi}}{\partial t} + k \cdot \nabla_x \bar{\phi} = \langle \nabla V(z) \cdot \nabla_k \phi_1 \rangle. \tag{7.49}$$

The term in the right side above may be written using expression (7.48) for ϕ_1:

$$\langle \nabla V(z) \cdot \nabla_k \phi_1 \rangle = \Big\langle \sum_{m,n=1}^d \frac{\partial V(z)}{\partial z_m} \frac{\partial}{\partial k_m} \Big(\frac{1}{|k|} \int_0^\infty \frac{\partial V(z + s\hat{k})}{\partial z_n} \frac{\partial \bar{\phi}(t,x,k)}{\partial k_n} e^{-\theta s} ds \Big) \Big\rangle.$$

Using spatial stationarity of $H_1(z,k)$ we may rewrite the above as

$$
\begin{aligned}
&-\Big\langle \sum_{m,n=1}^{d} V(z)\frac{\partial}{\partial z_m}\frac{\partial}{\partial k_m}\Big(\frac{1}{|k|}\int_0^\infty \frac{\partial V(z+s\hat{k})}{\partial z_n}\frac{\partial\bar{\phi}(t,x,k)}{\partial k_n}\,e^{-\theta s}ds\Big)\Big\rangle \\
&= -\sum_{m,n=1}^{d}\frac{\partial}{\partial k_m}\Big(\frac{1}{|k|}\int_0^\infty \Big\langle V(z,k)\frac{\partial^2 V(z+s\hat{k})}{\partial z_n\partial z_m}\Big\rangle\frac{\partial\bar{\phi}(t,x,k)}{\partial k_n}\,e^{-\theta s}ds\Big) \\
&= -\sum_{m,n=1}^{d}\frac{\partial}{\partial k_m}\Big(\frac{1}{|k|}\int_0^\infty \frac{\partial^2 R(s\hat{k})}{\partial x_n\partial x_m}\frac{\partial\bar{\phi}(t,x,k)}{\partial k_n}\,e^{-\theta s}ds\Big) \\
&\to -\frac{1}{2}\sum_{m,n=1}^{d}\frac{\partial}{\partial k_m}\left(\frac{1}{|k|}\int_{-\infty}^\infty \frac{\partial^2 R(s\hat{k})}{\partial x_n\partial x_m}\frac{\partial\bar{\phi}(t,x,k)}{\partial k_n}ds\right),\ \text{as } \theta\to 0^+.
\end{aligned}
$$

We insert the above expression into (7.49) and obtain

$$
\frac{\partial\bar{\phi}}{\partial t} = \sum_{m,n=1}^{d}\frac{\partial}{\partial k_n}\left(D_{nm}(k)\frac{\partial\bar{\phi}}{\partial k_m}\right) + k\cdot\nabla_x\bar{\phi} \tag{7.50}
$$

with the diffusion matrix $D(\hat{k},k)$ as in (7.35). Observe that (7.50) is nothing but (7.37). However, the naive asymptotic expansion (7.46) may not be justified directly, to the best of my knowledge. The rigorous proof is based on a completely different method.

8 The Schrödinger equation: slowly varying initial conditions

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

$$
i\phi_t + \frac{1}{2}\Delta\phi - \varepsilon V(t,x)\phi = 0, \tag{8.1}
$$

with initial conditions of the form $\phi(0,x) = \phi_0(\delta x)$, with $\delta \ll 1$. In other words, the initial condition varies on a scale much larger than the correlation length of the random potential. A common choice is to take $\delta = \varepsilon$ so that the diffusively rescaled solution $\phi^\varepsilon(t,x) = \phi(t/\varepsilon^2, x/\varepsilon)$ satisfies

$$
i\phi^\varepsilon_t + \frac{1}{2}\Delta\phi^\varepsilon - \frac{1}{\varepsilon}V(\frac{t}{\varepsilon^2},\frac{x}{\varepsilon})\phi^\varepsilon = 0,\quad \phi^\varepsilon(0,x) = \phi_0(x). \tag{8.2}
$$

However, this is not a unique possible choice – it is interesting to consider any $\delta \ll 1$. Recall that $\delta = 1$ corresponds to the kinetic regime that we are avoiding in these notes. For convenience, we will set $\delta = \varepsilon^\alpha$ with some $\alpha > 0$. Our goal is to analyze the long time behavior of $\phi(t,x)$, and understand the energy transfer from the low to high frequencies that comes about from the inhomogeneities in the random media.

We assume that $V(t,x)$ is a stationary mean-zero Gaussian random field with a spectral representation

$$
V(t,x) = \int_{\mathbb{R}^d} e^{ip\cdot x}\frac{\tilde{V}(t,dp)}{(2\pi)^d}. \tag{8.3}
$$

Here $\tilde{V}(t, dp)$ is the stochastic measure and $\tilde{V}(t, dp) = \tilde{V}^*(t, -dp)$, so V is real-valued. Its covariance function and power spectrum are

$$R(t,x) = \mathbb{E}\{V(s,y)V(s+t,y+x)\}, \quad \hat{R}(\omega,\xi) = \int_{\mathbb{R}^{d+1}} R(t,x)e^{-i\omega t - i\xi\cdot x}dtdx.$$

We assume that the spatial power spectrum (the Fourier transform of $R(t,x)$ in x only) has the form

$$\tilde{R}(t,\xi) = \int_{\mathbb{R}^d} R(t,x)e^{-i\xi\cdot x}dx = e^{-\mathfrak{g}(\xi)|t|}\hat{R}(\xi), \tag{8.4}$$

where $\hat{R}(\xi) \in L^1(\mathbb{R}^d)$, and the spectral gap $\mathfrak{g}(\xi) \geq 0$, so that

$$\hat{R}(\omega,\xi) = \frac{2\mathfrak{g}(\xi)\hat{R}(\xi)}{\omega^2 + \mathfrak{g}^2(\xi)}. \tag{8.5}$$

By Bochner's theorem, we have $\hat{R}, \tilde{R} \geq 0$, and we assume that

$$\frac{\hat{R}(p)}{\mathfrak{g}(p)} \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d). \tag{8.6}$$

The compensated wave function

The standard approach to an understanding of the behavior of the solutions of the weakly random Schrödinger equation is in the context of the kinetic limit [6, 7, 3, 53, 19, 46, 4], through the study of the Wigner transform of the solution (the phase space resolved energy density) [30]. We will proceed here as in [5, 45], focusing not on the weak limit of the energy density of the solution but on the strong limit of the wave field itself. In order to motivate the "correct" way to this end, let us mention that after a long time the phase of the wave field acquires a large factor: for instance, setting $V = 0$ leads to an explicit expression

$$\hat{\phi}(t,\xi) = e^{-i|\xi|^2 t/2}\hat{\phi}(0,\xi)$$

for the Fourier transform of the solution. Thus, a convenient object in the context of long time behaviors is the compensated wave function

$$\hat{\psi}(t,\xi) = e^{i|\xi|^2 t/2}\hat{\phi}(t,\xi), \tag{8.7}$$

which eliminates the deterministic component of the phase. This procedure is also known as phase conjugation in the engineering and physical literature. The surprising miracle is that after this simple-minded phase compensation, the wave field has a non-trivial limit. This is surprisingly similar to what we have done at the very beginning for a particle in a strong mean flow: there, we have subtracted the mean displacement. Here, we do the same – factor out the background dynamics.

Loose end #1: the high frequency initial data

In order to put things into some context, we first describe the results of [5] obtained when the initial condition is not slowly varying:

$$\phi(0,x) = \phi_0(x).$$

Let us set

$$D(p,\xi) = \frac{2\hat{R}(p)}{(2\pi)^d[\mathfrak{g}(p) - i(|\xi|^2 - |\xi - p|^2)/2]}, \qquad D(\xi) = \int_{\mathbb{R}^d} D(p,\xi)dp. \tag{8.8}$$

It is straightforward to check that

$$\mathrm{Re}D(p,\xi) = \frac{2\hat{R}(p)\mathfrak{g}(p)}{(2\pi)^d[\mathfrak{g}^2(p) + (|\xi|^2 - |\xi - p|^2)^2/4]} = \frac{1}{(2\pi)^d}\hat{R}(\frac{|\xi|^2 - |\xi - p|^2}{2}, p). \tag{8.9}$$

One of the results of [5] is that if

$$\frac{\hat{R}(p)}{\mathfrak{g}(p)} \in L^1(\mathbb{R}^d),$$

then on the time scale $t \sim \varepsilon^{-2}$, the compensated wave function corresponding to the initial data with $\alpha = 0$ converges pointwise in distribution to a Gaussian random variable:

$$\hat{\phi}(\frac{t}{\varepsilon^2}, \xi)e^{\frac{i|\xi|^2 t}{2\varepsilon^2}} \Rightarrow \hat{\phi}_0(\xi)e^{-\frac{1}{2}D(\xi)t} + Z(t,\xi). \tag{8.10}$$

Here, $Z(t,\xi)$ is a centered, complex valued Gaussian with the variance

$$\mathbb{E}\{|Z(t,\xi)|^2\} = \widehat{W}(t,\xi) - |\hat{\phi}_0(\xi)|^2 e^{-\mathrm{Re}D(\xi)t}. \tag{8.11}$$

The function $\widehat{W}$ solves a (space-homogeneous) kinetic equation

$$\partial_t \widehat{W} = \int_{\mathbb{R}^d} \hat{R}(\frac{|p|^2 - |\xi|^2}{2}, p - \xi)(\widehat{W}(t,p) - \widehat{W}(t,\xi))\frac{dp}{(2\pi)^d}, \tag{8.12}$$

with the initial condition

$$\widehat{W}(0,\xi) = |\hat{\phi}_0(\xi)|^2.$$

This result is consistent with the aforementioned "traditional" kinetic equation approaches.

Loose end #2: homogenization of the very low frequencies

The results in the high frequency regime ($\alpha = 0$) should be contrasted with the analysis of Bal and Zhang in [57, 58] for the case $\alpha = 1$, performed for time-independent potentials. For the initial value problem

$$\begin{aligned} &i\phi_t + \frac{1}{2}\Delta\phi - \varepsilon V(x)\phi = 0, \\ &\phi(0,x) = \phi_0(\varepsilon x), \end{aligned} \tag{8.13}$$

with a mean-zero Gaussian random potential $V(x)$, they have established a homogenization result:

$$\phi^\varepsilon(t,x) := \phi\Big(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\Big)$$

converges in probability, as $\varepsilon \to 0$ to a deterministic limit $\bar\phi(t,x)$, which satisfies the Schrödinger equation

$$\begin{aligned} &i\bar\phi_t + \frac{1}{2}\Delta\bar\phi - \bar V\bar\phi = 0, \\ &\bar\phi(0,x) = \phi_0(x). \end{aligned} \tag{8.14}$$

The effective potential is constant and is given by

$$\bar V = \int_{\mathbb{R}^d} \frac{\hat R(p)dp}{|p|^2}.$$

Let us mention that the choice $\alpha = 1$ is special, as then the overall phase of the solution at the times $t \sim \varepsilon^{-2}$ is

$$\frac{t}{\varepsilon^2}\varepsilon^2|\xi|^2 = O(1),$$

so that no phase compensation is needed.

Homogenization of the low frequencies

Summarizing the above results, while solutions with the high frequency initial data have a random limit on the time scale $t \sim \varepsilon^{-2}$, as in (8.10), solutions with the "very slowly varying" initial data as in (8.13) are homogenized on this time scale – their limit is deterministic. Our first goal is to understand where the transition between the two regimes occurs – this is the motivation for introducing a general $\alpha > 0$. It will turn out that the homogenization result (formulated for the compensated wave function) holds for all $\alpha > 0$ – that is, no matter how "relatively high" the low frequency of the initial condition is, solution has a deterministic limit at times $t \sim \varepsilon^{-2}$. However, we will see that, unlike in the setting of [57, 58], the temporal fluctuations of the random potential lead to an effective potential with a non-trivial imaginary part. This means that the homogenized field loses mass in the limit. This loss of mass is attributed to the energy transfer to the high frequencies, which, as we show, account for the mass missing in the low frequencies, do not homogenize, and satisfy a kinetic type limit. We also analyze the random fluctuations of the low frequency component of the wave field and characterize the corrector to the homogenized limit.

More precisely, we consider the Schrödinger equation

$$i\partial_t\phi(t,x) + \frac{1}{2}\Delta\phi(t,x) - \varepsilon V(t,x)\phi(t,x) = 0 \tag{8.15}$$

with a low frequency initial condition

$$\phi(0,x) = \phi_0(\kappa x), \tag{8.16}$$

with $\kappa \ll 1$. The Fourier transform of the initial condition is

$$\hat{\phi}(0,\xi) = \kappa^{-d}\hat{\phi}_0(\frac{\xi}{\kappa}).$$

Thus, if the function $\hat{\phi}_0(\xi)$ is of the Schwartz class, $\hat{\phi}(0,\xi)$ is concentrated on the wave vectors ξ of the size $O(\kappa)$. While the Schrödinger equation with a time-dependent potential conserves the total mass:

$$M(t) = \int_{\mathbb{R}^d} |\phi(t,x)|^2 dx = \int_{\mathbb{R}^d} |\phi(0,x)|^2 dx, \tag{8.17}$$

the total energy

$$E(t) = \int_{\mathbb{R}^d} [|\nabla\phi|^2 + \varepsilon V|\phi|^2] dx \tag{8.18}$$

is not conserved, unlike for time-independent potentials. Thus, even if the mass is initially concentrated in the low wave numbers, after a long time evolution it may spread to $O(1)$ frequencies as well. As the potential is weak, the time it takes for the mass to spread over a range of frequencies will be long.

We consider the long time behavior of the solution, on the time scale of the order $t \sim \varepsilon^{-2}$, when the effect of the weak random potential will be non-trivial. We will first consider the "low frequency" rescaled compensated wave function:

$$\psi_\varepsilon(t,\xi) = \kappa^d \hat{\phi}(\frac{t}{\varepsilon^2}, \kappa\xi) e^{\frac{i\kappa^2|\xi|^2 t}{2\varepsilon^2}} \tag{8.19}$$

with the initial data $\psi_\varepsilon(0,\xi) = \hat{\phi}_0(\xi)$. This allows us to study the low frequency component of the solution – wave numbers of the order $O(\kappa)$. A straightforward computation shows that this function is a solution of the following integral equation

$$\psi_\varepsilon(t,\xi) = \hat{\phi}_0(\xi) + \frac{1}{i\varepsilon}\int_0^t \int_{\mathbb{R}^d} \frac{\tilde{V}(\frac{s}{\varepsilon^2}, dp)}{(2\pi)^d} e^{i\kappa^2(|\xi|^2 - |\xi - \frac{p}{\kappa}|^2)\frac{s}{2\varepsilon^2}} \psi_\varepsilon(s, \xi - \frac{p}{\kappa}) ds. \tag{8.20}$$

We have the following result for the low frequencies.

Theorem 8.1 *Assume that $\kappa = \varepsilon^\alpha$ with $\alpha > 0$. Then, for fixed $t > 0$ and $\xi \in \mathbb{R}^d$,*

$$\psi_\varepsilon(t,\xi) \to \bar{\psi}(t,\xi) = \hat{\phi}_0(\xi) e^{-\frac{1}{2}D(0)t} \text{ in probability as } \varepsilon \to 0. \tag{8.21}$$

Let us stress that $\xi = O(1)$ in the argument of the function $\psi_\varepsilon(t,\xi)$ corresponds to $\xi = O(\kappa)$ in the argument of the function ϕ – Theorem 8.1 addresses the evolution of the low frequencies of the solution of the Schrödinger equation with a slowly varying initial condition. Recall that

$$D(0) = \int_{\mathbb{R}^d} \frac{2\hat{R}(p)}{(2\pi)^d(\mathfrak{g}(p) + i|p|^2/2)} dp, \tag{8.22}$$

and, as $\mathfrak{g}(p) \geq 0$, we have $\mathrm{Re}D(0) > 0$. Therefore, the passage to limit $\varepsilon \to 0$ in (8.21) induces a loss of the $L^2(\mathbb{R}^d)$ norm: while

$$\|\psi_\varepsilon(t,\cdot)\|_{L^2} = \|\phi_0\|_{L^2},$$

as can be seen simply from the definition of $\psi_\varepsilon(t,\xi)$, we have

$$\|\bar{\psi}(t,\cdot)\|_{L^2} = \|\phi_0\|_{L^2} e^{-\mathrm{Re}D(0)t/2} < \|\phi_0\|_{L^2}.$$

The natural question is how does the loss of mass happen, and where does the mass go? Mathematically, there is no contradiction, as the convergence in Theorem 8.1 is not uniform with respect to $\xi \in \mathbb{R}^d$. From a physical point of view, as we have mentioned, the time dependence of the random potential breaks the conservation of the energy (8.18), which allows the mass to escape to the high frequencies. Let us mention that in the *time-independent* case [11], where the conservation of the energy prevents the escape of mass from the low frequencies, it is shown that the mass is conserved as well.

Generation of the high frequencies

We now investigate the generation of the high frequencies in the above setting. Once again, we consider the solution $\phi(t,x)$ of (8.15) with the initial data (8.16). We stress that in all our results the initial condition (8.16) is the same – various rescalings in Theorem 8.1 above and Theorems 8.2, 8.3 and 8.4 below correspond to zooming into various frequency ranges in the same solution. Our next goal is to understand how the mass escapes from the low frequencies (those of the initial condition) to the high frequencies, generated by the interaction with the random potential. As we are now interested in the high and not the low frequencies, we define the compensated wave function not quite as in (8.19), but as

$$\Psi_\varepsilon(t,\xi) = \kappa^{\frac{d}{2}} \hat{\phi}(\frac{t}{\varepsilon^2},\xi) e^{\frac{i|\xi|^2 t}{2\varepsilon^2}}, \tag{8.23}$$

so that the frequency is not rescaled. The initial condition for Ψ_ε is

$$\Psi_\varepsilon(0,\xi) = \kappa^{-d/2} \hat{\phi}_0(\xi/\kappa).$$

The pre-factor $\kappa^{d/2}$ in (8.23) is chosen so that we get a non-trivial limit. This function solves the integral equation

$$\Psi_\varepsilon(t,\xi) = \frac{1}{\kappa^{d/2}} \hat{\phi}_0(\frac{\xi}{\kappa}) + \frac{1}{i\varepsilon} \int_0^t \int_{\mathbb{R}^d} \frac{\tilde{V}(\frac{s}{\varepsilon^2}, dp)}{(2\pi)^d} e^{i(|\xi|^2 - |\xi-p|^2)\frac{s}{2\varepsilon^2}} \Psi_\varepsilon(s,\xi-p) ds. \tag{8.24}$$

The following result explains the loss of mass observed in Theorem 8.1, and tracks the generation of the high frequencies.

Theorem 8.2 *Assume that $\kappa = \varepsilon^\alpha$ with $\alpha > 0$, then for fixed $t > 0$ and $\xi \neq 0$, we have*

$$\Psi_\varepsilon(t,\xi) \Rightarrow \bar{Z}(t,\xi) \text{ in law as } \varepsilon \to 0,$$

where $\bar{Z}(t,\xi)$ is a centered, complex valued Gaussian random variable. Its variance $\widehat{W}_\delta(t,\xi)$ is the solution of (8.12) with the initial condition $\widehat{W}_\delta(0,\xi) = \|\hat{\phi}_0\|_2^2 \delta(\xi)$.

The variance $\widehat{W}_\delta(t,\xi)$ can be explicitly written as a series expansion

$$\widehat{W}_\delta(t,\xi) = \widehat{W}_{\delta,b}(t,\xi) + \widehat{W}_{\delta,s}(t,\xi), \tag{8.25}$$

with the ballistic part

$$\widehat{W}_{\delta,b}(t,\xi) = \|\hat{\phi}_0\|^2 e^{-\mathrm{Re}D(0)t}\delta(\xi),$$

and the scattering part

$$\widehat{W}_{\delta,s}(t,\xi) = \sum_{k=1}^{\infty} \|\hat{\phi}_0\|_2^2 \int_{0=v_{k+1}\le v_k\le\ldots\le v_1\le v_0=t} dv \int_{\mathbb{R}^{kd}} dP \prod_{j=0}^{k} e^{-(v_j-v_{j+1})\mathrm{Re}D(\xi-\ldots-P_j)} \times \prod_{j=1}^{k} \mathrm{Re}D(P_j, \xi-\ldots-P_{j-1})\delta(\xi-P_1-\ldots-P_k).$$

Let us mention that $\widehat{W}_{\delta}(t,\xi) = \widehat{W}_{\delta,s}(t,\xi)$ when $\xi \neq 0$, that is, only the scattering part contributes to the variance in Theorem 8.2. We also observe

$$\int_{\mathbb{R}^d} \widehat{W}_{\delta,b}(t,\xi)d\xi = \|\hat{\phi}_0\|_2^2 e^{-\mathrm{Re}D(0)t},$$

which equals to the mass of the low frequency waves.

Theorems 8.1 and 8.2 describe the dynamics of on different scales of the frequency domain. In the former case, the low frequencies are zoomed in, and we find a deterministic evolution (homogenzation). In the latter, we track the high frequency component of the solution, so that the low frequency initial condition shrinks to a point source at the origin, which generates the high frequency waves.

The fluctuation analysis in homogenization regime

We now return to the analysis of the behavior of the low frequencies. According to Theorem 8.1, the compensated wave function homogenizes for the low frequencies, hence the next interesting object is the fluctuation, which we define as

$$\mathcal{U}_{\varepsilon}(t,\xi) = \frac{1}{\kappa^{d/2}}(\psi_{\varepsilon}(t,\xi) - \mathbb{E}\{\psi_{\varepsilon}(t,\xi)\}).$$

Here, $\psi_{\varepsilon}(t,\xi)$ is defined as in (8.19). Heuristically, since the homogenization limit in Theorem 8.1 captures the ballistic component of the wave field, we expect small random fluctuations consisting of the remaining scattering components. Indeed, we will see that the fluctuation exhibits a kinetic-like behavior. Let us set

$$\mathcal{W}_{\alpha}(t,\xi) = \begin{cases} 0 & \text{if } \alpha \in (0,1), \\ -D(0,0)e^{-D(0)t}\displaystyle\int_0^t \int_{\mathbb{R}^d} \hat{\phi}_0(\xi-p)\hat{\phi}_0(\xi+p)e^{-i|p|^2 v}dpdv & \text{if } \alpha = 1, \\ -D(0,0)te^{-D(0)t}\displaystyle\int_{\mathbb{R}^d} \hat{\phi}_0(\xi-p)\hat{\phi}_0(\xi+p)dp & \text{if } \alpha > 1. \end{cases} \tag{8.26}$$

Theorem 8.3 *Assume that $\kappa = \varepsilon^{\alpha}$, then for fixed $t > 0$ and $\xi \in \mathbb{R}^d$, we have*

$$\mathcal{U}_{\varepsilon}(t,\xi) \Rightarrow Z_{\delta}(t,\xi) = X_{\delta}(t,\xi) + iY_{\delta}(t,\xi) \text{ as } \varepsilon \to 0,$$

where X_δ, Y_δ are centered, jointly Gaussian random variables such that

$$\mathbb{E}\{|Z_\delta(t,\xi)|^2\} = \widehat{W}_{\delta,s}(t,0),$$

and

$$\mathbb{E}\{Z_\delta(t,\xi)^2\} = \mathcal{W}_\alpha(t,\xi).$$

Therefore, we can write

$$\psi_\varepsilon(t,\xi) = \mathbb{E}\{\psi_\varepsilon(t,\xi)\} + \kappa^{d/2}\mathcal{U}_\varepsilon(t,\xi),$$

and Theorem 8.3 shows that when $\kappa = \varepsilon^\alpha$, with $\alpha < 1$, the fluctuation $\mathcal{U}_\varepsilon(t,\xi)$ is approximately distributed as $Z_\delta(t,0)$, a centered complex Gaussian random variable with variance $\widehat{W}_{\delta,s}(t,0)$. This is similar to the result of Theorem 8.2 for the high frequency, albeit the variance is now given by the transport solution evaluated at the origin $\xi = 0$, since we are now in the low frequency regime. If we let $\alpha \to 0$ (which is the same as $\kappa \to 1$, so that the initial condition is less and less slowly varying), then, formally, $\psi_\varepsilon(t,\xi)$ is distributed as

$$\hat{\phi}_0(\xi)e^{-\frac{1}{2}D(0)t} + Z_\delta(t,0),$$

which is consistent with (8.10). That is, Theorem 8.3 also interpolates between the deterministic limit for the low frequencies and the random behavior of the high frequency component of the solution.

The Wigner transform of the random fluctuation

Besides the pointwise fluctuation for a fixed $\xi \in \mathbb{R}^d$, we also consider the fluctuation of $\psi_\varepsilon(t,\xi)$ as a wave field. The tool we use is the Wigner transform for some $\beta \geq 0$:

$$W_\varepsilon(t,x,\xi) = \int_{\mathbb{R}^d} \mathcal{U}_\varepsilon(t,\xi + \frac{\varepsilon^\beta \eta}{2})\mathcal{U}_\varepsilon^*(t,\xi - \frac{\varepsilon^\beta \eta}{2})e^{i\eta\cdot x}\frac{d\eta}{(2\pi)^d}. \tag{8.27}$$

Let $\bar{W}_\delta$ be the solution to the kinetic equation

$$\partial_t \bar{W} + \xi\cdot\nabla_x \bar{W} = \int_{\mathbb{R}^d} \hat{R}(\frac{|p|^2-|\xi|^2}{2}, p-\xi)(\bar{W}(t,x,p) - \bar{W}(t,x,\xi))\frac{dp}{(2\pi)^d}, \tag{8.28}$$

with the initial condition

$$\bar{W}_\delta(0,x,\xi) = \|\hat{\phi}_0\|_2^2\delta(\xi)\delta(x),$$

and $\bar{W}_{\delta,b}, \bar{W}_{\delta,s}$ be the ballistic and scattering component of $\bar{W}_\delta$, respectively:

$$\bar{W}_{\delta,b}(t,x,\xi) = \|\hat{\phi}_0\|_2^2\delta(\xi)\delta(x)e^{-\mathrm{Re}D(0)t},$$

and

$$\bar{W}_{\delta,s}(t,x,\xi) = \sum_{k=1}^\infty \|\hat{\phi}_0\|_2^2 \int_{0=v_{k+1}\leq v_k\leq\ldots\leq v_1\leq v_0=t} dv \int_{\mathbb{R}^{kd}} dP \prod_{j=0}^k e^{-(v_j-v_{j+1})\mathrm{Re}D(\xi-\ldots-P_j)}$$
$$\times \prod_{j=1}^k \mathrm{Re}D(P_j, \xi-\ldots-P_{j-1})\delta(\xi-P_1-\ldots-P_k)\delta(x-\xi t+\sum_{j=1}^k P_j v_j).$$

Theorem 8.4 *Assume that $\kappa = \varepsilon^\alpha$, $\alpha \in (0,1)$ and $\alpha + \beta = 2$, then for any test function $\varphi \in \mathcal{S}(\mathbb{R}^{2d})$ and $t > 0$,*

$$\int_{\mathbb{R}^{2d}} W_\varepsilon(t,x,\xi)\varphi^*(x,\xi)dxd\xi \to \int_{\mathbb{R}^{2d}} \bar{W}_{\delta,s}(t,x,0)\varphi^*(x,\xi)dxd\xi$$

in probability as $\varepsilon \to 0$.

As Theorem 8.1 indicates that the ballistic component of transport solution gives the low frequency behavior, we conclude from Theorems 8.3 and 8.4 that the small random fluctuations are described by the scattering component of the solution of the kinetic equation.

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] G. Bal, T. Komorowski, and L. Ryzhik, Kinetic limits for waves in random media, Kin. Rel. Mod., 3(4) (2010), pp. 529 – 644.

[5] G. Bal, T. Komorowski, and L. Ryzhik, Asymptotics of the phase of the solutions of the random Schrödinger equation, Arch. Rat. Mech. Anal., 200(2) (2011), pp. 613–664.

[6] G. Bal, G. C. Papanicolaou, and L. Ryzhik, Radiative transport limit for the random Schrödinger equation, Nonlinearity, 15 (2002), pp. 513–529.

[7] G. Bal, G. C. Papanicolaou, and L. Ryzhik, Self-averaging in time reversal for the parabolic wave equation, Stoch. Dyn., 4 (2002), pp. 507–531.

[8] 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.

[9] H. Berestycki, A. Kiselev, A. Novikov and L. Ryzhik, The explosion problem in a flow, Jour. d'Anal. Mathématique, **110**, 2010, 31–65.

[10] P. Billingsley, Convergence of Probability Measures, Wiley, 1999.

[11] T. Chen, T. Komorowski and L. Ryzhik, The weak coupling limit for the random Schrdinger equation: The average wave function, Arch. Ration. Mech. Anal. **227**, 2018, 387–422..

[12] P. Constantin, A. Kiselev and L. Ryzhik, Quenching of flames by fluid advection, Comm. Pure Appl. Math., **54**, 2001, 1320–1342.

[13] P. Constantin, A. Kiselev, L. Ryzhik and A. Zlatos, Diffusion and mixing in fluid flow. Ann. of Math., **168**, 2008, 643–674.

[14] I.P. Cornfeld, S.V. Fomin and Ya.G. Sinai, *Ergodic Theory,* Springer-Verlag, New York, 1982.

[15] H. Cycon, R. Froese, W. Kirsch and B. Simon, *Schrödinger Operators,* Springer-Verlag, 1987.

[16] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, Vol. 3: Spectral Theory and Applications, Springer, 2000.

[17] 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.

[18] V. Enss, Asymptotic completeness for quantum mechanical potential scattering. I. Short range potentials. Comm. Math. Phys. **61**, 1978, 285–291.

[19] L. Erdös and H. T. Yau, Linear Boltzmann equation as the weak coupling limit of a random Schrödinger equation, Comm. Pure Appl. Math., 53(6) (2000), pp. 667–735.

[20] L.C. Evans, Partial Differential Equations, Second Edition, AMS, 2010.

[21] A. Fannjiang and L. Wo?owski, Noise induced dissipation in Lebesgue-measure preserving maps on d-dimensional torus. J. Statist. Phys. **113**, 2003, 335–378.

[22] A. Fannjiang, S. Nonnenmacher, L. Woĺowski, Relaxation time of quantized toral maps. Ann. Henri Poincar **7**, 2006, 161–198.

[23] A. Fannjiang, S. Nonnenmacher, L. Woĺowski, Dissipation time and decay of correlations. Nonlinearity **17**, 2004, 1481–1508.

[24] B. Fayad, *Weak mixing for reparameterized linear flows on the torus,* Ergodic Theory Dynam. Systems **22**, 2002, 187–201.

[25] B. Fayad, Fayad, Bassam Smooth mixing flows with purely singular spectra. Duke Math. J. **132**, 2006, 371–391.

[26] 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.

[27] 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.

[28] 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.

[29] P. Gérard, Microlocal defect measures, Comm. PDEs, **16**, 1991, 1761-1794.

[30] P.Gérard, P.Markovich, N.Mauser and F.Poupaud, Homogenization limits and Wigner transforms, Comm.Pure Appl. Math., **50**, 1997, 323-380.

[31] Ichihara, K. snd Kunita, H., *A classification of the second order degenerate elliptic operators and its probabilistic characterization*, Z. Wahrscheinlichkeitstheorie verw., **30**, 1974, 235-254.

[32] Hörmander, L., Hypoelliptic second order differential equations, Acta Math., **119**, 1967, 147-171.

[33] 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.

[34] Y. Katznelson, *An Introduction to Harmonic Analysis*, Cambridge University Press, 2004.

[35] J.B. Keller, Stochastic equations and wave propagation in random media, Proc. Sympos. Appl. Math. AMS, Vol. XVI, 1964, 145–170.

[36] H. Kesten and G. Papanicolaou, A limit theorem for turbulent diffusion, Comm. Math. Phys., **65**, 1979, 97-128.

[37] H. Kesten, G. C. Papanicolaou, A Limit Theorem for Stochastic Acceleration, Comm. Math. Phys., **78**, 1980, 19–63.

[38] R. Khasminskii, A limit theorem for solutions of differential equations with a random right hand side, Theory Probab. Appl. **11**, 390-406.

[39] Y. Kifer, *Random Perturbations of Dynamical Systems*, Birkhäuser, Boston, 1988.

[40] A. Kiselev and A. Zlatos, Quenching of combustion by shear flows, Duke Math. J. **132**, 2006, 49–72.

[41] A.N. Kolmogorov, *On dynamical systems with an integral invariant on the torus (Russian),* Doklady Akad. Nauk SSSR (N.S.), **93**, 1953, 763–766.

[42] T. Komorowski, Diffusion approximation for the advection of particles in a a strongly turbulent random environment, Ann. Probab., **24**, 1996, 346–376.

[43] T. Komorowski and L. Ryzhik, Diffusion in a weakly random Hamiltonian flow, Comm. Math. Phys., **262**, 2006, 277–323.

[44] T. Komorowski and L. Ryzhik, The stochastic acceleration problem in two dimensions, Israel Jour. Math., **155**, 2006, 157–204.

[45] T. Komorowski, S. Olla, and L. Ryzhik, Asymptotics of the solutions of the stochastic lattice wave equation, Arch. Rat. Mech. Anal., 209 (2013), pp. 455–494.

[46] J. Lukkarinen and H. Spohn, Kinetic limit for wave propagation in a random medium, Arch. Rat. Mech. Anal., 183(1) (2007), pp. 93–162.

[47] J. Moser, *On the volume elements on a manifold,* Trans. Amer. Math. Soc., **120**, 1965, 286–294.

[48] J. von Neumann, *Zur operatorenmethode in der classichen mechanik,* Annals of Math (2), **33**, 1932, 587–642.

[49] J.R. Norris, Long-time behaviour of heat flow: global estimates and exact asymptotics. Arch. Rational Mech. Anal. **140**, 1997, 161–195.

[50] G. Papanicolaou and W. Kohler, Asymptotic theory of mixing stochastic ordinary differential equations, Comm. Pure Appl. Math., **27**, 1974, 641-668.

[51] D. Ruelle, A remark on bound states in potential-scattering theory. Nuovo Cimento A (10) **61**, 1969, 655–662.

[52] M.D. Šklover, *Classical dynamical systems on the torus with continuous spectrum,* Izv. Vuzov, **10**, 1967, 113–124.

[53] H. Spohn, Derivation of the transport equation for electrons moving through random impurities, J. Stat. Phys., 17 (1977(6)), pp. 385–412.

[54] S. Sternberg, *On differential equations on the torus,* Amer. J. Math., **79**, 1957, 397–402.

[55] B.S. White, The stochastic caustic, SIAM J. Appl. Math. **44**, 1984, 127–149.

[56] E. Wigner, On the quantum correction for thermodynamic equilibrium, Physical Rev., **40**, 1932, 749-759.

[57] N. Zhang and G. Bal, Convergence to spde of the schrödinger equation with large, random potential, Comm. Math. Sci., 12 (2014).

[58] N. Zhang and G. Bal, Homogenization of the schroedinger equation with large, random potential, Stoch. Dyn., 14 (2014).