

Lecture Notes on Kinetic Models for Waves in Random Media

Guillaume Bal¹

Tomasz Komorowski²

Lenya Ryzhik³

June 12, 2006

¹Department of Applied Physics and Applied Mathematics, Columbia University, New York NY, 10027; gb2030@columbia.edu

²Institute of Mathematics, UMCS, pl. Marii Curie Skłodowskiej 1, 20-031 Lublin, Poland; e-mail: komorow@hektor.umcs.lublin.pl

³Department of Mathematics, University of Chicago, Chicago IL, 60637; ryzhik@math.uchicago.edu

Chapter 1

Lecture I. Random geometric optics I: short time asymptotics

1.1 Rays without the wave equation – the formal theory

We consider in this section the very basic ray theory in a random emdium without any references to the wave equation – this material is based on the classical paper by J.B. Keller [38].

1.1.1 Perturbative ray theory

Fermat's principle postulates that light goes from point A to B as fast as possible. Such fastest path is called a ray connecting points A and B . The medium in which light is propagating is described in terms of the local speed of light $c(x)$. Let Γ be a ray from A to B , then

$$\int_{\Gamma} \frac{dl}{c(X)} = \inf_{\gamma} \int_{\gamma} \frac{dl}{c(X)}.$$

Here the infimum is taken over all continuous curves γ connecting A and B . Equivalently, parameterizing γ by $x(t) = (x_1(t), x_2(t), x_3(t))$, $0 \leq t \leq 1$, we need to minimize the functional

$$\int_0^1 n(x(s)) |\dot{x}(s)| ds \tag{1.1}$$

with $n(x) = c_0/c(x)$ being the refractive index. Here $c_0 = \text{const}$ is a reference speed that is some typical speed of propagation in the medium. This will be sometimes formalized by requiring that $n(x)$ does not deviate from $n_0 = 1$ too much but that is not required a priori. The Euler-Lagrange equations for the functional (1.1) are

$$\frac{d}{dt} \frac{\partial F}{\partial \dot{x}_j} - \frac{\partial F}{\partial x_j} = 0$$

with $F(x, \dot{x}) = n(x(s)) |\dot{x}(s)|$. This may be rewritten as

$$\frac{d}{dt} \left(\frac{n \dot{x}_j}{|\dot{x}|} \right) - |\dot{x}| \frac{\partial n}{\partial x_j} = 0.$$

Let θ be the unit vector along the ray: $\theta = \dot{x}/|\dot{x}|$, then the above equations take the form

$$\frac{d}{dt} (n\theta) - |\dot{x}| \nabla n = 0. \tag{1.2}$$

A convenient way to parameterize the path is to use the arclength parameter l along the curve $x(t)$, then $dl = |\dot{x}|dt$ and we obtain the ray equation

$$\frac{d}{dl}(n\theta) - \nabla n = 0. \quad (1.3)$$

It may be also written as

$$\frac{d}{dl} \left(n \frac{dx}{dl} \right) - \nabla n = 0. \quad (1.4)$$

Equation (1.4) should be supplemented by the initial conditions:

$$x(0) = 0, \quad \frac{dx}{dl}(0) = \theta_0, \quad |\theta_0| = 1. \quad (1.5)$$

This is the fundamental equation of the ray optics that describes the geometry of rays connecting different points in an inhomogeneous medium. Observe that if $n(x) = \text{const}$, then $d\theta/dl = 0$ and the direction θ doesn't change along the ray. Therefore rays in a homogeneous medium are straight lines. Similarly, if the medium is layered, that is, the refraction variable depends only on the variable x_1 : $n = n(x_1)$ then rays that point initially in the direction of x_1 are straight lines – this also follows immediately from (1.5) with $n = n(x_1)$ and the initial data $\theta_0 = e_1$.

Let us now consider the case when index of refraction deviates slightly from unity:

$$n(x) = 1 + \varepsilon\mu(x).$$

We assume that ε is a small parameter: $\varepsilon \ll 1$ and employ the formal perturbation theory to determine the perturbed path $x(l, \varepsilon)$ expanding it as

$$x(l, \varepsilon) = x_0(l) + \varepsilon x_1(l) + \varepsilon^2 x_2(l) + \dots$$

We insert this expansion in the ray equations (1.4) and get in the order ε^0 : $x_0(l) = l\theta_0$. The first order correction in ε is given by

$$x_1(l) = \int_0^l (l-s) \nabla_{\perp} \mu(\theta_0 s) ds. \quad (1.6)$$

It follows that $(x_1 \cdot x_0) = 0$ – this is typical for a first order correction in the perturbation series. A straightforward computation using the fact that x_1 is perpendicular to θ_0 shows that the second order term x_2 is

$$x_2(l) = \int_0^l ds (l-s) \left[(x_1(s) \cdot \nabla_{\perp}) \nabla_{\perp} \mu(s\theta_0) - \frac{1}{2} \nabla_{\perp} \mu^2(s\theta_0) - \theta_0 \int_0^s (\nabla_{\perp} \mu(s\theta_0) \cdot \nabla_{\perp} \mu(\tau\theta_0)) d\tau \right. \\ \left. - (\theta_0 \cdot \nabla \mu(s\theta_0)) \int_0^s \nabla_{\perp} \mu(\tau\theta_0) d\tau \right]. \quad (1.7)$$

It is useful to observe that both x_1 and x_2 depend only on the perturbation μ (and its gradient) along the line in the direction θ_0 . This is the most transparent example of the role of the unperturbed dynamical system – it moves the particle around, “taking it places” and allowing to sample various regions of fluctuations.

1.1.2 Weakly perturbed rays in a random medium

Expressions for the corrections x_1 and x_2 obtained above are valid for any perturbation $\mu(x)$. Let us now specify that $\mu(x)$ is a random function that has mean zero and its statistics is spatially homogeneous and isotropic:

$$\langle \mu(x) \rangle = 0, \quad \langle \mu(x)\mu(y) \rangle = R(|x - y|), \quad \langle \mu(p)\mu(q) \rangle = (2\pi)^n \hat{R}(p)\delta(p + q). \quad (1.8)$$

The correlation function $R(|x|)$ is smooth, has maximum at zero, is a decreasing and rapidly decaying function of $|x|$, and the power spectrum $\hat{R}(p)$ is its Fourier transform.

The mean ray position

Let us first compute the average ray position using expressions obtained in Section 1.1. The first order correction has mean zero: $\langle x_1 \rangle = 0$ so that

$$\langle x(l) \rangle = l\theta_0 + \varepsilon^2 \langle x_2 \rangle + O(\varepsilon^3).$$

The expected value of x_2 may be computed explicitly using expression (1.7). Observe that $\langle \nabla_\perp \mu^2 \rangle = 0$ since $\langle \mu^2 \rangle$ is constant because of spatial homogeneity. After a straightforward computation we obtain

$$\langle x(l) \rangle = l\theta_0 - \varepsilon^2 \theta_0 \int_0^l ds (l-s) \int_0^s d\tau \langle (\nabla_\perp \mu(s\theta_0) \cdot \nabla_\perp \mu(\tau\theta_0)) \rangle + O(\varepsilon^3).$$

The second term may be evaluated:

$$\langle \nabla_\perp \mu(s\theta_0) \cdot \nabla_\perp \mu(\tau\theta_0) \rangle = -\Delta_\perp R((s - \tau)\theta_0).$$

Here Δ_\perp is the Laplacian in the direction transverse to θ_0 : with our usual choice of θ_0 along the last coordinate x_n we have

$$\Delta_\perp R = \sum_{j=1}^{n-1} \frac{\partial^2 R}{\partial x_j^2}.$$

When $R(|x|)$ is radially symmetric this is

$$\Delta_\perp R(r\theta_0) = \frac{(n-1)R'(r)}{r}.$$

This leads to

$$\langle x(l) \rangle = \theta_0 \left(l + \frac{(n-1)\varepsilon^2}{2} \int_0^l (l-r)^2 \frac{R'(r)}{r} \right) dr + O(\varepsilon^3). \quad (1.9)$$

We see that the mean location of the endpoint of a ray of length l which starts from the origin is in the direction θ_0 . However, its distance from the origin is less than when the ray is not perturbed since $R'(r) < 0$. Physically this is expected because the presence of the inhomogeneities slows down the propagation as light no longer propagates along a straight line.

The mean square fluctuations

Let us denote by ρ deviation of the ray from the straight line in the initial direction θ_0 : $\rho(l) = x_\varepsilon(l) - \theta_0 l$. Then we have

$$\langle \rho(l)^2 \rangle = \varepsilon^2 \langle x_1^2 \rangle + o(\varepsilon^2).$$

We set now space dimension $n = 3$ and compute the above average in the same manner as before:

$$\langle x_1^2(l) \rangle = - \int_0^l \int_0^l (l-s)(l-\tau) \Delta_\perp R(|s-\tau|) d\tau ds = - \int_0^l R'(r) \left[\frac{2r^2}{3} - 2l^2 + \frac{4l^3}{3r} \right] dr.$$

For l large compared to the correlation length a the last term above dominates so that

$$\langle \rho(l)^2 \rangle \approx - \frac{4\varepsilon^2 l^3}{3} \int_0^\infty \frac{R'(r)}{r} dr + O(\varepsilon^3) \text{ for } l \gg a. \quad (1.10)$$

Similarly one may compute the average deviation α of the direction of the ray from the mean direction θ_0 :

$$\begin{aligned} \langle \alpha^2(l) \rangle &= \left\langle \left(\frac{dx}{dl} - \theta_0 \right)^2 \right\rangle = \varepsilon^2 \left\langle \left| \frac{dx_1}{dl} \right|^2 \right\rangle + O(\varepsilon^3) = \varepsilon^2 \left\langle \int_0^l \int_0^l \nabla_\perp \mu(\theta_0 s) \nabla_\perp \mu(\theta_0 \tau) \right\rangle d\tau ds \\ &= -2\varepsilon^2 \int_0^l \int_0^l \frac{R'(|\tau-s|)}{|\tau-s|} d\tau ds = 4\varepsilon^2 \left[R(l) - R(0) - l \int_0^l \frac{R'(r)}{r} dr \right] + O(\varepsilon^3). \end{aligned}$$

For l large compared to the correlation length a this becomes

$$\langle \alpha^2(l) \rangle \approx -4\varepsilon^2 l \int_0^\infty \frac{R'(r)}{r} dr + O(\varepsilon^3) \text{ for } l \gg a. \quad (1.11)$$

Expressions (1.10) and (1.11) may be written as

$$\langle \rho^2(l) \rangle \approx \frac{1}{3} D l^3 + O(\varepsilon^3), \quad \langle \alpha^2(l) \rangle \approx D l + O(\varepsilon^3). \quad (1.12)$$

Here we introduced the ray diffusion coefficient

$$D = -4\varepsilon^2 \int_0^\infty \frac{R'(r)}{r} dr. \quad (1.13)$$

Note that (1.12) may also be obtained by treating the ray direction $\alpha(l)$ as a Brownian motion with the diffusion coefficient D and $\rho(l)$ as its time integral:

$$d\alpha = \sqrt{D} dB_l, \quad d\rho = \alpha(l) dl.$$

Here $B(l)$ is the standard Brownian motion. Then a simple calculation shows that

$$\langle \rho(l)^2 \rangle = D \left\langle \int_0^l \int_0^l B(s) B(s') ds' ds \right\rangle = 2D \left\langle \int_0^l \int_0^{s'} B(s) B(s') ds ds' \right\rangle = 2D \int_0^l \int_0^s s ds ds' = \frac{D l^3}{3}.$$

However, the a priori assumption that ray direction may be described in terms of such Markov process is not easy to justify. Nevertheless, this concept is important and the ray direction does behave as a Markov process in a certain asymptotic limit that we will discuss in the rest of this chapter.

1.1.3 Random Liouville equations: small time formal asymptotics

Reduction to a time-dependent stochastic acceleration problem

In order to make the above discussion of the diffusive ray behavior somewhat more careful (albeit not yet rigorous) we consider the Liouville equations in phase space

$$\frac{\partial \phi}{\partial t} + c(x) \hat{k} \cdot \nabla_x \phi - |k| \nabla c(x) \cdot \nabla_k \phi = 0 \quad (1.14)$$

with the speed $c(x) = 1 + \delta \mu(x)$. Here $\mu(x)$ is a spatially homogeneous random process with the correlation function as in (1.8). Then (1.14) becomes

$$\frac{\partial \phi}{\partial t} + [1 + \delta \mu(x)] \hat{k} \cdot \nabla_x \phi - \delta \nabla \mu(x) |k| \cdot \nabla_k \phi = 0 \quad (1.15)$$

and solutions are close to those of

$$\frac{\partial \phi}{\partial t} + \hat{k} \cdot \nabla_x \phi = 0 \quad (1.16)$$

for times $t = O(1)$. In order to see some more interesting phenomena, in particular, the ray diffusion mentioned above, we look at the bicharacteristics of (1.15):

$$\dot{X}(t) = -(1 + \delta \mu(X)) \hat{K}(t), \quad \dot{K}(t) = \delta \nabla \mu(X(t)) |K(t)|, \quad X(0) = x_0, \quad K(0) = k_0. \quad (1.17)$$

It is convenient to re-write this system in terms of the unit vector $\hat{K}(t) = K(t)/|K(t)|$:

$$\frac{dX(t)}{dt} = -(1 + \delta \mu(X)) \hat{K}(t), \quad \frac{d\hat{K}(t)}{dt} = \delta [\nabla \mu(X(t)) - (\hat{K} \cdot \nabla \mu(X)) \hat{K}]. \quad (1.18)$$

Let us introduce the rescaled quantities $Y(t) = X(t) + \hat{k}_0 t - x_0$ and $P = (\hat{K}(t) - \hat{k}_0)/\delta^\alpha$ with $\alpha > 0$ to be chosen. Naively, one would expect that over a time T the direction \hat{K} deviates from its initial value by δT which means that the trajectory deviates from $X_0(t) = -\hat{k}_0 t$ by $T \cdot \delta T = \delta T^2$. Hence we would expect that $Y(t)$ behaves non-trivially on the time scale $O(\delta^{-1/2})$. We will see, however, that because the random perturbation has mean zero, the effect takes place on a longer time scale. In the slow time variable $t' = \delta^\alpha t$ the above system is

$$\begin{aligned} \frac{d\tilde{Y}(t')}{dt'} &= -\tilde{P}(t') - \delta^{1-\alpha} \mu \left(x_0 - \frac{k_0 t'}{\delta^\alpha} + \tilde{Y}(t') \right) \hat{k}_0 - \delta \mu \left(x_0 - \frac{k_0 t'}{\delta^\alpha} + \tilde{Y}(t') \right) \tilde{P}(t'), \\ \frac{d\tilde{P}(t')}{dt'} &= \delta^{1-2\alpha} \left[I - (\hat{k}_0 + \delta^\alpha \tilde{P}(t')) \otimes (\hat{k}_0 + \delta^\alpha \tilde{P}(t')) \right] \nabla \mu \left(x_0 - \frac{k_0 t'}{\delta^\alpha} + \tilde{Y}(t') \right) \end{aligned} \quad (1.19)$$

with $\tilde{Y}(t') = Y(t'/\delta^\alpha)$ and $\tilde{P}(t') = P(t'/\delta^\alpha)$. We choose $\alpha = 2/3$ so that $\delta^{1-2\alpha} = \delta^{-\alpha/2}$ and introduce $\varepsilon = \delta^{1/3}$:

$$\begin{aligned} \frac{d\tilde{Y}(t')}{dt'} &= -\tilde{P}(t') - \varepsilon \mu \left(x_0 - \frac{k_0 t'}{\varepsilon^2} + \tilde{Y}(t') \right) \hat{k}_0 - \varepsilon^3 \mu \left(x_0 - \frac{k_0 t'}{\varepsilon^2} + \tilde{Y}(t') \right) \tilde{P}(t'), \\ \frac{d\tilde{P}(t')}{dt'} &= \frac{1}{\varepsilon} \left[I - (\hat{k}_0 + \varepsilon^2 \tilde{P}(t')) \otimes (\hat{k}_0 + \varepsilon^2 \tilde{P}(t')) \right] \nabla \mu \left(x_0 - \frac{k_0 t'}{\varepsilon^2} + \tilde{Y}(t') \right). \end{aligned} \quad (1.20)$$

Let us keep only the leading order terms in (1.20). The analysis that we perform on the simplified system may be applied to the full problem as well albeit at the price of somewhat

lengthier calculations that we are not willing to pay at the moment. Then (1.20) becomes (we now drop both the primes and tildes)

$$\begin{aligned}\dot{Y}(t) &= -P(t), \quad Y(0) = 0 \\ \dot{P}(t) &= \frac{1}{\varepsilon} \left[I - (\hat{k}_0 \otimes \hat{k}_0) \right] \nabla \mu \left(x_0 - \frac{k_0 t}{\varepsilon^2} + Y(t) \right), \quad P(0) = 0,\end{aligned}\tag{1.21}$$

which is the system we will study. The vector $\dot{P}(t)$ is orthogonal to k_0 for all $t \geq 0$ – hence so is $P(t)$ and thus $(Y(t) \cdot k_0) = 0$ for all $t \geq 0$ as well. This is a familiar phenomenon for the perturbation theory – the first order correction is orthogonal to the mean displacement. It is convenient to set $x_0 = 0$ and choose the coordinate axes so that $k_0 = e_n$, the unit vector in the direction of x_n . Then $Y(t) = (Y_1, \dots, Y_{n-1}, 0)$, $P(t) = (P_1, \dots, P_{n-1}, 0)$ and (1.21) may be rewritten as the following system for $Z(t) = (Y_1(t), \dots, Y_{n-1}(t))$ and $Q(t) = (P_1(t), \dots, P_{n-1}(t))$:

$$\begin{aligned}\dot{Z}(t) &= -Q(t), \quad Z(0) = 0 \\ \dot{Q}(t) &= \frac{1}{\varepsilon} G \left(\frac{t}{\varepsilon^2}, Z(t) \right), \quad Q(0) = 0,\end{aligned}\tag{1.22}$$

where

$$G(x_n, x') = \left(\frac{\partial \mu(x', -x_n)}{\partial x_1}, \dots, \frac{\partial \mu(x', -x_n)}{\partial x_{n-1}} \right), \quad x' \in \mathbb{R}^{n-1}.$$

A very formal derivation of the diffusive limit

We now describe a very formal but quick and effective way to obtain the limit of (1.22) as $\varepsilon \rightarrow 0$. Let us write the corresponding Liouville equation

$$\frac{\partial \phi}{\partial t} + q \cdot \nabla_z \phi - \frac{1}{\varepsilon} G \left(\frac{t}{\varepsilon^2}, z \right) \cdot \nabla_q \phi = 0.\tag{1.23}$$

Instead of assuming that the random function $G(s, x)$ is as in (1.22) we make a more general hypothesis that for each $x \in \mathbb{R}^n$ the process $G(s, x) \in \mathbb{R}^n$ is stationary in time with the two-point correlation tensor

$$R_{ml}(s, x) = \langle G_m(t, x) G_l(t + s, x) \rangle.$$

We seek the solution as a multiple time scale expansion

$$\phi = \phi_0(t, z) + \varepsilon \phi_1(t, \tau, z) + \varepsilon^2 \phi_2(t, \tau, z) + \dots, \quad \tau = t/\varepsilon^2.\tag{1.24}$$

As usual in such expansions in random media we assume that the leading order term is independent of the fast variable and is deterministic. The higher order corrections are assumed to be stationary in the fast variable τ . These assumptions are typically very hard to justify rigorously – nevertheless they often provide the correct answer. We insert the expansion into (1.23) and obtain in the leading order $O(1/\varepsilon)$

$$\frac{\partial \phi_1}{\partial \tau} = G(\tau, z) \cdot \nabla_q \phi_0(t, z)$$

so that

$$\phi_1(t, z, \tau) = \chi(\tau, z) \cdot \nabla_q \phi_0(t, z)\tag{1.25}$$

with the corrector $\chi(\tau, z)$ that solves $\dot{\chi} = G(\tau, z)$. It is very convenient to introduce a regularization parameter θ that we will send to zero later and write

$$\chi_m(\tau, z) = \int_{-\infty}^{\tau} e^{\theta s} G_m(s, z) ds.\tag{1.26}$$

The terms of the order $O(1)$ in (1.23) are

$$\frac{\partial \phi_0}{\partial t} + q \cdot \nabla_z \phi_0 - G(\tau, z) \cdot \nabla_q \phi_1 + \frac{\partial \phi_2}{\partial \tau} = 0.$$

We take the expectation of this equation using the fact that ϕ_0 is deterministic and argue that because ϕ_2 is stationary in τ we have

$$\left\langle \frac{\partial \phi_2}{\partial \tau} \right\rangle = 0.$$

With these two closure assumptions we obtain

$$\frac{\partial \phi_0}{\partial t} + q \cdot \nabla_z \phi_0 = \langle G(\tau, z) \cdot \nabla_q \phi_1 \rangle.$$

The term on the right side is computed explicitly using expression (1.25)-(1.26) for ϕ_1 :

$$\begin{aligned} \langle G(\tau, z) \cdot \nabla_q \phi_1 \rangle &= \left\langle G_m(\tau, z) \frac{\partial}{\partial q_m} \left[\int_{-\infty}^{\tau} e^{\theta s} G_p(s, z) ds \frac{\partial \phi_0}{\partial q_p} \right] \right\rangle \\ &= \int_{-\infty}^{\tau} e^{\theta s} R_{mp}(s - \tau, z) ds \frac{\partial^2 \phi_0}{\partial q_m \partial q_p} \rightarrow \int_{-\infty}^0 R_{mp}(s, z) ds \frac{\partial^2 \phi_0}{\partial q_m \partial q_p} \text{ as } \theta \rightarrow 0. \end{aligned}$$

Therefore, the function $\phi_0(t, q, z)$ satisfies a degenerate parabolic equation

$$\frac{\partial \phi_0}{\partial t} + q \cdot \nabla_z \phi_0 = D_{mp}(z) \frac{\partial^2 \phi_0}{\partial q_m \partial q_p} \quad (1.27)$$

with the symmetrized diffusion coefficient

$$D_{mp}(z) = \frac{1}{2} \left[\int_{-\infty}^0 R_{mp}(s, z) ds + \int_{-\infty}^0 R_{pm}(s, z) ds \right] = \frac{1}{2} \int_{-\infty}^{\infty} R_{mp}(s, z) ds.$$

If the statistics of $G(s, x)$ is identical for all $x \in \mathbb{R}^n$ then the diffusion matrix is constant in space. This means that in the limit $\varepsilon \rightarrow 0$ the process $Q(t)$ becomes a diffusion while $Z(t)$ is its integral in time.

Going back to the short time asymptotics for the geometric optics we see that the rescaled deviation of the wave vector from its original value k_0 converges to a diffusion process $Q(t)$ and the deviation of the spatial position from its average $k_0 t$ converges to the time integral of $Q(t)$ on a time scale of the order $O(\delta^{-2/3})$. This time is much longer than the naive prediction $O(\delta^{-1/2})$ discussed below (1.18). Here δ is the relative size of the variations of the refraction index. This provides a formalization of the ray diffusion we have discussed in Section 1.1.2, at least for short times. It turns out that the randomization of the wave vector on the time scale $O(\delta^{-2/3})$ is related to the appearance of a caustic. It has been shown by B. White [61] appears on this time scale with probability one. This means that the ray approach in a random medium works only on a very short time scale as caustics appear very quickly. On the other hand, one may follow the solutions of the Liouville equations for arbitrarily long times. This will be very useful when we discuss the long time behavior of the geometric optics limit, where caustics become too numerous and we can not keep track of them.

1.2 Basic facts on weak convergence in C and D

Weak convergence

We recall in this section basic facts from [11] 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

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

The space D

The cadlag functions. 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 consists of functions that have limits on the left and are continuous on the right:

- (i) For $0 \leq t < 1$ the right limit $x(t^+) = \lim_{s \rightarrow t^+} x(s)$ exists and $x(t) = x(t^+)$.
 - (ii) For $0 < t \leq 1$ the left limit $x(t^-) = \lim_{s \rightarrow t^-} x(s)$.
- (1.28)

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

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

$$w_x(S) = \sup_{s, t \in S} |x(s) - x(t)|. \quad (1.29)$$

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 $n \in \mathbb{N}$.

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

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

and

$$\sup_{t \in [0, 1]} |x(t) - y(\lambda(t))| = \sup_{t \in [0, 1]} |x(\lambda^{-1}(t)) - y(t)| < \varepsilon. \quad (1.30)$$

This metric defines the Skorohod topology.

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

$$|x_n(t) - x(t)| \leq |x_n(t) - x(\lambda_n(t))| + |x(\lambda_n(t)) - x(t)|, \quad (1.31)$$

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, Skorohod convergence implies pointwise convergence except on a countable set of points. In addition (1.31) implies that if the limit $x(t)$ is continuous on $[0, 1]$ (and hence uniformly continuous) then the Skorohod convergence implies the uniform convergence.

Compactness in D

Compactness in terms of $w'_x(\delta)$. Modulus of continuity is not a right notion for a function in D as $w_x(\delta)$ does not vanish in the limit $\delta \rightarrow 0$. An alternative modulus which allows for jumps is defined as follows. A useful generalization of the modulus of continuity is the following modulus

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

This is a relaxation of the continuity modulus as it is not hard to see that $w''_x(\delta) \leq w_x(\delta)$. However, there is no inequality in the opposite direction: for instance, the functions

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

we have $w''_{x_n}(\delta) = 0$ while $w_{x_n}(\delta) = 1$.

A direct analog of the Arzela-Ascoli theorem is the following.

Theorem 1.2.1 *A necessary and sufficient condition for a sequence P_n of probability measures on D to be tight is that*

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

and

$$(ii.1) \quad \lim_{\delta \rightarrow 0} \limsup_n P_n [x : w''_x(\delta) \geq \varepsilon] = 0 \text{ for all } \varepsilon > 0$$

and

$$(ii.2) \quad \begin{cases} \lim_{\delta \rightarrow 0} \limsup_n P_n [x : |x(\delta) - x(0)| \geq \varepsilon] = 0 \\ \lim_{\delta \rightarrow 0} \limsup_n P_n [x : |x(1^-) - x(1 - \delta)| \geq \varepsilon] = 0. \end{cases}$$

A convenient and more practical criterion for tightness is the following: assume that there exists $\beta \geq 0$ and $\alpha > 1/2$ so that for all $r \leq s \leq t$ and $\lambda > 0$ we have

$$P [\min\{|X_s^n - X_r^n|, |X_t^n - X_s^n|\} \geq \lambda] \leq \frac{C}{\lambda^{4\beta}} |F(t) - F(s)|^{2\alpha}, \quad (1.32)$$

where F is a non-decreasing continuous function on $[0, 1]$. The key estimate in the proof of this criterion is that (1.32) implies that there exists a constant K that depends only on C , α and β so that

$$P [w''_{X^n}(\delta) \geq \varepsilon] \leq \frac{K}{\varepsilon^{4\beta}} (F(1) - F(0)) [w_F(2\delta)]^{2\alpha-1}, \quad (1.33)$$

where w_F is the modulus of continuity of the function F . This means that (1.32) ensures that condition (ii.1) of Theorem 1.2.1 holds. A useful and verifiable condition that guarantees (1.32) is that there exist $\beta > 0$, $\alpha > 1/2$ and $C > 0$ so that

$$\mathbb{E} \left\{ |X_s^n - X_r^n|^{2\beta} |X_t^n - X_s^n|^{2\beta} \right\} \leq C |t - r|^{2\alpha} \quad (1.34)$$

for all n . Then we may take $F(t) = t$ and (1.33) becomes

$$P[w''_{X_n}(\delta) \geq \varepsilon] \leq \frac{K}{\varepsilon^{4\beta}} \delta^{2\alpha-1}. \quad (1.35)$$

This is why we need $\alpha > 1/2$ in (1.34). It follows that we may use (1.34) as a substitute for condition (ii.1) in Theorem 1.2.1.

In turn, the following condition is sufficient to ensure that (1.34) holds: for any $T > 0$ and $\nu > 0$ there exists a constant $C(T, \nu)$ so that for all n , and all $0 \leq s \leq t \leq u \leq T$, we have

$$\mathbb{E}\{|X_n(u) - X_n(t)|^2 |X_n(t) - X_n(s)|^\nu\} \leq C(T, \nu)(u - t)\mathbb{E}\{|X_n(t) - X_n(s)|^\nu\}. \quad (1.36)$$

Indeed, when $\nu = 0$ in (1.36) we have

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

Taking $\nu = 2$ in (1.36) we get, using the above:

$$\begin{aligned} \mathbb{E}\{|X_n(u) - X_n(t)|^2 |X_n(t) - X_n(s)|^2\} &\leq C(T, \nu)(u - t)\mathbb{E}\{|X_n(t) - X_n(s)|^2\} \\ &\leq C(T, \nu)(u - t)(t - s) \leq C(T, \nu)(u - s)^2 \end{aligned} \quad (1.37)$$

and thus (1.34) indeed holds. A somewhat more general estimate than (1.36) is a reformulation in terms of the conditional expectation

$$\mathbb{E}\{|X_n(t) - X_n(s)|^2 | \mathcal{F}_s\} \leq C(T)(t - s). \quad (1.38)$$

A practical advantage of working with the conditional expectation in (1.38) is that the power of $(t - s)$ on the right is equal to one, not larger than one as in (1.34).

1.3 A limit theorem for a particle in a random flow

We now return to the question of the limiting behavior of solutions of the ray equations (1.22)

$$\begin{aligned} \dot{Z}(t) &= -Q(t), \quad Z(0) = 0 \\ \dot{Q}(t) &= \frac{1}{\varepsilon} G\left(\frac{t}{\varepsilon^2}, Z(t)\right), \quad Q(0) = 0. \end{aligned} \quad (1.39)$$

The rigorous approach to this problem lies via the more general problem of the behavior of a particle in a rapidly varying in time random flow:

$$\dot{X} = \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, X\right), \quad X(0) = x, \quad (1.40)$$

with a random function V when $\varepsilon \ll 1$. This question goes back to the classical papers by Khasminskii [41] from the 60's with subsequent contributions by various authors: without any attempt at completeness we mention the work of Papanicolaou and Kohler [52], and Kesten and Papanicolaou [39]. We present a version of the limit theorem due to T. Komorowski [43].

Let us explain where the scaling in (1.40) comes from, apart from coinciding with that in (1.39). 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 function $V(s, x)$ and introduce non-dimensional space-time variables $X = Y/x_0$, $s = T/t_0$:

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

Let us now assume that $\varepsilon \ll 1$ is a small parameter – physically this means that the time it takes the particle to pass one spatial correlation length is much larger than the correlation time of the random fluctuations. Therefore, in this regime the temporal randomness of $V(s, x)$ “dominates” the spatial variations. If we now introduce a slow time t so that $t = \varepsilon^2 s$, then in the variables (t, x) the particle obeys (1.40). The limit $\varepsilon \rightarrow 0$ now corresponds to observing the particle at times much larger than the correlation time of the random fluctuations and on the spatial scale of the order of the correlation length of the medium.

The first order equation corresponding to (1.40) is

$$\frac{\partial \phi}{\partial t} + \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, x\right) \cdot \nabla \phi = 0, \quad \phi(0, x) = \phi_0(x). \quad (1.41)$$

Its solution is $\phi(T, x) = \phi_0(X(0; x, T))$, where $X(t; x, T)$ is the solution of (1.40) with the final data prescribed at time T : $X(T; x, T) = x$.

When does one expect the trajectories of (1.40) 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.

The ray equations (1.39) are not quite of the form (1.40): one should consider a slightly more general dynamics with an additional slow component $F(t, x)$:

$$\frac{dX}{dt} = \frac{1}{\varepsilon} G\left(\frac{t}{\varepsilon^2}, X\right) + F(t, X), \quad X(0) = x, \quad (1.42)$$

with a function F which we will assume to be deterministic for simplicity. Then equations (1.39) are of the form (1.42) with $X = (Z, Q)$, $G = (0, V)$ and $F = (-Q, 0)$.

Assumptions on the random field

Stationarity. The random field $V(t, x)$ is strictly stationary in time and space. This means that for any $t_1, t_2, \dots, t_m \in \mathbb{R}$, $x_1, \dots, x_m \in \mathbb{R}^n$, and each $h \in \mathbb{R}$ and $y \in \mathbb{R}^n$ the joint distribution of $V(t_1 + h, x + y), V(t_2 + h, x + y), \dots, V(t_m + h, x + y)$ is the same as that of $V(t_1, x), V(t_2, x), \dots, V(t_m, x)$. We will denote by $R_{nm}(t, x)$ the two-point correlation tensor of $V(t, x)$:

$$R_{nm}(t, x) = \mathbb{E} \{ V_n(s, y) V_m(t + s, y + x) \}. \quad (1.43)$$

Mixing. Given $C > 0$ and $\rho > 0$ let us denote by $\mathcal{V}_a^b(C, \rho)$ the σ -algebra generated by the sets of the form $\{\omega : V(t, x, \omega) \in A\}$ where $a \leq t \leq b$, $|x| \leq C(1 + t^\rho)$ and A is a Borel set in \mathbb{R}^n . We will assume that there exists $C > 0$ and $1/2 < \rho < 1$ such that for any $m \geq 0$ the mixing coefficient

$$\beta(h; C, \rho) = \sup_t \sup_{A \in \mathcal{V}_{t+h}^\infty(C, \rho), B \in \mathcal{V}_0^t(C, \rho)} \frac{|P(A \cap B) - P(A)P(B)|}{P(B)} \quad (1.44)$$

satisfies

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

Boundedness. 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

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

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

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 mixing assumption is sometimes strengthened considering larger σ -algebras $\tilde{\mathcal{V}}_a^b$ generated by the sets of the form $\{\omega : V(t, x, \omega) \in A\}$ where $a \leq t \leq b$, $x \in \mathbb{R}^n$ (there is no restriction on x now) and A is a Borel set in \mathbb{R}^n with the corresponding mixing coefficient

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

The stronger assumption does not apply to shifts by a mean flow, that is, random fields of the form $V(t, x) = U(x - \bar{u}t)$, where $U(x)$ is a field that is mixing in space and \bar{u} is a mean flow. This is an important and interesting class of random fields that we would like to include in our consideration. The small price to pay for its inclusion is the modification of the mixing condition as in (1.44).

The spatial stationarity of $V(t, x)$ is not a necessary assumption but it allows to simplify a few expressions in what follows. This can be seen already from the formal computation in Section 1.1.3. 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.

The limit theorem

Let us define the diffusion matrix

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

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

Theorem 1.3.1 *Suppose that the random field $V(t, x)$ satisfies the assumptions above. Then the process $X_\varepsilon(t)$ converges weakly as $\varepsilon \rightarrow 0$ to the limit process $\bar{X}(t)$ that satisfies a stochastic differential equation*

$$d\bar{X}(t) = F(t, \bar{X}(t))dt + \sigma dW_t.$$

Here W_t is the standard Brownian motion.

The main result of [43] is actually much more general – it applies also to non-divergence free velocities. 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. Komorowski also accounts 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.

We will present the proof of Theorem 1.3.1 under two simplifying assumptions: first, the drift $F = 0$ and second, the matrix σ is invertible. While they do not subtract any of the essential aspects of the proof, they do shorten many expressions and calculations which are sufficiently long even without them. The proof proceeds in several steps. First, we establish a mixing lemma that translates the mixing properties of the random field into a “loss-of-memory” effect for the trajectories. Second, using the mixing lemma we establish the tightness of the family of processes $X_\varepsilon(t)$. This is done in the space D . However, as the processes $X_\varepsilon(t)$ are all continuous the limit process also has to be continuous and convergence take place in C . In the last step we identify the limit as a Brownian motion multiplied by the matrix σ by means of the martingale characterization of the Brownian motion.

The mixing lemmas

A crucial component in many proofs of this kind is some sort of a mixing lemma. It translates the mixing properties of the random field into the mixing properties of the trajectories. At the end of the day this allows us to split expectations into product of expectations and either “justify”, or explain away the closure assumptions that are often made formally such as what we did at the end of Section 1.1.3. In our particular problem it explains why the formal assumption that the leading order term in the asymptotic expansion (1.24) is deterministic produced the correct answer.

We set $G_0(s_1, x) = V(s_1, x)$ and

$$G_{1,j}(s_1, s_2, x) = \sum_{p=1}^n V_p(s_2, x) \frac{\partial V_j(s_1, x)}{\partial x_p}, \quad j = 1, \dots, n.$$

Incompressibility of $V(t, x)$ and its spatial stationarity imply that $\mathbb{E}\{G_1(s_1, s_2, x)\} = 0$. In the next lemma we drop C and ρ in the notation for the σ -algebras $\mathcal{V}_0^s(C, \rho)$.

Lemma 1.3.2 *Fix $T \geq 0$ and let $0 \leq u \leq s \leq T$. Assume that Y is a $\mathcal{V}_0^{s/\varepsilon^2}$ -measurable random vector function. Then there exists $\varepsilon_0 > 0$ and a constant $C > 0$ such that for any $0 \leq u \leq s \leq s_2 \leq s_1 \leq T$ and $0 < \varepsilon < \varepsilon_0$ we have*

$$\left| \mathbb{E} \left\{ V \left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u) \right) Y \left(\frac{s}{\varepsilon^2} \right) \right\} \right| \leq C \beta(s_1 - s) \mathbb{E} \left| Y \left(\frac{s}{\varepsilon^2} \right) \right|, \quad (1.45)$$

$$\left| \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[V \left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u) \right) \right] Y \left(\frac{s}{\varepsilon^2} \right) \right\} \right| \leq C \beta(s_1 - s) \mathbb{E} \left| Y \left(\frac{s}{\varepsilon^2} \right) \right|, \quad (1.46)$$

and

$$\left| \mathbb{E} \left\{ G_1 \left(\frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u) \right) Y \left(\frac{s}{\varepsilon^2} \right) \right\} \right| \leq C \beta^{1/2}(s_1 - s_2) \beta^{1/2}(s_2 - s) \mathbb{E} \left| Y \left(\frac{s}{\varepsilon^2} \right) \right|, \quad (1.47)$$

$$\left| \mathbb{E} \left\{ \frac{\partial}{\partial x_k} \left[G_1 \left(\frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u) \right) \right] Y \left(\frac{s}{\varepsilon^2} \right) \right\} \right| \leq C \beta^{1/2}(s_1 - s_2) \beta^{1/2}(s_2 - s) \mathbb{E} \left| Y \left(\frac{s}{\varepsilon^2} \right) \right|, \quad (1.48)$$

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

The proof of tightness

We will establish the inequality

$$\mathbb{E} \{ |X_\varepsilon(t) - X_\varepsilon(s)|^2 |X_\varepsilon(s) - X_\varepsilon(u)|^2 \} \leq C(t-u)^{1+\nu} \quad (1.49)$$

with $\nu > 0$. This is the criterion (1.34) for tightness in the space D . The main step in the proof is to find $\gamma \in (1, 2)$ such that for all times t and s such that $t - s > 10\varepsilon^\gamma$ we have an estimate for the conditional expectation

$$\mathbb{E} \{ |X_\varepsilon(t) - X_\varepsilon(s)|^2 | \mathcal{F}_s \} \leq C(t-s) \text{ for } t-s > 10\varepsilon^\gamma. \quad (1.50)$$

Step 0. Nearby times. As we have explained before, the estimate (1.50) itself is sufficient to establish tightness in D for the family $X_\varepsilon(t)$ if it were to hold for all $t > s$. 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

$$\mathbb{E} \{ |X_\varepsilon(t) - X_\varepsilon(s)|^2 |X_\varepsilon(s) - X_\varepsilon(u)|^2 \} \leq C(t-u)^{1+\nu} \text{ for } t-s > 10\varepsilon^\gamma \text{ and } s-u > 10\varepsilon^\gamma.$$

However, with an appropriate choice of $\gamma \in (1, 2)$, if either $t - s \leq 10\varepsilon^\gamma$ or $s - u \leq 10\varepsilon^\gamma$ (or both), the estimate (1.49) follows from (1.50) together with the dynamical system (1.42) governing $X_\varepsilon(t)$: this is a straightforward computation.

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 of the interval $[s, t]$ into subintervals of the length

$$\Delta t = l_\varepsilon = (t-s) \left(\left\lfloor \frac{t-s}{\varepsilon^\gamma} \right\rfloor \right)^{-1},$$

where $[x]$ is the integer part of x . Then the time step l_ε is such that $\varepsilon^\gamma/2 \leq l_\varepsilon \leq 2\varepsilon^\gamma$ and the partition is $s = t_0 < t_1 < \dots < t_{M+1} = t$ with a time step $\Delta t = t_{i+1} - t_i = l_\varepsilon$. The parameter $\gamma \in (1, 2)$ is to be defined later. The important aspect is that $\gamma < 2$ so that Δt is much larger than the velocity correlation time ε^2 . The basic idea in the proof of (1.50) is “to expand $X_\varepsilon(t) - X_\varepsilon(s)$ in a Taylor series” with a “large” time step $O(\Delta t)$. This will produce explicitly computable terms which are the first two terms in this expansion. The error terms which are nominally large are shown to be small using the mixing Lemma 1.3.2.

Dropping the subscript ε of X_ε we write for $t > s$, using the second-order Taylor expansion :

$$\begin{aligned} X(t) - X(s) &= \frac{1}{\varepsilon} \int_s^t V\left(\frac{u}{\varepsilon^2}, X(u)\right) du = \frac{1}{\varepsilon} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} V\left(\frac{u}{\varepsilon^2}, X(u)\right) du \\ &= \frac{1}{\varepsilon} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} V\left(\frac{u}{\varepsilon^2}, X(t_{i-1})\right) du + \frac{1}{\varepsilon^2} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} \left[\int_{t_{i-1}}^u G_1\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(u_1)\right) du_1 \right] du \\ &= \frac{1}{\varepsilon} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} V\left(\frac{u}{\varepsilon^2}, X(t_{i-1})\right) du + \frac{1}{\varepsilon^2} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} \left[\int_{t_{i-1}}^u G_1\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1})\right) du_1 \right] du \\ &\quad + \frac{1}{\varepsilon^3} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} \left[\int_{t_{i-1}}^u \left[\int_{t_{i-1}}^{u_1} G_2\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, \frac{u_2}{\varepsilon^2}, X(u_2)\right) du_2 \right] du_1 \right] du \end{aligned} \quad (1.51)$$

with

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

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 $t-s \geq 10\varepsilon^\gamma$. Therefore, the last integral is bounded by

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

which is small if $\gamma > 3/2$. This is a general idea in proofs of weak coupling limits: pull back one time step and expand the integrands until they become almost surely small, then compute the limit of the (very) finite number of surviving terms. In our present case we have shown that, for $3/2 < \gamma < 2$,

$$X(t) - X(s) = L_1(s, t) + L_2(s, t) + E(s, t)$$

where

$$L_1(s, t) = \frac{1}{\varepsilon} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} V \left(\frac{u}{\varepsilon^2}, X(t_{i-1}) \right) du$$

and

$$L_2(s, t) = \frac{1}{\varepsilon^2} \sum_{i=0}^M \int_{t_i}^{t_{i+1}} \left[\int_{t_{i-1}}^u G_1 \left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1}) \right) du_1 \right] du,$$

while $|E(s, t)| \leq C\varepsilon^\alpha(t-s)$ with some $\alpha > 0$ and a deterministic constant $C > 0$. This finishes the first preliminary step in the proof of tightness.

Step 2. Application of the tightness criterion. Now we are ready to prove (1.50). That is, we have to verify that for any non-negative and $\mathcal{V}_0^{s/\varepsilon^2}$ -measurable random variable Y we have for all $0 \leq s \leq t \leq T$ such that $t \geq s + 10\varepsilon^\gamma$:

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

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

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

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

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

The idea is to use separation between t_{i-1} and t_{j-1} and apply the mixing lemma. Accordingly we look at the cases $i \leq j-2$, $i = j-1$ and $i = j$ separately as the terms end up being of a different order. The terms with $i \leq j-2$ may be estimated with the help of the mixing Lemma

1.3.2 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 :

$$\begin{aligned} \sum_{j=0}^M \sum_{i \leq j-2} |I_{ij}| &\leq \frac{C}{\varepsilon^2} \sum_{j=0}^M \sum_{i \leq j-2} \int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} \beta\left(\frac{u' - t_{j-1}}{\varepsilon^2}\right) \mathbb{E}\{Y\} du' du \\ &\leq \frac{C}{\varepsilon^2} \beta(\varepsilon^{2-\gamma}) (t-s)^2 E\{Y\} \leq C\varepsilon^p (t-s) E\{Y\} \end{aligned}$$

for any $p > 0$ since $\gamma < 2$ and $\beta(s)$ decays faster than any power of s . The term I_3 corresponding to $i = j$ can be estimated using the mixing lemma again, using the fact that t_{j-1} is smaller than both u and u' :

$$\begin{aligned} \sum_{j=0}^M |I_{jj}| &\leq \frac{C}{\varepsilon^2} \sum_{j=0}^M \int_{t_j}^{t_{j+1}} \int_{t_j}^{t_{j+1}} \mathbb{E}\left\{V_p\left(\frac{u}{\varepsilon^2}, X(t_{j-1})\right) V_p\left(\frac{u'}{\varepsilon^2}, X(t_{j-1})\right) Y\right\} du' du \quad (1.52) \\ &\leq \frac{2C}{\varepsilon^2} \sum_{j \in I} \int_{t_j}^{t_{j+1}} \int_{u'}^{t_{j+1}} \beta\left(\frac{u - u'}{\varepsilon^2}\right) du du' E\{Y\} \leq C(t-s) E\{Y\} \int_0^\infty \beta(u) du. \end{aligned}$$

The integral I_2 with $i = j - 1$ is estimated similarly.

A better estimate estimate for L_1 . Let us now go one step further and actually identify the limit of $E\{L_{1,j}(s, t)L_{1,m}(s, t)Y\}$ with $1 \leq j, m \leq n$. The previous calculations already show that the term corresponding to the previous I_1 (but now with V_j and V_m replacing V_p and V_p) satisfies $|I_1| \leq C\varepsilon^\alpha (t-s) E\{Y\}$ with $\alpha > 0$ so we are interested only in the limit of I_2 and I_3 . The term I_3 is computed as in (1.52) with the help of the mixing lemma:

$$\begin{aligned} \sum_{j \in I} |I_{jj}| &= \frac{1}{\varepsilon^2} \sum_{j=0}^M \int_{t_j}^{t_{j+1}} \int_{t_j}^{t_{j+1}} \mathbb{E}\left\{V_j\left(\frac{u}{\varepsilon^2}, X(t_{j-1})\right) V_m\left(\frac{u'}{\varepsilon^2}, X(t_{j-1})\right) Y\right\} du' du \quad (1.53) \\ &= \frac{1}{\varepsilon^2} \sum_{j=0}^M \int_{t_j}^{t_{j+1}} \int_{t_j}^{t_{j+1}} R_{jm}\left(\frac{u - u'}{\varepsilon^2}, 0\right) du du' E\{Y\} + o(1)(t-s) E\{Y\} \\ &= \left[\int_{-\infty}^\infty R_{jm}(\tau, 0) d\tau + o(1) \right] (t-s) E\{Y\}. \end{aligned}$$

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

$$\begin{aligned} \sum_{j \in I} |I_{j-1,j}| &= \frac{1}{\varepsilon^2} \sum_{j=0}^M \int_{t_j}^{t_{j+1}} \int_{t_{j-1}}^{t_j} \mathbb{E}\left\{V_j\left(\frac{u}{\varepsilon^2}, X(t_{j-1})\right) V_m\left(\frac{u'}{\varepsilon^2}, X(t_{j-2})\right) Y\right\} du' du \quad (1.54) \\ &= \frac{1}{\varepsilon^2} \sum_{j \in I} \int_{t_j}^{t_{j+1}} \int_{t_{j-1}}^{t_j} R_{jm}\left(\frac{u - u'}{\varepsilon^2}, 0\right) du du' E\{Y\} + o(1)(t-s) E\{Y\} = o(1)(t-s) E\{Y\}. \end{aligned}$$

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

$$\mathbb{E}\{(L_{1,j}(s, t)L_{1,m}(s, t))Y\} = \left[\int_{-\infty}^\infty R_{jm}(\tau, 0) d\tau + o(1) \right] (t-s) E\{Y\}. \quad (1.55)$$

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

$$\mathbb{E} \{ (L_2(s, t))^2 Y \} \leq C(t - s) E \{ Y \}. \quad (1.56)$$

There is no reason to repeat these calculations separately for L_2 except that an even stronger estimate than (1.56) holds with an appropriate choice of γ :

$$\mathbb{E} \{ (L_2(s, t))^2 Y \} \leq C\varepsilon^\alpha (t - s) E \{ Y \} \quad (1.57)$$

with $\alpha > 0$. We will need (1.57) in the identification of the limit, thus we will show it now: $\mathbb{E} \{ (L_2(s, t))^2 Y \}$ is equal to

$$\frac{1}{\varepsilon^4} \sum_{i,j} \int_{t_i}^{t_{i+1}} du \int_{t_j}^{t_{j+1}} du' \int_{t_{i-1}}^u du_1 \int_{t_{j-1}}^{u'} du'_1 \mathbb{E} \left\{ G_1 \left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1}) \right) G_1 \left(\frac{u'}{\varepsilon^2}, \frac{u'_1}{\varepsilon^2}, X(t_{j-1}) \right) Y \right\}.$$

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

$$\begin{aligned} & \frac{1}{\varepsilon^4} \sum_{i \leq j-2} \int_{t_i}^{t_{i+1}} du \int_{t_j}^{t_{j+1}} du' \int_{t_{i-1}}^u du_1 \int_{t_{j-1}}^{u'} du'_1 \mathbb{E} \left\{ G_1 \left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1}) \right) G_1 \left(\frac{u'}{\varepsilon^2}, \frac{u'_1}{\varepsilon^2}, X(t_{j-1}) \right) Y \right\} \\ & \leq C\varepsilon^{2\gamma-4} \beta (\varepsilon^{\gamma-2}) (t - s)^2 \mathbb{E} \{ Y \}. \end{aligned}$$

We used in the above estimate the mixing lemma with the gap between t_{i-1} and t_{j-1} as well as the fact that the length of each time interval is 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:

$$\begin{aligned} & \frac{1}{\varepsilon^4} \sum_i \int_{t_i}^{t_{i+1}} du \int_{t_i}^{t_{i+1}} du' \int_{t_{i-1}}^u du_1 \int_{t_{i-1}}^{u'} du'_1 \mathbb{E} \left\{ G_1 \left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1}) \right) G_1 \left(\frac{u'}{\varepsilon^2}, \frac{u'_1}{\varepsilon^2}, X(t_{i-1}) \right) Y \right\} \\ & \leq C\varepsilon^{3\gamma-4} (t - s) \mathbb{E} \{ Y \} \end{aligned}$$

simply because now the number of summands is bounded by $(2(t - s)/\varepsilon^\gamma)$ (without the square). This means that if we take $\gamma > 4/3$ this term is bounded by the right side of (1.57). The contribution of the terms with $i = j - 1$ is estimated identically – hence (1.57) indeed holds.

Summarizing our work so far (and restoring the missing indices) we have shown that

$$\mathbb{E} \{ (X_m(t) - X_m(s))(X_n(t) - X_n(s))Y \} = \left[\int_{-\infty}^{\infty} R_{mn}(\tau, 0) d\tau + o(1) \right] (t - s) E \{ Y \} \quad (1.58)$$

for all $t > s$ with $t - s \geq 10\varepsilon^\gamma$. This, of course, implies (1.50) and hence the tightness of the family $X_\varepsilon(t)$ follows.

Identification of the limit

In order to identify the limit, using the Levy theorem (the martingale characterization of the Brownian motion) (see, for instance, Theorem 3.16 in [37]) all we have to do is verify that the limit is continuous (that we already know) and the following two conditions hold: first,

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

for all bounded non-negative continuous functions $\Psi = \Psi(X_\varepsilon(t_1), \dots, X_\varepsilon(t_n))$ with $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq s < t \leq T$. Second, we need

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

for all $t > 0$. These conditions allow us to conclude that the limit process is a martingale. 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 1.3.1. \square

Chapter 2

Random geometric optics II: the long time limit, from rays to diffusion

Here we study the long time asymptotics of rays in a weakly random medium. This problem is analyzed in the general setting of a particle in a weakly random Hamiltonian field. This chapter is based on the results of [6] and [44].

2.1 A particle in a random Hamiltonian

We have considered in Chapter 1 the asymptotic behavior of a ray in a medium with weakly random sound speed and have seen that on a short time scale the rescaled deviation of the direction of the ray from its original value becomes a diffusion process. The long time behavior of this system is an example of the analysis of the long time, large distance behavior of a particle in a weakly random time-independent Hamiltonian flow. It turns out that this limit is also described by the momentum diffusion but now, of course, without rescaling of the momentum: the particle momentum itself undergoes the Brownian motion on the energy sphere. This intuitive result has been first proved in [40] for a classical particle in dimensions higher than two, and later extended to two dimensions with the Poisson distribution of scatterers in [22], and in a general two-dimensional setting in [45]. On the other hand, the long time limit of a momentum diffusion is the standard spatial Brownian motion. Hence, a natural question arises if it is possible to obtain such a Brownian motion directly as the limiting description in the original problem of a particle in a quenched random potential. This necessitates the control of the particle behavior over times longer than those when the momentum diffusion holds. This is what we do in this chapter.

We consider a particle that moves in an isotropic weakly random Hamiltonian flow with the Hamiltonian of the form $H_\delta(x, k) = H_0(k) + \sqrt{\delta}H_1(x, k)$, $k = |k|$, and $x, k \in \mathbb{R}^d$ with $d \geq 3$:

$$\frac{dX^\delta}{dt} = \nabla_k H_\delta, \quad \frac{dK^\delta}{dt} = -\nabla_x H_\delta, \quad X^\delta(0) = 0, \quad K^\delta(0) = k_0. \quad (2.1)$$

Here $H_0(k)$ is the background Hamiltonian and $H_1(x, k)$ is a random perturbation, while the small parameter $\delta \ll 1$ measures the relative strength of random fluctuations. One expects that the effect of the random fluctuation would be of order one on the time scale of the order

$t \sim O(\delta^{-1})$. And indeed, as we have mentioned, it has been shown in [40] that, when

$$H_\delta(x, k) = \frac{k^2}{2} + \sqrt{\delta}V(x),$$

and under certain mixing assumptions on the random potential $V(x)$, the momentum process $K^\delta(t/\delta)$ converges to a diffusion process $K(t)$ on the sphere $k = k_0$ and the rescaled spatial component $\tilde{X}^\delta(t) = \delta X^\delta(t/\delta)$ converges to $X(t) = \int_0^t K(s)ds$. This is the momentum diffusion mentioned above. Another special case,

$$H_\delta(x, k) = (c_0 + \sqrt{\delta}c_1(x))|k|, \tag{2.2}$$

arises in the geometrical optics limit of wave propagation and this is the problem we are mostly interested in these notes. Here c_0 is the background sound speed, and $c_1(x)$ is a random perturbation. This case has been considered in [6], where it has been shown that, once again, $K^\delta(t/\delta)$ converges to a diffusion process $K(t)$ on the sphere $\{k = k_0\}$ while $\tilde{X}^\delta(t) = \delta X^\delta(t/\delta)$ converges to $X(t) = c_0 \int_0^t \hat{K}(s)ds$, $\hat{K}(t) := K(t)/|K(t)|$.

We show in this chapter how the momentum diffusion may be obtained and that this analysis may be pushed beyond the time of the momentum diffusion, so that under certain assumptions concerning the mixing properties of H_1 in the spatial variable there exists $\alpha_0 > 0$ such that the process $\delta^{1+\alpha} X^\delta(t/\delta^{1+2\alpha})$ converges to the standard Brownian motion in \mathbb{R}^d for all $\alpha \in (0, \alpha_0)$. The main difficulty of the proof is to obtain error estimates in the convergence of $K^\delta(\cdot)$ to the momentum diffusion. The error estimates allow us to push the analysis to times much longer than δ^{-1} where the momentum diffusion converges to the standard Brownian motion. The method of the proof is a modification of the cut-off technique used in [6] and [40].

A similar question arises in the semi-classical limit of the quantum mechanics and high frequency wave propagation. The Wigner transform, or the phase space energy density of the solution of the Schrödinger equation, is approximated in a weakly random medium by the solution of a deterministic linear Boltzmann equation. This behavior is also conjectured for the acoustic and other waves in a weakly random medium [53]. As in the momentum diffusion model for a particle, the long time limit of the Boltzmann equation is the spatial diffusion equation. It has been recently shown by Erdős, Salmhofer and Yau in [24, 25] that, indeed, one may push the Erdős-Yau analysis of [23] beyond the times on which the Boltzmann equation holds and obtain the diffusive behavior of the energy density of the solutions of the Schrödinger equation in the weak coupling limit.

We also discuss in some detail the application our results to the problem of multiple scattering of the acoustic waves. Our approach is different from that of [24, 25] mentioned above: we first consider the random geometrical optics approximation of the wave phase space energy density. The rays in the phase space satisfy the Hamiltonian equations (2.1) with the Hamiltonian given by (2.2). Therefore, the aforementioned convergence result of the solutions of (2.1) to the standard Brownian motion, combined with the standard error estimates [6, 49] on the geometrical optics approximation of the Wigner distribution of the solutions of the wave equation, allows us to establish rigorously the diffusive behavior of the wave energy density. To the best of our knowledge, this is the first result of such kind for classical waves.

2.2 The main result and preliminaries

2.2.1 The notation

As we will avoid the singular point $k = 0$, where the dynamics may not be well-defined, it is convenient to introduce $\mathbb{R}_*^d := \mathbb{R}^d \setminus \{0\}$ and $\mathbb{R}_*^{2d} := \mathbb{R}^d \times \mathbb{R}_*^d$. Also $\mathbb{S}_R^{d-1}(x)$ ($\mathbb{B}_R(x)$) shall stand

for a sphere (open ball) in \mathbb{R}^d of radius $R > 0$ centered at x . We shall drop writing either x , or R in the notation of the sphere (ball) in the particular cases when either $x = 0$, or $R = 1$. For a fixed $M > 0$ we define the spherical shell $A(M) := [k \in \mathbb{R}_*^d : M^{-1} \leq |k| \leq M]$ in the k -space, and $\mathcal{A}(M) := \mathbb{R}^d \times A(M)$ in the whole phase space. Given a vector $\mathbf{v} \in \mathbb{R}_*^d$ we denote by $\hat{\mathbf{v}} := \mathbf{v}/|\mathbf{v}| \in \mathbb{S}^{d-1}$ the unit vector in the direction of \mathbf{v} .

For any non-negative integers p, q, r , positive times $T > T_* \geq 0$ and a function $G : [T_*, T] \times \mathbb{R}_*^{2d} \rightarrow \mathbb{R}$ that has p, q and r derivatives in the respective variables we define

$$\|G\|_{p,q,r}^{[T_*, T]} := \sum_{(t,x,k) \in [T_*, T] \times \mathbb{R}^{2d}} \sup_{|\beta| \leq q, |\gamma| \leq r} |\partial_t^\alpha \partial_x^\beta \partial_k^\gamma G(t, x, k)|. \quad (2.3)$$

The summation range covers all integers $0 \leq \alpha \leq p$ and all integer valued multi-indices $|\beta| \leq q$ and $|\gamma| \leq r$. In the special case when $T_* = 0, T = +\infty$ we write $\|G\|_{p,q,r} = \|G\|_{p,q,r}^{[0, +\infty)}$. We denote by $C_b^{p,q,r}([0, +\infty) \times \mathbb{R}_*^{2d})$ the space of all functions G with $\|G\|_{p,q,r} < +\infty$. We shall also consider spaces of bounded and a suitable number of times continuously differentiable functions $C_b^{p,q}(\mathbb{R}_*^{2d})$ and $C_b^p(\mathbb{R}_*^d)$ with the respective norms $\|\cdot\|_{p,q}$ and $\|\cdot\|_p$.

2.2.2 The background Hamiltonian

We assume that the background Hamiltonian $H_0(k)$ is isotropic, that is, it depends only on $k = |k|$, and is uniform in space. Moreover, we assume that $H_0 : [0, +\infty) \rightarrow \mathbb{R}$ is a strictly increasing function satisfying $H_0(0) \geq 0$ and such that it is of C^3 -class of regularity in $(0, +\infty)$ with $H'_0(k) > 0$ for all $k > 0$, and let

$$h^*(M) := \max_{k \in [M^{-1}, M]} (H'_0(k) + |H''_0(k)| + |H'''_0(k)|), \quad h_*(M) := \min_{k \in [M^{-1}, M]} H'_0(k). \quad (2.4)$$

Two examples of such Hamiltonians are the quantum Hamiltonian $H_0(k) = k^2/2$ and the acoustic wave Hamiltonian $H_0(k) = c_0 k$. The qualitative reason for the above assumptions on $H_0(k)$ is that we need the background dynamics ‘to take the particle to various regions where it will sample the nearly independent random fluctuations. The overall effect will then lead to a Markovian limit. This makes the problem much simpler than a seemingly similar problem

$$\dot{X} = V(X), \quad (2.5)$$

with a mixing in space and time-independent random field $V(x)$. Unlike our problem, (2.5) lack any mechanism to move the particle around which makes it extremely difficult to obtain any rigorous, or even formal results for the particle behavior in (2.5).

2.2.3 The random medium

Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, and let \mathbb{E} denote the expectation with respect to \mathbb{P} . We denote by $\|X\|_{L^p(\Omega)}$ the L^p -norm of a given random variable $X : \Omega \rightarrow \mathbb{R}$, $p \in [1, +\infty]$. Let $H_1 : \mathbb{R}^d \times [0, +\infty) \times \Omega \rightarrow \mathbb{R}$ be a random field that is measurable and strictly stationary in the first variable. This means that for any shift $x \in \mathbb{R}^d$, $k \in [0, +\infty)$, and a collection of points $x_1, \dots, x_n \in \mathbb{R}^d$ the laws of $(H_1(x_1 + x, k), \dots, H_1(x_n + x, k))$ and $(H_1(x_1, k), \dots, H_1(x_n, k))$ are identical. In addition, we assume that $\mathbb{E}H_1(x, k) = 0$ for all $k \geq 0$, $x \in \mathbb{R}^d$, the realizations of $H_1(x, k)$ are \mathbb{P} -a.s. C^2 -smooth in $(x, k) \in \mathbb{R}^d \times (0, +\infty)$ and they satisfy

$$D_{i,j}(M) := \max_{|\alpha|=i} \operatorname{ess-sup}_{(x,k,\omega) \in \mathbb{R}^d \times [M^{-1}, M] \times \Omega} |\partial_x^\alpha \partial_k^j H_1(\mathbf{x}, k; \omega)| < +\infty, \quad i, j = 0, 1, 2. \quad (2.6)$$

We define $\tilde{D}(M) := \sum_{0 \leq i+j \leq 2} D_{i,j}(M)$.

We suppose further that the random field is strongly mixing in the uniform sense. More precisely, for any $R > 0$ we let \mathcal{C}_R^i and \mathcal{C}_R^e be the σ -algebras generated by random variables $H_1(\mathbf{x}, k)$ with $k \in [0, +\infty)$, $\mathbf{x} \in \mathbb{B}_R$ and $\mathbf{x} \in \mathbb{B}_R^c$ respectively. The uniform mixing coefficient between the σ -algebras is

$$\phi(\rho) := \sup[|\mathbb{P}(B) - \mathbb{P}(B|A)| : R > 0, A \in \mathcal{C}_R^i, B \in \mathcal{C}_{R+\rho}^e],$$

for all $\rho > 0$. We suppose that $\phi(\rho)$ decays faster than any power: for each $p > 0$

$$h_p := \sup_{\rho \geq 0} \rho^p \phi(\rho) < +\infty. \quad (2.7)$$

The two-point spatial correlation function of the random field H_1 is

$$R(\mathbf{y}, k) := \mathbb{E}[H_1(\mathbf{y}, k)H_1(\mathbf{0}, k)].$$

We also assume that the correlation function $R(y, l)$ is of the C^∞ -class for a fixed $l > 0$, is sufficiently smooth in l , and that for any fixed $l > 0$

$$\hat{R}(k, l) \text{ does not vanish identically on any hyperplane } H_p = \{k : (k \cdot p) = 0\}. \quad (2.8)$$

Here $\hat{R}(k, l) = \int R(x, l) \exp(-ik \cdot x) dx$ is the power spectrum of H_1 .

The above assumptions are satisfied, for example, if $H_1(x, k) = c_1(x)h(k)$, where $c_1(x)$ is a stationary uniformly mixing random field with a smooth correlation function, and $h(k)$ is a smooth deterministic function.

2.2.4 The path-spaces

For fixed integers $d, m \geq 1$ we let $\mathcal{C}^{d,m} := C([0, +\infty); \mathbb{R}^d \times \mathbb{R}_*^m)$: we shall omit the subscripts in the notation of the path space if $m = d$. We define $(X(t), K(t)) : \mathcal{C}^{d,m} \rightarrow \mathbb{R}^d \times \mathbb{R}_*^m$ as the canonical mapping $(X(t; \pi), K(t; \pi)) := \pi(t)$, $\pi \in \mathcal{C}^{d,m}$ and also let $\theta_s(\pi)(\cdot) := \pi(\cdot + s)$ be the standard shift transformation.

For any $u \leq v$ denote by \mathcal{M}_u^v the σ -algebra of subsets of \mathcal{C} generated by $(X(t), K(t))$, $t \in [u, v]$. We write $\mathcal{M}^v := \mathcal{M}_0^v$ and \mathcal{M} for the σ algebra of Borel subsets of \mathcal{C} . It coincides with the smallest σ -algebra that contains all \mathcal{M}^t , $t \geq 0$.

Let $\delta_*(M) := H_0(M^{-1}) / (2\tilde{D}(M))$. For a given $M > 0$ and $\delta \in (0, \delta_*(M)]$ we let

$$M_\delta := \max \left\{ H_0^{-1}(H_0(M) + 2\sqrt{\delta}\tilde{D}(M)), \left[H_0^{-1} \left(H_0 \left(\frac{1}{M} \right) - 2\sqrt{\delta}\tilde{D}(M) \right) \right]^{-1} \right\}. \quad (2.9)$$

For a particle that is governed by the Hamiltonian flow generated by $H_\delta(x, k)$ we have

$$M_\delta^{-1} \leq |K(t)| \leq M_\delta$$

for all t provided that $K(0) \in A(M)$. Accordingly, we define $\mathcal{C}(T, \delta)$ as the set of paths $\pi \in \mathcal{C}$ so that both $(2M_\delta)^{-1} \leq |K(t)| \leq 2M_\delta$, and

$$\left| X(t) - X(u) - \int_u^t H'_0(K(s))\hat{K}(s)ds \right| \leq \tilde{D}(2M_\delta)\sqrt{\delta}(t-u), \text{ for all } 0 \leq u < t \leq T.$$

In the case when $\delta = 1$, or $T = +\infty$ we shall write simply $\mathcal{C}(T)$, or $\mathcal{C}(\delta)$ respectively.

2.2.5 The main results

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

$$\begin{aligned} \frac{\partial \phi^\delta}{\partial t} + \nabla_x H_\delta(x, k) \cdot \nabla_k \phi^\delta - \nabla_k H_\delta(x, k) \cdot \nabla_x \phi^\delta &= 0, \\ \phi^\delta(0, x, k) &= \phi_0(\delta x, k). \end{aligned} \quad (2.10)$$

We assume that the initial data $\phi_0(x, k)$ is a compactly supported function four times differentiable in k , twice differentiable in x whose support is contained inside a spherical shell $\mathcal{A}(M) = \{(x, k) : M^{-1} < |k| < M\}$ for some positive $M > 0$.

Let us define the diffusion matrix D_{mn} by

$$D_{mn}(\hat{k}, l) = -\frac{1}{2} \int_{-\infty}^{\infty} \frac{\partial^2 R(H'_0(l)s\hat{k}, l)}{\partial x_n \partial x_m} ds = -\frac{1}{2H'_0(l)} \int_{-\infty}^{\infty} \frac{\partial^2 R(s\hat{k}, l)}{\partial x_n \partial x_m} ds, \quad m, n = 1, \dots, d. \quad (2.11)$$

Then we have the following result.

Theorem 2.2.1 *Let ϕ^δ be the solution of (2.10) and let $\bar{\phi}$ satisfy*

$$\begin{aligned} \frac{\partial \bar{\phi}}{\partial t} &= \sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left(D_{mn}(\hat{k}, k) \frac{\partial \bar{\phi}}{\partial k_n} \right) + H'_0(k) \hat{k} \cdot \nabla_x \bar{\phi} \\ \bar{\phi}(0, x, k) &= \phi_0(x, k). \end{aligned} \quad (2.12)$$

Suppose that $M \geq M_0 > 0$ and $T \geq T_0 > 0$. Then, there exist two constants $C, \alpha_0 > 0$ such that for all $T \geq T_0$

$$\sup_{(t,x,k) \in [0,T] \times K} \left| \mathbb{E} \phi^\delta \left(\frac{t}{\delta}, \frac{x}{\delta}, k \right) - \bar{\phi}(t, x, k) \right| \leq CT(1 + \|\phi_0\|_{1,4})\delta^{\alpha_0} \quad (2.13)$$

for all compact sets $K \subset \mathcal{A}(M)$.

Note that

$$\begin{aligned} \sum_{m=1}^d D_{nm}(\hat{k}, k) \hat{k}_m &= - \sum_{m=1}^d \frac{1}{2H'_0(k)} \int_{-\infty}^{\infty} \frac{\partial^2 R(s\hat{k}, k)}{\partial x_n \partial x_m} \hat{k}_m ds \\ &= - \sum_{m=1}^d \frac{1}{2H'_0(k)} \int_{-\infty}^{\infty} \frac{d}{ds} \left(\frac{\partial R(s\hat{k}, k)}{\partial x_n} \right) ds = 0 \end{aligned}$$

and thus the K -process generated by (2.12) is indeed a diffusion process on a sphere $k = \text{const}$, or, equivalently, equations (2.12) for different values of k are decoupled. Assumption (2.8) implies that the matrix $D(\hat{k}, l)$ has rank $d - 1$ for each $\hat{k} \in \mathbb{S}^{d-1}$ and each $l > 0$.

Solutions of (2.12) 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 (2.12) with an initial data $\bar{\phi}_\gamma(0, t, x, k) = \phi_0(\gamma x, k)$. We also let $w(t, x, k)$ be the solution of the spatial diffusion equation:

$$\begin{aligned} \frac{\partial w}{\partial t} &= \sum_{m,n=1}^d a_{mn}(k) \frac{\partial^2 w}{\partial x_n \partial x_m}, \\ w(0, x, k) &= \bar{\phi}_0(x, k) \end{aligned} \quad (2.14)$$

with the averaged initial data

$$\bar{\phi}_0(x, k) = \frac{1}{\Gamma_{d-1}} \int_{\mathbb{S}^{d-1}} \phi_0(x, k) d\Omega(\hat{k}).$$

Here $d\Omega(\hat{k})$ is the surface measure on the unit sphere \mathbb{S}^{d-1} and Γ_n is the area of an n -dimensional sphere. The diffusion matrix $A := [a_{nm}]$ in (2.14) is given explicitly as

$$a_{nm}(k) = \frac{1}{\Gamma_{d-1}} \int_{\mathbb{S}^{d-1}} H'_0(k) \hat{k}_n \chi_m(k) d\Omega(\hat{k}). \quad (2.15)$$

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

$$\sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left(D_{mn}(\hat{k}, k) \frac{\partial \chi_j}{\partial k_n} \right) = -H'_0(k) \hat{k}_j. \quad (2.16)$$

Note that equations (2.16) for χ_m are elliptic on each sphere $\{|k| = k\}$. This follows from the fact that the equations for each such sphere are all decoupled and non-degeneracy of the matrix D along the sphere discussed above. Also note that the matrix A is positive definite. Indeed, let $\mathbf{c} = (c_1, \dots, c_d) \in \mathbb{R}^d$ be a fixed vector and let $\chi_{\mathbf{c}} := \sum_{m=1}^d c_m \chi_m$. Since the matrix D is non-negative we have

$$\begin{aligned} (A\mathbf{c}, \mathbf{c})_{\mathbb{R}^d} &= -\frac{1}{\Gamma_{d-1}} \sum_{m,n=1}^d \int_{\mathbb{S}^{d-1}} \chi_{\mathbf{c}}(\hat{k}, l) \frac{\partial}{\partial k_m} \left(D_{mn}(\hat{k}, l) \frac{\partial \chi_{\mathbf{c}}(\hat{k}, l)}{\partial k_n} \right) d\Omega(\hat{k}) \\ &= -\frac{1}{\Gamma_{d-1}} \sum_{m,n=1}^d \int_{\mathbb{R}^d} \chi_{\mathbf{c}}(\hat{k}, l) \frac{\partial}{\partial k_m} \left(D_{mn}(\hat{k}, l) \frac{\partial \chi_{\mathbf{c}}(\hat{k}, l)}{\partial k_n} \right) \delta(k - l) \frac{dk}{l^{d-1}} \\ &= \frac{1}{\Gamma_{d-1}} \int_{\mathbb{S}^{d-1}} (D(\hat{k}, l) \nabla \chi_{\mathbf{c}}(\hat{k}, l), \nabla \chi_{\mathbf{c}}(\hat{k}, l))_{\mathbb{R}^d} d\Omega(\hat{k}) \geq 0. \end{aligned} \quad (2.17)$$

The last equality holds after integration by parts because $D(\hat{k}, l) \hat{k} = 0$. Moreover, the inequality appearing in the last line of (2.17) is strict. This can be seen as follows. Since the null-space of the matrix $D(\hat{k}, l)$ is one-dimensional and consists of the vectors parallel to \hat{k} , in order for $(A\mathbf{c}, \mathbf{c})_{\mathbb{R}^d}$ to vanish one needs that the gradient $\nabla \chi_{\mathbf{c}}(\hat{k}, l)$ is parallel to \hat{k} for all $\hat{k} \in \mathbb{S}^{d-1}$. This, however, together with (2.16) would imply that $\hat{k} \cdot \mathbf{c} = 0$ for all \hat{k} , which is impossible.

The following theorem holds.

Theorem 2.2.2 *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 (2.12) converges as $\gamma \rightarrow 0$ in $C([T_*, T]; L^\infty(\mathbb{R}^{2d}))$ to $w(t, x, k)$. Moreover, there exists a constant $C > 0$ so that we have*

$$\|w(t, \cdot) - \bar{\phi}_\gamma(t, \cdot)\|_{0,0} \leq C(\gamma T + \sqrt{\gamma}) \|\phi_0\|_{1,1} \quad (2.18)$$

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

The proof of Theorem 2.2.2 is based on some classical asymptotic expansions and is quite straightforward so we will not give it here. In fact, we have a stronger result, namely T_* can be made to vanish as $\gamma \rightarrow 0$. For instance, we can choose $T_* = \gamma^{3/2}$. We can not set $T_* = 0$ since there is a small initial layer when the solution of (2.12) adjusts to become independent of the direction \hat{k} .

As an immediate corollary of Theorems 2.2.1 and 2.2.2 we obtain the following result, which is the main result of this chapter.

Theorem 2.2.3 Let ϕ_δ be solution of (2.10) with the initial data $\phi_\delta(0, x, k) = \phi_0(\delta^{1+\alpha}x, k)$ and let $\bar{w}(t, x)$ be the solution of the diffusion equation (2.14) with the initial data $w(0, x, k) = \bar{\phi}_0(x, k)$. Then, there exists $\alpha_0 > 0$ and a constant $C > 0$ so that for all $0 \leq \alpha < \alpha_0$ and all $0 < T_* \leq T$ we have for all compact sets $K \subset \mathcal{A}(M)$:

$$\sup_{(t,x,k) \in [T_*, T] \times K} |w(t, x, k) - \mathbb{E}\bar{\phi}_\delta(t, x, k)| \leq CT\delta^{\alpha_0 - \alpha}, \quad (2.19)$$

where $\bar{\phi}_\delta(t, x, k) := \phi_\delta(t/\delta^{1+2\alpha}, x/\delta^{1+\alpha}, k)$.

Theorem 2.2.3 shows that the movement of a particle in a weakly random quenched Hamiltonian is, indeed, approximated by a Brownian motion in the long time-large space limit, at least for times $T \ll \delta^{-\alpha_0}$. In fact, we can allow T_* to vanish as $\delta \rightarrow 0$ choosing $T_* = \delta^{3\alpha/2}$.

The diffusion matrix. In the isotropic case when $R = R(|x|, k)$ we may simplify the above expressions for the diffusion matrices D_{mn} and a_{mn} . In that case we have

$$\begin{aligned} D_{mn}(\hat{k}, k) &= -\frac{1}{2} \int_{-\infty}^{\infty} \frac{\partial^2 R(H'_0(k)s\hat{k}, k)}{\partial x_n \partial x_m} ds \\ &= -\int_0^{\infty} \left[\frac{k_n k_m}{k^2} R''(H'_0(k)s, k) + \left(\delta_{nm} - \frac{k_n k_m}{k^2} \right) \frac{R'(H'_0(k)s, k)}{H'_0(k)s} \right] ds \\ &= -\frac{1}{H'_0(k)} \int_0^{\infty} \frac{R'(s, k)}{s} ds \left(\delta_{nm} - \frac{k_n k_m}{k^2} \right), \end{aligned}$$

so that the matrix $[D_{mn}(\hat{k}, k)]$ has the form

$$D(\hat{k}, k) = D_0(k) \left(I - \hat{k} \otimes \hat{k} \right), \quad D_0(k) = -\frac{1}{H'_0(k)} \int_0^{\infty} \frac{R'(s, k)}{s} ds.$$

The functions χ_j are given explicitly in this case by

$$\chi_j(\hat{k}, k) = -\frac{|H'_0(k)|^2 |k|^2 \hat{k}_j}{(d-1)\bar{D}_0(k)}, \quad \bar{D}_0(k) = -\int_0^{\infty} \frac{R'(s, k)}{s} ds$$

and

$$a_{nm}(k) = \frac{|H'_0(k)|^3 |k|^2}{\Gamma_{d-1}(d-1)\bar{D}_0(k)} \int_{\mathbb{S}^{d-1}} \hat{k}_n \hat{k}_m d\Omega(\hat{k}) = \frac{|H'_0(k)|^3 |k|^2}{d(d-1)\bar{D}_0(k)} \delta_{nm}.$$

2.2.6 A formal derivation of the momentum diffusion

We now recall how the diffusion operator in (2.12) can be derived in a quick formal way. This is very similar to what we did at the end of Section 1.1.3. We represent the solution of (2.10) as $\phi^\delta(t, x, k) = \psi^\delta(\delta t, \delta x, k)$ and write an asymptotic multiple scale expansion for ψ^δ

$$\psi^\delta(t, x, k) = \bar{\phi}(t, x, k) + \sqrt{\delta}\phi_1\left(t, x, \frac{x}{\delta}, k\right) + \delta\phi_2\left(t, x, \frac{x}{\delta}, k\right) + \dots \quad (2.20)$$

We assume formally that the leading order term $\bar{\phi}$ is deterministic and independent of the fast variable $z = x/\delta$. We insert this expansion into (2.10) and obtain in the order $O(\delta^{-1/2})$:

$$\nabla_z H_1(z, k) \cdot \nabla_k \bar{\phi} - H'_0(k) \hat{k} \cdot \nabla_z \phi_1 = 0. \quad (2.21)$$

Let $\theta \ll 1$ be a small positive regularization parameter that will be later sent to zero, and consider a regularized version of (2.21):

$$\frac{1}{H'_0(k)} \nabla_z H_1(z, k) \cdot \nabla_k \bar{\phi} - \hat{k} \cdot \nabla_z \phi_1 + \theta \phi_1 = 0,$$

Its solution is

$$\phi_1(z, k) = -\frac{1}{H'_0(k)} \int_0^\infty \sum_{m=1}^d \frac{\partial H_1(z + s\hat{k}, k)}{\partial z_m} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_m} e^{-\theta s} ds. \quad (2.22)$$

The next order equation becomes upon averaging

$$\frac{\partial \bar{\phi}}{\partial t} = \mathbb{E} \left(\frac{\partial H_1(z, k)}{\partial k} \hat{k} \cdot \nabla_z \phi_1 \right) - \mathbb{E} (\nabla_z H_1(z, k) \cdot \nabla_k \phi_1) + H'_0(k) \hat{k} \cdot \nabla_x \bar{\phi}. \quad (2.23)$$

The first two terms on the right hand side above may be computed explicitly using expression (2.22) for ϕ_1 :

$$\begin{aligned} & \mathbb{E} \left(\frac{\partial H_1(z, k)}{\partial k} \hat{k} \cdot \nabla_z \phi_1 \right) - \mathbb{E} (\nabla_z H_1(z, k) \cdot \nabla_k \phi_1) \\ &= -\mathbb{E} \left[\sum_{m,n=1}^d \frac{\partial H_1(z, k)}{\partial k} \hat{k}_m \frac{\partial}{\partial z_m} \left(\frac{1}{H'_0(k)} \int_0^\infty \frac{\partial H_1(z + s\hat{k}, k)}{\partial z_n} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right) \right] \\ &+ \mathbb{E} \left[\sum_{m,n=1}^d \frac{\partial H_1(z, k)}{\partial z_m} \frac{\partial}{\partial k_m} \left(\frac{1}{H'_0(k)} \int_0^\infty \frac{\partial H_1(z + s\hat{k}, k)}{\partial z_n} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right) \right]. \end{aligned}$$

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

$$\begin{aligned} & -\mathbb{E} \left[\sum_{m,n=1}^d \frac{\partial H_1(z, k)}{\partial k} \hat{k}_m \frac{\partial}{\partial z_m} \left(\frac{1}{H'_0(k)} \int_0^\infty \frac{\partial H_1(z + s\hat{k}, k)}{\partial z_n} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right) \right] \\ & -\mathbb{E} \left[\sum_{m,n=1}^d H_1(z, k) \frac{\partial}{\partial z_m} \frac{\partial}{\partial k_m} \left(\frac{1}{H'_0(k)} \int_0^\infty \frac{\partial H_1(z + s\hat{k}, k)}{\partial z_n} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right) \right] \\ &= -\sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left[\frac{1}{H'_0(k)} \int_0^\infty \mathbb{E} \left(H_1(z, k) \frac{\partial^2 H_1(z + s\hat{k}, k)}{\partial z_n \partial z_m} \right) \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right] \\ &= -\sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left(\frac{1}{H'_0(k)} \int_0^\infty \frac{\partial^2 R(s\hat{k}, k)}{\partial x_n \partial x_m} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} e^{-\theta s} ds \right) \\ &\rightarrow -\frac{1}{2} \sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left(\frac{1}{H'_0(k)} \int_{-\infty}^\infty \frac{\partial^2 R(s\hat{k}, k)}{\partial x_n \partial x_m} \frac{\partial \bar{\phi}(t, x, k)}{\partial k_n} ds \right), \text{ as } \theta \rightarrow 0^+. \end{aligned}$$

We insert the above expression into (2.23) and obtain

$$\frac{\partial \bar{\phi}}{\partial t} = \sum_{m,n=1}^d \frac{\partial}{\partial k_n} \left(D_{nm}(\hat{k}, k) \frac{\partial \bar{\phi}}{\partial k_m} \right) + H'_0(k) \hat{k} \cdot \nabla_x \bar{\phi} \quad (2.24)$$

with the diffusion matrix $D(\hat{k}, k)$ as in (2.11). Observe that (2.24) is nothing but (2.12). However, the naive asymptotic expansion (2.20) may not be justified. The rigorous proof presented in the next section is based on a quite different method.

2.3 From the Liouville equation to the momentum diffusion: estimation of the convergence rates

Outline of the proof

The basic idea of the proof of Theorem 2.2.1 is a modification of that of [6, 40]. We consider the trajectories corresponding to the Liouville equation (2.10) and introduce a stopping time, called τ_δ , that, among others, prevents near self-intersection of trajectories. This fact ensures that until the stopping time occurs the particle is “exploring a new territory” and, thanks to the strong mixing properties of the medium, “memory effects” are lost. Therefore, roughly speaking, until the stopping time the process is approximately characterized by the Markov property. Furthermore, since the amplitude of the random Hamiltonian is not strong enough to destroy the continuity of its path, it becomes a diffusion in the limit, as $\delta \rightarrow 0$. We introduce also an augmented process that follows the trajectories of the Hamiltonian flow until the stopping time τ_δ and becomes a diffusion after $t = \tau_\delta$. We show that the law of the augmented process is close to the law of a diffusion, see Proposition 2.3.3, with an explicit error bound. We also prove that the stopping time tends to infinity as $\delta \rightarrow 0$, once again with the error bound that is proved in Theorem 2.3.5. The combination of these two results allows us to estimate the difference between the solutions of the Liouville and the diffusion equations in a rather straightforward manner (see Section 2.3.2): they are close until the stopping time as the law of the diffusion is always close to that of the augmented process, while the latter coincides with the true process until τ_δ . On the other hand, the fact that $\tau_\delta \rightarrow \infty$ as $\delta \rightarrow 0$ shows that with a large probability the augmented process is close to the true process. This combination finishes the proof.

The random characteristics corresponding to (2.10)

Consider the motion of a particle governed by a Hamiltonian system of equations

$$\begin{cases} \frac{dz^{(\delta)}(t; \mathbf{x}, \mathbf{k})}{dt} = (\nabla_{\mathbf{k}} H_\delta) \left(\frac{z^{(\delta)}(t; \mathbf{x}, \mathbf{k})}{\delta}, \mathbf{m}^{(\delta)}(t; \mathbf{x}, \mathbf{k}) \right) \\ \frac{d\mathbf{m}^{(\delta)}(t; \mathbf{x}, \mathbf{k})}{dt} = -\frac{1}{\sqrt{\delta}} (\nabla_z H_\delta) \left(\frac{z^{(\delta)}(t; \mathbf{x}, \mathbf{k})}{\delta}, \mathbf{m}^{(\delta)}(t; \mathbf{x}, \mathbf{k}) \right) \\ z^{(\delta)}(0; \mathbf{x}, \mathbf{k}) = \mathbf{x}, \quad \mathbf{m}^{(\delta)}(0; \mathbf{x}, \mathbf{k}) = \mathbf{k}, \end{cases} \quad (2.25)$$

where the Hamiltonian $H_\delta(x, k) := H_0(k) + \sqrt{\delta} H_1(x, k)$, $k = |k|$. The trajectories of (2.25) are the characteristics of the Liouville equation (2.10). The hypotheses made in Section 2.2 imply that the trajectory $(z^{(\delta)}(t; \mathbf{x}, k), \mathbf{m}^{(\delta)}(t; \mathbf{x}, k))$ necessarily lies in $\mathcal{C}(T, \delta)$ for each $T > 0$, $\delta \in (0, \delta_*(M)]$, provided that the initial data $(\mathbf{x}, k) \in \mathcal{A}(M)$. Indeed, it follows from the Hamiltonian structure of (2.25) that the Hamiltonian $H_\delta(x, m) = H_0(m) + \sqrt{\delta} H_1(z, m)$ must be conserved along the trajectory. Hence, the definition (2.9) implies that $M_\delta^{-1} \leq |\mathbf{m}^{(\delta)}(\cdot; \mathbf{x}, k)| \leq M_\delta$. We denote by $Q_{s, \mathbf{x}, \mathbf{k}}^\delta(\cdot)$ the law over \mathcal{C} of the process corresponding to (2.25) starting at $t = s$ from (\mathbf{x}, \mathbf{k}) (this law is actually supported in $\mathcal{C}(\delta)$). We shall omit writing the subscript s when it equals to 0.

The stopping times

We now define the stopping time τ_δ , described in Section 2.3, that prevents the trajectories of (2.25) to have near self-intersections (recall that the intent of the stopping time is to prevent

any “memory effects” of the trajectories). As we have already mentioned, we will later show that the probability of the event $[\tau_\delta < T]$ for a fixed $T > 0$ goes to zero, as $\delta \rightarrow 0$.

Let $0 < \epsilon_1 < \epsilon_2 < 1/2$, $\epsilon_3 \in (0, 1/2 - \epsilon_2)$, $\epsilon_4 \in (1/2, 1 - \epsilon_1 - \epsilon_2)$ be small positive constants that will be further determined later and set

$$N = [\delta^{-\epsilon_1}], \quad p = [\delta^{-\epsilon_2}], \quad q = p[\delta^{-\epsilon_3}], \quad N_1 = Np[\delta^{-\epsilon_4}]. \quad (2.26)$$

We will specify additional restrictions on the constants ϵ_j as the need for such constraints arises. However, the basic requirement is that ϵ_i , $i = 1, 2, 3$ should be sufficiently small and ϵ_4 is bigger than $1/2$, less than one and can be made as close to one as we would need it. It is important that $\epsilon_1 < \epsilon_2$ so that $N \ll p$ when $\delta \ll 1$. We introduce the following $(\mathcal{M}^t)_{t \geq 0}$ -stopping times. Let $t_k^{(p)} := kp^{-1}$ be a mesh of times, and $\pi \in \mathcal{C}$ be a path. We define the “violent turn” stopping time

$$S_\delta(\pi) := \inf \left[t \geq 0 : \text{for some } k \geq 0 \text{ we have } t \in [t_k^{(p)}, t_{k+1}^{(p)}) \text{ and} \right. \\ \left. \hat{K}(t_{k-1}^{(p)}) \cdot \hat{K}(t) \leq 1 - \frac{1}{N}, \text{ or } \hat{K}\left(t_k^{(p)} - \frac{1}{N_1}\right) \cdot \hat{K}(t) \leq 1 - \frac{1}{N} \right], \quad (2.27)$$

where by convention we set $\hat{K}(-1/p) := \hat{K}(0)$. Note that with the above choice of ϵ_4 we have $\hat{K}\left(t_k^{(p)} - 1/N_1\right) \cdot \hat{K}(t_k^{(p)}) > 1 - 1/N$, provided that $\delta \in (0, \delta_0]$ and δ_0 is sufficiently small. We adopt in (2.27) a customary convention that the infimum of an empty set equals $+\infty$. The stopping time S_δ is triggered when the trajectory performs a sudden turn – this is undesirable as the trajectory may then return back to the region it has already visited and create correlations with the past.

For each $t \geq 0$, we denote by $\mathfrak{X}_t(\pi) := \bigcup_{0 \leq s \leq t} X(s; \pi)$ the trace of the spatial component of the path π up to time t , and by $\mathfrak{X}_t(q; \pi) := [x : \text{dist}(x, \mathfrak{X}_t(\pi)) \leq 1/q]$ a tubular region around the path. We introduce the stopping time

$$U_\delta(\pi) := \inf \left[t \geq 0 : \exists k \geq 1 \text{ and } t \in [t_k^{(p)}, t_{k+1}^{(p)}) \text{ for which } X(t) \in \mathfrak{X}_{t_{k-1}^{(p)}}(q) \right]. \quad (2.28)$$

It is associated with the return of the X component of the trajectory to the tube around its past – this is again an undesirable way to create correlations with the past. Finally, we set the stopping time

$$\tau_\delta(\pi) := S_\delta(\pi) \wedge U_\delta(\pi). \quad (2.29)$$

The cut-off functions and the corresponding dynamics

Let $M > 0$ be fixed and p, q, N, N_1 be the positive integers defined above. We define now several auxiliary functions that will be used to introduce the cut-offs in the dynamics. These cut-offs will ensure that the particle moving under the modified dynamics will avoid self-intersections, will have no violent turns and the changes of its momentum will be under control. In addition, up to the stopping time τ_δ the motion of the particle will coincide with the motion under the original Hamiltonian flow.

Let $a_1 = 2$ and $a_2 = 3/2$. The functions $\psi_j : \mathbb{R}^d \times \mathbb{S}_1^{d-1} \rightarrow [0, 1]$, $j = 1, 2$ are of C^∞ class and satisfy

$$\psi_j(\mathbf{k}, \mathbf{l}) = \begin{cases} 1, & \text{if } \hat{\mathbf{k}} \cdot \mathbf{l} \geq 1 - 1/N \quad \text{and} \quad M_\delta^{-1} \leq |\mathbf{k}| \leq M_\delta \\ 0, & \text{if } \hat{\mathbf{k}} \cdot \mathbf{l} \leq 1 - a_j/N, \quad \text{or} \quad |\mathbf{k}| \leq (2M_\delta)^{-1}, \quad \text{or} \quad |\mathbf{k}| \geq 2M_\delta. \end{cases} \quad (2.30)$$

One can construct ψ_j in such a way that for arbitrary nonnegative integers m, n it is possible to find a constant $C_{m,n}$ for which $\|\psi_j\|_{m,n} \leq C_{m,n} N^{m+n}$. The cut-off function

$$\Psi(t, \mathbf{k}; \pi) := \begin{cases} \psi_1\left(\mathbf{k}, \hat{K}\left(t_{k-1}^{(p)}\right)\right) \psi_2\left(\mathbf{k}, \hat{K}\left(t_k^{(p)} - 1/N_1\right)\right) & \text{for } t \in [t_k^{(p)}, t_{k+1}^{(p)}) \text{ and } k \geq 1 \\ \psi_2(\mathbf{k}, \hat{K}(0)) & \text{for } t \in [0, t_1^{(p)}) \end{cases} \quad (2.31)$$

will allow us to control the direction of the particle motion over each interval of the partition as well as not to allow the trajectory to escape to the regions where the change of the size of the velocity can be uncontrollable.

Let $\phi : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, 1]$ be a function of the C^∞ class that satisfies $\phi(\mathbf{y}, \mathbf{x}) = 1$, when $|\mathbf{y} - \mathbf{x}| \geq 3/q$ and $\phi(\mathbf{y}, \mathbf{x}) = 0$, when $|\mathbf{y} - \mathbf{x}| \leq 2/q$. Again, in this case we can construct ϕ in such a way that $\|\phi\|_{m,n} \leq Cq^{m+n}$ for arbitrary integers m, n and a suitably chosen constant C . The function $\phi_k : \mathbb{R}^d \times \mathcal{C} \rightarrow [0, 1]$ for a fixed path π is given by

$$\phi_k(\mathbf{y}; \pi) = \prod_{0 \leq l/q \leq t_{k-1}^{(p)}} \phi\left(\mathbf{y}, X\left(\frac{l}{q}\right)\right). \quad (2.32)$$

We set

$$\Phi(t, \mathbf{y}; \pi) := \begin{cases} 1, & \text{if } 0 \leq t < t_1^{(p)} \\ \phi_k(\mathbf{y}; \pi), & \text{if } t_k^{(p)} \leq t < t_{k+1}^{(p)}. \end{cases} \quad (2.33)$$

The function Φ shall be used to modify the dynamics of the particle in order to avoid a possibility of near self-intersections of its trajectory.

For a given $t \geq 0$, $(\mathbf{y}, \mathbf{k}) \in \mathbb{R}_*^{2d}$ and $\pi \in \mathcal{C}$ let us denote $\Theta(t, \mathbf{y}, \mathbf{k}; \pi) := \Psi(t, \mathbf{k}; \pi) \Phi(t, \mathbf{y}; \pi)$. Then we have

$$|\partial_{\mathbf{y}}^{\beta_1} \partial_{\mathbf{k}}^{\beta_2} \Theta(t, \mathbf{y}, \mathbf{k}; \pi)| \leq CT^{|\beta_1|} q^{2|\beta_1|} N^{|\beta_2|} \text{ for all } t \in [0, T], (\mathbf{y}, \mathbf{k}) \in \mathcal{A}(2M), \pi \in \mathcal{C},$$

with a constant C depending only on $m = |\beta_1| + |\beta_2|$ and M .

Finally, let us set

$$F_\delta(t, \mathbf{y}, \mathbf{l}; \pi, \omega) = \Theta(t, \delta \mathbf{y}, \mathbf{l}; \pi) \nabla_{\mathbf{y}} H_1(\mathbf{y}, |\mathbf{l}|; \omega). \quad (2.34)$$

For a fixed $(\mathbf{x}, \mathbf{k}) \in \mathbb{R}_*^{2d}$, $\delta > 0$ and $\omega \in \Omega$ we consider the modified particle dynamics with the cut-off that is described by the stochastic process $(y^{(\delta)}(t; \mathbf{x}, \mathbf{k}, \omega), \mathbf{l}^{(\delta)}(t; \mathbf{x}, \mathbf{k}, \omega))_{t \geq 0}$ whose paths are the solutions of the following equation

$$\begin{cases} \frac{dy^{(\delta)}(t; \mathbf{x}, \mathbf{k})}{dt} = \left[H'_0(|\mathbf{l}^{(\delta)}(t; \mathbf{x}, \mathbf{k})|) + \sqrt{\delta} \partial_l H_1\left(\frac{y^{(\delta)}(t; \mathbf{x}, \mathbf{k})}{\delta}, |\mathbf{l}^{(\delta)}(t; \mathbf{x}, \mathbf{k})|\right) \right] \hat{\mathbf{l}}^{(\delta)}(t; \mathbf{x}, \mathbf{k},) \\ \frac{d\mathbf{l}^{(\delta)}(t; \mathbf{x}, \mathbf{k})}{dt} = -\frac{1}{\sqrt{\delta}} F_\delta\left(t, \frac{y^{(\delta)}(t; \mathbf{x}, \mathbf{k})}{\delta}, \mathbf{l}^{(\delta)}(t; \mathbf{x}, \mathbf{k}); y^{(\delta)}(\cdot; \mathbf{x}, \mathbf{k}), \mathbf{l}^{(\delta)}(\cdot; \mathbf{x}, \mathbf{k})\right) \\ y^{(\delta)}(0; \mathbf{x}, \mathbf{k}) = \mathbf{x}, \quad \mathbf{l}^{(\delta)}(0; \mathbf{x}, \mathbf{k}) = \mathbf{k}. \end{cases} \quad (2.35)$$

We will denote by $\tilde{Q}_{\mathbf{x}, \mathbf{k}}^{(\delta)}$ the law of the modified process $(y^{(\delta)}(\cdot; \mathbf{x}, \mathbf{k}), \mathbf{l}^{(\delta)}(\cdot; \mathbf{x}, \mathbf{k}))$ over \mathcal{C} for a given $\delta > 0$ and by $\tilde{E}_{\mathbf{x}, \mathbf{k}}^{(\delta)}$ the corresponding expectation. We assume that the initial momentum $\mathbf{k} \in A(M)$. From the construction of the cut-offs we immediately conclude that

$$\hat{\mathbf{l}}^{(\delta)}(t) \cdot \hat{\mathbf{l}}^{(\delta)}(t_{k-1}^{(p)}) \geq 1 - \frac{2}{N}, \quad t \in [t_{k-1}^{(p)}, t_{k+1}^{(p)}), \quad \forall k \geq 0. \quad (2.36)$$

Some consequences of the mixing assumption

For any $t \geq 0$ we denote by \mathcal{F}_t the σ -algebra generated by $(y^{(\delta)}(s), \mathbf{l}^{(\delta)}(s))$, $s \leq t$. Here we suppress, for the sake of abbreviation, writing the initial data in the notation of the trajectory. In this section we assume that $M > 0$ is fixed, $X_1, X_2 : (\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d^2})^2 \rightarrow \mathbb{R}$ are certain continuous functions, Z is a random variable and g_1, g_2 are $\mathbb{R}^d \times [M^{-1}, M]$ -valued random vectors. We suppose further that Z, g_1, g_2 , are \mathcal{F}_t -measurable, while \tilde{X}_1, \tilde{X}_2 are random fields of the form

$$\tilde{X}_i(\mathbf{x}, k) = X_i \left(\left(\partial_k^j H_1(\mathbf{x}, k), \nabla_{\mathbf{x}} \partial_k^j H_1(\mathbf{x}, k), \nabla_{\mathbf{x}}^2 \partial_k^j H_1(\mathbf{x}, k) \right)_{j=0,1} \right).$$

For $i = 1, 2$ we denote $g_i := (g_i^{(1)}, g_i^{(2)})$ where $g_i^{(1)} \in \mathbb{R}^d$ and $g_i^{(2)} \in [M^{-1}, M]$. We also let

$$U(\theta_1, \theta_2) := \mathbb{E} \left[\tilde{X}_1(\theta_1) \tilde{X}_2(\theta_2) \right], \quad \theta_1, \theta_2 \in \mathbb{R}^d \times [M^{-1}, M]. \quad (2.37)$$

The following mixing lemma is useful in formalizing the “memory loss effect” and can be proved in the same way as Lemmas 5.2 and 5.3 of [6]. It is also similar in spirit to Lemma 1.3.2.

Lemma 2.3.1 (i) Assume that $r, t \geq 0$ and

$$\inf_{u \leq t} \left| g_i^{(1)} - \frac{y^{(\delta)}(u)}{\delta} \right| \geq \frac{r}{\delta}, \quad (2.38)$$

\mathbb{P} -a.s. on the set $Z \neq 0$ for $i = 1, 2$. Then, we have

$$\left| \mathbb{E} \left[\tilde{X}_1(g_1) \tilde{X}_2(g_2) Z \right] - \mathbb{E} [U(g_1, g_2) Z] \right| \leq 2\phi \left(\frac{r}{2\delta} \right) \|X_1\|_{L^\infty} \|X_2\|_{L^\infty} \|Z\|_{L^1(\Omega)}. \quad (2.39)$$

(ii) Let $\mathbb{E} X_1(\mathbf{0}, k) = 0$ for all $k \in [M^{-1}, M]$. Furthermore, we assume that g_2 satisfies (2.38),

$$\inf_{u \leq t} \left| g_1^{(1)} - \frac{y^{(\delta)}(u)}{\delta} \right| \geq \frac{r + r_1}{\delta} \quad (2.40)$$

and $|g_1^{(1)} - g_2^{(1)}| \geq r_1 \delta^{-1}$ for some $r_1 \geq 0$, \mathbb{P} -a.s. on the event $Z \neq 0$. Then, we have

$$\left| \mathbb{E} \left[\tilde{X}_1(g_1) \tilde{X}_2(g_2) Z \right] - \mathbb{E} [U(g_1, g_2) Z] \right| \leq C \phi^{1/2} \left(\frac{r}{2\delta} \right) \phi^{1/2} \left(\frac{r_1}{2\delta} \right) \|X_1\|_{L^\infty} \|X_2\|_{L^\infty} \|Z\|_{L^1(\Omega)} \quad (2.41)$$

for some absolute constant $C > 0$. Here the function U is given by (2.37).

The momentum diffusion

Let $\mathbf{k}(t)$ be a diffusion, starting at $k \in \mathbb{R}_*^d$ at $t = 0$, with the generator of the form

$$\begin{aligned} \mathcal{L}F(\mathbf{k}) &= \sum_{m,n=1}^d D_{mn}(\hat{\mathbf{k}}, |k|) \partial_{k_m, k_n}^2 F(\mathbf{k}) + \sum_{m=1}^d E_m(\hat{\mathbf{k}}, |k|) \partial_{k_m} F(\mathbf{k}) \\ &= \sum_{m,n=1}^d \partial_{k_m} \left(D_{m,n}(\hat{\mathbf{k}}, |k|) \partial_{k_n} F(\mathbf{k}) \right), \quad F \in C_0^\infty(\mathbb{R}_*^d). \end{aligned} \quad (2.42)$$

Here the diffusion matrix is given by (2.11) and the drift vector is

$$E_m(\hat{\mathbf{k}}, l) = -\frac{1}{H'_0(l)l} \sum_{n=1}^d \int_0^{+\infty} s \frac{\partial^3 R(s\hat{\mathbf{k}}, l)}{\partial x_m \partial x_n^2} ds, \quad m = 1, \dots, d.$$

It can be easily seen that this diffusion is supported on \mathbb{S}_k^{d-1} , where $k = |\mathbf{k}|$. Moreover, it is non-degenerate on the sphere, for instance, under the assumption (2.8).

Let $\mathfrak{Q}_{\mathbf{x}, \mathbf{k}}$ be the law of the process $(\mathbf{x}(t), \mathbf{k}(t))$ that starts at $t = 0$ from (\mathbf{x}, \mathbf{k}) given by $\mathbf{x}(t) = x + \int_0^t H'_0(|\mathbf{k}(s)|) \hat{\mathbf{k}}(s) ds$, where $\mathbf{k}(t)$ is the diffusion described by (2.42). This process is a degenerate diffusion whose generator is given by

$$\tilde{\mathcal{L}}F(\mathbf{x}, \mathbf{k}) = \mathcal{L}_k F(\mathbf{x}, \mathbf{k}) + H'_0(|\mathbf{k}|) \hat{\mathbf{k}} \cdot \nabla_x F(\mathbf{x}, \mathbf{k}), \quad F \in C_0^\infty(\mathbb{R}_*^{2d}). \quad (2.43)$$

Here the notation \mathcal{L}_k stresses that the operator \mathcal{L} defined in (2.42) acts on the respective function in the k variable. We denote by $\mathfrak{M}_{x, k}$ the expectation corresponding to the path measure $\mathfrak{Q}_{\mathbf{x}, \mathbf{k}}$.

The augmented process

The following construction of the augmentation of path measures has been carried out in Section 6.1 of [58]. Let $s \geq 0$ be fixed and $\pi \in \mathcal{C}$. Then, according to Lemma 6.1.1 of *ibid.* there exists a unique probability measure, that is denoted by $\delta_\pi \otimes_s \mathfrak{Q}_{X(s), K(s)}$, such that for any pair of events $A \in \mathcal{M}^s$, $B \in \mathcal{M}$ we have $\delta_\pi \otimes_s \mathfrak{Q}_{X(s), K(s)}[A] = \mathbf{1}_A(\pi)$ and $\delta_\pi \otimes_s \mathfrak{Q}_{X(s), K(s)}[\theta_s(B)] = \mathfrak{Q}_{X(s), K(s)}[B]$. The following result is a direct consequence of Theorem 6.2.1 of [58].

Proposition 2.3.2 *There exists a unique probability measure $R_{x, k}^{(\delta)}$ on \mathcal{C} such that $R_{x, k}^{(\delta)}[A] := Q_{x, k}^{(\delta)}[A]$ for all $A \in \mathcal{M}^{\tau_\delta}$ and the regular conditional probability distribution of $R_{x, k}^{(\delta)}[\cdot | \mathcal{M}^{\tau_\delta}]$ is given by $\delta_\pi \otimes_{\tau_\delta(\pi)} \mathfrak{Q}_{X(\tau_\delta(\pi)), K(\tau_\delta(\pi))}$, $\pi \in \mathcal{C}$. This measure shall be also denoted by $Q_{x, k}^{(\delta)} \otimes_{\tau_\delta} \mathfrak{Q}_{X(\tau_\delta), K(\tau_\delta)}$.*

Note that for any $(\mathbf{x}, \mathbf{k}) \in \mathcal{A}(M)$ and $A \in \mathcal{M}^{\tau_\delta}$ we have

$$R_{\mathbf{x}, \mathbf{k}}^{(\delta)}[A] = Q_{\mathbf{x}, \mathbf{k}}^{(\delta)}[A] = \tilde{Q}_{\mathbf{x}, \mathbf{k}}^{(\delta)}[A], \quad (2.44)$$

that is, the law of the augmented process coincides with that of the true process, and of the modified process with the cut-offs until the stopping time τ_δ . Hence, according to the uniqueness part of Proposition 2.3.2, in such a case $Q_{x, k}^{(\delta)} \otimes_{\tau_\delta} \mathfrak{Q}_{X(\tau_\delta), K(\tau_\delta)} = \tilde{Q}_{x, k}^{(\delta)} \otimes_{\tau_\delta} \mathfrak{Q}_{X(\tau_\delta), K(\tau_\delta)}$. We denote by $E_{x, k}^{(\delta)}$ the expectation with respect to the augmented measure described by the above proposition. Let also $R_{x, k, \pi}^{(\delta)}$, $E_{x, k, \pi}^{(\delta)}$ denote the respective conditional law and expectation obtained by conditioning $R_{x, k}^{(\delta)}$ on $\mathcal{M}^{\tau_\delta}$.

The following proposition is of crucial importance for us, as it shows that the law of the augmented process is close to that of the momentum diffusion as $\delta \rightarrow 0$. To abbreviate the notation we let

$$N_t(G) := G(t, X(t), K(t)) - G(0, X(0), K(0)) - \int_0^t (\partial_\varrho + \tilde{\mathcal{L}})G(\varrho, X(\varrho), K(\varrho)) d\varrho$$

for any $G \in C_b^{1,1,3}([0, +\infty) \times \mathbb{R}_*^{2d})$ and $t \geq 0$.

Proposition 2.3.3 Suppose that $(x, k) \in \mathcal{A}(M)$ and $\zeta \in C_b((\mathbb{R}_*^{2d})^n)$ is nonnegative. Let $\gamma_0 \in (0, 1/2)$ and let $0 \leq t_1 < \dots < t_n \leq T_* \leq t < v \leq T$. We assume further that $v - t \geq \delta^{\gamma_0}$. Then, there exist constants γ_1, C such that for any function $G \in C^{1,1,3}([T_*, T] \times \mathbb{R}_*^{2d})$ we have

$$\left| E_{x,k}^{(\delta)} \left\{ [N_v(G) - N_t(G)] \tilde{\zeta} \right\} \right| \leq C \delta^{\gamma_1} (v - t) \|G\|_{1,1,3}^{[T_*, T]} T^2 E_{x,k}^{(\delta)} \tilde{\zeta}. \quad (2.45)$$

Here $\tilde{\zeta}(\pi) := \zeta(X(t_1), K(t_1), \dots, X(t_n), K(t_n))$, $\pi \in \mathcal{C}(T, \delta)$. The choice of the constants γ_1, C does not depend on (x, k) , $\delta \in (0, 1]$, ζ , times $t_1, \dots, t_n, T_*, T, v, t$, or the function G .

Proof. Let $0 = s_0 \leq s_1 \leq \dots \leq s_n \leq t$ and $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R}_*^{2d})$ be Borel sets. We denote $A_0 := \mathcal{C}$ and for any $k \in \{1, \dots, n\}$, $s \leq s_k$ we define the events

$$A_k := [\pi : (X(s_1), K(s_1)) \in B_1, \dots, (X(s_k), K(s_k)) \in B_k].$$

We will show that

$$\left| E_{x,k}^{(\delta)} [N_v(G) - N_t(G), A_n] \right| \leq C \delta^{\gamma_1} \|G\|_{1,1,3}^{[T_*, T]} T^2 R_{x,k}^{(\delta)} [A_n] (v - t) \quad (2.46)$$

for times $v - t \geq \delta^{\gamma_0}$, which is equivalent to (2.45). It is convenient to define the shifted “future” counterparts of the sets A_n :

$$A_k^{(s)} := [\pi : (X(s_k - s), K(s_k - s)) \in B_k, \dots, (X(s_n - s), K(s_n - s)) \in B_n].$$

For $(\mathbf{x}, \mathbf{k}) \in \mathbb{R}_*^{2d}$, $\pi \in \mathcal{C}$ and $G \in C^{1,1,2}([0, +\infty) \times \mathbb{R}_*^{2d})$ we let

$$\begin{aligned} \hat{\mathcal{L}}_t G(t, x, \mathbf{k}; \pi) &:= H'_0(|k|) \hat{\mathbf{k}} \cdot \nabla_x G(t, \mathbf{x}, \mathbf{k}) + \Theta^2(t, X(t), K(t); \pi) \mathcal{L}_k G(t, x, \mathbf{k}) \\ &\quad - \Theta(t, X(t), K(t); \pi) \sum_{m,n=1}^d \partial_{K_m} \Theta(t, X(t), K(t); \pi) D_{m,n}(\hat{\mathbf{k}}, |k|) \partial_{k_n} G(t, x, \mathbf{k}) \end{aligned}$$

and

$$\hat{N}_t(G) := G(t, X(t), K(t)) - G(0, X(0), K(0)) - \int_0^t (\partial_\varrho + \hat{\mathcal{L}}_\varrho) G(\varrho, X(\varrho), K(\varrho); \pi) d\varrho.$$

Note that unlike \mathcal{L} , $\hat{\mathcal{L}}$ is not a generator of a diffusion and is not local in time – its coefficients depend on the whole random path via the dependence of the cut-off Θ on the whole path π up to the time t . It follows, however, from the definition of the stopping time $\tau_\delta(\pi)$ and the cut-off function Θ that

$$\nabla_K \Theta(t, X(t), K(t); \pi) = \mathbf{0}, \quad t \in [0, \tau_\delta(\pi)],$$

hence

$$\hat{\mathcal{L}}_t G(t, X(t), K(t); \pi) = \tilde{\mathcal{L}} G(t, X(t), K(t); \pi), \quad t \in [0, \tau_\delta(\pi)].$$

We need the following result.

Lemma 2.3.4 Suppose that $(x, k) \in \mathcal{A}(M)$ and $\zeta \in C_b((\mathbb{R}_*^{2d})^n)$ is nonnegative. Let $\gamma'_0 \in (0, 1)$, $0 \leq t_1 < \dots < t_n \leq T_* \leq t < v \leq T$ and $t - T_* \geq \delta^{\gamma'_0}$. Then, there exist constants $\gamma'_1, C' > 0$ such that for any function $G \in C^{1,1,3}([T_*, T] \times \mathbb{R}_*^{2d})$ we have

$$\left| \tilde{E}_{x,k}^{(\delta)} \left\{ [\hat{N}_v(G) - \hat{N}_t(G)] \tilde{\zeta} \right\} \right| \leq C' \delta^{\gamma'_1} (v - t) \|G\|_{1,1,3}^{[T_*, T]} T^2 \tilde{E}_{x,k}^{(\delta)} \tilde{\zeta}. \quad (2.47)$$

The choice of the constants γ'_1, C' does not depend on (x, k) , $\delta \in (0, 1]$, times $t_1, \dots, t_n, T_*, T, v, t$, or function G .

The proof of this lemma is quite long and technical so we leave it out of the notes. However, the basic idea is not that different from what we did in the proof of Theorem 1.3.1 – one has to make a “not so small” partition in time, use the dynamics of the particle on each time interval, make a Taylor expansion of the second order until the number of integrations make the corrections small, and use mixing lemmas to kill the apparently large terms. The details are quite involved but this is the blueprint. The reason for the introduction of the augmented process is seen here, as the difference between Proposition 2.3.3 and Lemma 2.3.4: the augmented process satisfies approximately a martingale problem with the “conventional” generator \mathcal{L} while the process with the cut-off has this property with a much less convenient object $\widehat{\mathcal{L}}$.

In the meantime we apply this result to conclude the proof of Proposition 2.3.3. We write

$$\begin{aligned} E_{x,\mathbf{k},\pi}^{(\delta)}[N_v(G) - N_{v \wedge \tau_\delta(\pi)}(G), A_n] \\ = \sum_{p=0}^{n-1} \mathbf{1}_{[s_p, s_{p+1})}(\tau_\delta(\pi)) \mathbf{1}_{A_p}(\pi) \mathfrak{M}_{X(\tau_\delta(\pi)), K(\tau_\delta(\pi))}[N_{v-\tau_\delta(\pi)}(G), A_{p+1}^{(\tau_\delta(\pi))}] \\ + \mathbf{1}_{[s_n, v)}(\tau_\delta(\pi)) \mathbf{1}_{A_n}(\pi) \mathfrak{M}_{X(\tau_\delta(\pi)), K(\tau_\delta(\pi))}[N_{v-\tau_\delta(\pi)}(G)], \end{aligned} \quad (2.48)$$

simply because $N_v(G) - N_{v \wedge \tau_\delta(\pi)}(G) = 0$ if $v \leq \tau_\delta(\pi)$. When $\tau_\delta(\pi) \in [s_p, s_{p+1})$ we obviously have (recall that $v \geq t \geq s_n$ and N is an \mathfrak{M} -martingale)

$$\mathfrak{M}_{X(\tau_\delta(\pi)), K(\tau_\delta(\pi))}[N_{v-\tau_\delta(\pi)}(G), A_{p+1}^{(\tau_\delta(\pi))}] = \mathfrak{M}_{X(\tau_\delta(\pi)), K(\tau_\delta(\pi))}[N_{t-\tau_\delta(\pi)}(G), A_{p+1}^{(\tau_\delta(\pi))}]$$

and $\mathfrak{M}_{X(\tau_\delta(\pi)), K(\tau_\delta(\pi))}[N_{v-\tau_\delta(\pi)}(G)] = 0$. Hence the left hand side of (2.48) equals

$$\sum_{p=0}^{n-1} \mathbf{1}_{[s_p, s_{p+1})}(\tau_\delta(\pi)) \mathbf{1}_{A_p}(\pi) \mathfrak{M}_{X(\tau_\delta(\pi)), K(\tau_\delta(\pi))}[N_{t-\tau_\delta(\pi)}(G), A_{p+1}^{(\tau_\delta(\pi))}],$$

which is nothing but (reading (2.48) backward with v replaced by t)

$$E_{x,\mathbf{k},\pi}^{(\delta)}[N_t(G) - N_{t \wedge \tau_\delta(\pi)}(G), A_n]$$

so that

$$E_{x,\mathbf{k},\pi}^{(\delta)}[N_v(G) - N_{v \wedge \tau_\delta(\pi)}(G), A_n] = E_{x,\mathbf{k},\pi}^{(\delta)}[N_t(G) - N_{t \wedge \tau_\delta(\pi)}(G), A_n] \quad (2.49)$$

We conclude from (2.49) that

$$\begin{aligned} E_{x,\mathbf{k},\pi}^{(\delta)}[N_v(G), A_n] &= E_{x,\mathbf{k},\pi}^{(\delta)}[N_{v \wedge \tau_\delta(\pi)}(G) + N_t(G) - N_{t \wedge \tau_\delta(\pi)}(G), A_n] \\ &= E_{x,\mathbf{k},\pi}^{(\delta)}[N_{(v \wedge \tau_\delta(\pi)) \vee t}(G), A_n] \end{aligned} \quad (2.50)$$

and therefore

$$\begin{aligned} E_{x,\mathbf{k}}^{(\delta)}[N_v(G), A_n] &= E_{x,\mathbf{k}}^{(\delta)} \left[E_{x,\mathbf{k},\pi}^{(\delta)}[N_{(v \wedge \tau_\delta(\pi)) \vee t}(G), A_n] \right] \\ &= E_{x,\mathbf{k}}^{(\delta)} \left[E_{x,\mathbf{k},\pi}^{(\delta)}[N_{(v \wedge \tau_\delta(\pi)) \vee t}(G), A_n], \tau_\delta(\pi) \leq t \right] \\ &\quad + E_{x,\mathbf{k}}^{(\delta)} \left[E_{x,\mathbf{k},\pi}^{(\delta)}[N_{(v \wedge \tau_\delta(\pi)) \vee t}(G), A_n], \tau_\delta(\pi) > t \right]. \end{aligned} \quad (2.51)$$

The first term on the utmost right hand side of (2.51) equals $E_{x,\mathbf{k}}^{(\delta)}[N_t(G), A_n, \tau_\delta \leq t]$, while the second one equals $\tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{(v \wedge \tau_\delta) \vee t}(G), B]$, where $B := A_n \cap [\tau_\delta > t]$ is an \mathcal{M}_t -measurable event:

$$E_{x,\mathbf{k}}^{(\delta)}[N_v(G), A_n] = E_{x,\mathbf{k}}^{(\delta)}[N_t(G), A_n, \tau_\delta \leq t] + \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{(v \wedge \tau_\delta) \vee t}(G), B]. \quad (2.52)$$

Therefore, we see that the difference we have to estimate is

$$\begin{aligned} \left| E_{x,\mathbf{k}}^{(\delta)}[N_v(G), A_n] - E_{x,\mathbf{k}}^{(\delta)}[N_t(G), A_n] \right| &= \left| E_{x,\mathbf{k}}^{(\delta)}[N_t(G), A_n, \tau_\delta \leq t] + \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{(v \wedge \tau_\delta) \vee t}(G), B] \right. \\ &\quad \left. - E_{x,\mathbf{k}}^{(\delta)}[N_t(G), A_n, \tau_\delta \leq t] - \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_t(G), B] \right| = \left| \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{(v \wedge \tau_\delta) \vee t}(G), B] - \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_t(G), B] \right|. \end{aligned} \quad (2.53)$$

In order to estimate the difference on the right we create yet another mesh with the time-step $1/L$, where $L := \lceil \delta^{-\gamma'_0} \rceil$ and $\gamma'_0 \in (\gamma_0 + 1/2, 1)$. We take a point σ on this mesh close to the time $(v \wedge \tau_\delta) \vee t$

$$\sigma := L^{-1}[(\lceil L(v \wedge \tau_\delta) \rceil + 2) \vee (\lceil Lt \rceil + 2)]$$

and note that

$$\tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_\sigma(G), B] = \sum_{p=\lceil Lt \rceil + 2}^{\lceil Lv \rceil + 2} \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{p/L}(G), B, \sigma = \frac{p}{L}]. \quad (2.54)$$

The reason for adding two in the definition of σ will become clear soon. Representing the event $[\sigma = p/L]$ as the difference of $[\sigma \geq p/L]$ and $[\sigma \geq (p+1)/L]$ (note that $[\sigma \geq (\lceil Lv \rceil + 3)/L] = \emptyset$) and grouping the terms of the sum that correspond to the same index p we obtain that the right hand side of (2.54) equals

$$\tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{(\lceil Lt \rceil + 2)/L}(G), B] + \sum_{p=\lceil Lt \rceil + 2}^{\lceil Lv \rceil + 2} \tilde{E}_{x,\mathbf{k}}^{(\delta)}\left[N_{p+1/L}(G) - N_{p/L}(G), B, \sigma \geq \frac{p+1}{L}\right]. \quad (2.55)$$

Since the event $B \cap [\sigma \geq (p+1)/L]$ is $\mathcal{M}^{(p-1)/L}$ -measurable, from Lemma 2.3.4 we extract a gap $1/L \geq \delta^{\gamma'_0}$ between (in the notation of this Lemma) “ $T_* = (p-1)/L$ ” and “ $t = p/L$ ”, apply this lemma with “ $t = p/L$ ” and “ $v = (p+1)/L$ ”, and conclude that the absolute value of each term appearing under the summation sign in (2.55) can be estimated by $C' \|G\|_{1,1,3} \delta^{\gamma'_1} L^{-1} T^2 \tilde{Q}_{x,\mathbf{k}}^{(\delta)}[B]$. This implies that we may “drop down” from the time σ to the time t :

$$\left| \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_\sigma(G), B] - \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{(\lceil Lt \rceil + 2)/L}(G), B] \right| \leq C' \delta^{\gamma'_1} \|G\|_{1,1,3}^{[T_*, T]} T^2 \tilde{Q}_{x,\mathbf{k}}^{(\delta)}[B] \frac{\lceil Lv \rceil + 1 - \lceil Lt \rceil}{L}.$$

A direct calculation using formulas (2.25) allows us to conclude also that both

$$|N_\sigma(G) - N_{(v \wedge \tau_\delta) \vee t}(G)| \leq C \|G\|_{1,1,3}^{[T_*, T]} \delta^{\gamma'_0 - 1/2}$$

and

$$|N_{(\lceil Lt \rceil + 2)L^{-1}}(G) - N_t(G)| \leq C \|G\|_{1,1,3}^{[T_*, T]} \delta^{\gamma'_0 - 1/2}.$$

This is because these pairs of points lie within $2/L$ of each other. Hence, (since $\gamma'_0 > 1/2 + \gamma_0$)

$$\begin{aligned} \left| \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{(v \wedge \tau_\delta) \vee t}(G), B] - \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_t(G), B] \right| &\leq \left| \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_\sigma(G) - N_{(v \wedge \tau_\delta) \vee t}(G), B] \right| \\ &+ \left| \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_\sigma(G), B] - \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{(\lceil Lt \rceil + 2)L^{-1}}(G), B] \right| + \left| \tilde{E}_{x,\mathbf{k}}^{(\delta)}[N_{(\lceil Lt \rceil + 2)L^{-1}}(G) - N_t(G), B] \right| \\ &\leq C \delta^{\gamma_1} \|G\|_{1,1,3}^{[T_*, T]} T^2 \tilde{Q}_{x,\mathbf{k}}^{(\delta)}[B] [(v - t) \vee \delta^{\gamma_0}] \end{aligned} \quad (2.56)$$

for a certain constant $C > 0$ and $\gamma_1 := \min[\gamma'_0 - \gamma_0 - 1/2, \gamma'_1]$.

Comining (2.53) and (2.56) we obtain

$$\left| E_{x,\mathbf{k}}^{(\delta)}[N_v(G) - N_t(G), A_n] \right| \leq C \delta^{\gamma_1} \|G\|_{1,1,3}^{[T_*, T]} T^2 R_{x,\mathbf{k}}^{(\delta)}[A_n] [(v - t) \vee \delta^{\gamma_0}],$$

which is nothing but (2.46) and thus the conclusion of Proposition 2.3.3 follows. \square

2.3.1 An estimate of the stopping time

The purpose of this section is to prove the following estimate for $R_{\mathbf{x},\mathbf{k}}^{(\delta)}[\tau_\delta < T]$.

Theorem 2.3.5 *Assume that the dimension $d \geq 3$. Then, one can choose $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$ in such a way that there exist constants $C, \gamma > 0$ for which*

$$R_{x,k}^{(\delta)}[\tau_\delta < T] \leq C\delta^\gamma T, \quad \forall \delta \in (0, 1], T \geq 1, (x, k) \in \mathcal{A}(M). \quad (2.57)$$

Proof. We obviously have

$$[\tau_\delta < T] = [U_\delta \leq \tau_\delta, U_\delta < T] \cup [S_\delta \leq \tau_\delta, S_\delta < T] \quad (2.58)$$

with the stopping times S_δ and U_δ defined in (2.27) and (2.28). Let us denote the first and second event appearing on the right hand side of (2.58) by $A(\delta)$ and $B(\delta)$ respectively. To show that (2.58) holds we prove that the $R_{x,k}^{(\delta)}$ probabilities of both events can be estimated by $C\delta^\gamma T$ for some $C, \gamma > 0$: see (2.66), (2.67) and (2.68).

An estimate of $R_{x,k}^{(\delta)}[A(\delta)]$

The first step towards obtaining the desired estimate will be to replace the event $A(\delta)$ whose definition involves a stopping time by an event $C(\delta)$ whose definition depends only on deterministic times, see (2.59) below. Next we use the estimate (2.45) of Proposition 2.3.3 for an appropriately chosen function G to reduce the question of bounding the $R_{x,k}^{(\delta)}$ probability of $\tilde{A}(\delta)$ by an easier problem of estimating its $\mathfrak{Q}_{x,k}$ probability ($\mathfrak{Q}_{x,k}$ corresponds to a degenerate diffusion determined by (2.43)). The latter is achieved by using bounds on heat kernels corresponding to hypoelliptic diffusions due to Kusuoka and Stroock.

We assume in this section to simplify the notation and without any loss of generality that $h^*(M) = 1$. Note that then

$$A(\delta) \subset \tilde{A}(\delta) := \left[\left| X\left(\frac{j}{q}\right) - X\left(\frac{i}{q}\right) \right| \leq \frac{3}{q} : 1 \leq i \leq j \leq [Tq], \quad |i - j| \geq \frac{q}{p} \right] \quad (2.59)$$

and thus

$$R_{x,k}^{(\delta)}[A(\delta)] \leq [Tq]^2 \max \left\{ R_{x,k}^{(\delta)} \left[\left| X\left(\frac{j}{q}\right) - X\left(\frac{i}{q}\right) \right| \leq \frac{3}{q} : 1 \leq i \leq j \leq [Tq], \quad |i - j| \geq \frac{q}{p} \right] \right\}. \quad (2.60)$$

Suppose that $f^{(\delta)} : \mathbb{R}^d \rightarrow [0, 1]$ is a C^∞ -regular function that satisfies $f(x) = 1$, if $|x| \leq 4h/q$ and $f^{(\delta)}(x) = 0$, if $|x| \geq 5/q$. We assume furthermore that i, j are positive integers such that $(j - i)/q \in [0, 1]$ and $\|f^{(\delta)}\|_3 \leq 2q^3$. For any $x_0 \in \mathbb{R}^d$ and $i/q \leq t \leq j/q$ define

$$G_j(t, x, k; x_0) := \mathfrak{M}_{x,k} f^{(\delta)} \left(X\left(\frac{j}{q} - t\right) - x_0 \right).$$

Obviously, we have

$$\partial_t G_j(t, x, k; x_0) + \tilde{\mathcal{L}} G_j(t, x, k; x_0) = 0.$$

Hence, using Proposition 2.3.3 with $v = j/q$ and $t = i/q$ (note that $v - t \geq 1/p \geq \delta^{\epsilon_2}$ and $\epsilon_2 \in (0, 1/2)$), we obtain that there exists $\gamma_1 > 0$ such that

$$\begin{aligned} & \left| E_{x,k}^{(\delta)} \left[f^{(\delta)} \left(X\left(\frac{j}{q}\right) - x_0 \right) - G_j \left(\frac{i}{q}, X\left(\frac{i}{q}\right), K\left(\frac{i}{q}\right); x_0 \right) \right] \mathcal{M}^{i/q} \right| \\ & \leq C \frac{j-i}{q} \|G_j(\cdot, \cdot, \cdot; x_0)\|_{1,1,3}^{[i/q, j/q]} T^2 \delta^{\gamma_1}, \quad \forall \delta \in (0, 1]. \end{aligned} \quad (2.61)$$

According to [57] Theorem 2.58, p. 53 we have

$$\|G_j(\cdot, \cdot, \cdot; x_0)\|_{1,1,3}^{[i/q, j/q]} \leq C\|f^{(\delta)}\|_3 \leq Cq^3 \leq C\delta^{-3(\epsilon_2+\epsilon_3)}, \quad j \in \{0, \dots, [qT]\}. \quad (2.62)$$

Hence combining (2.61) and (2.62) we obtain that the left hand side of (2.61) is less than, or equal to $C\delta^{\gamma_1-3(\epsilon_2+\epsilon_3)}$ for all $\delta \in (0, 1]$. Let now $i_0 = j - \frac{q}{p}$ so that $1 \leq i \leq i_0 \leq j \leq [Tq]$. We have

$$\begin{aligned} R_{x,k}^{(\delta)} \left[\left| X\left(\frac{j}{q}\right) - X\left(\frac{i}{q}\right) \right| \leq \frac{3}{q} \right] &\leq E_{x,k}^{(\delta)} \left[f^{(\delta)} \left(X\left(\frac{j}{q}\right) - X\left(\frac{i}{q}\right) \right) \right] \\ &= E_{x,k}^{(\delta)} \left[E_{x,k}^{(\delta)} \left[f^{(\delta)} \left(X\left(\frac{j}{q}\right) - \mathbf{y} \right) \mid \mathcal{M}^{i_0/q} \right] \right]_{\mathbf{y}=X(i/q)}. \end{aligned} \quad (2.63)$$

According to (2.61) and (2.62) we can estimate the utmost right hand side of (2.63) by

$$\sup_{x, \mathbf{y}, k} \left\{ \mathfrak{M}_{x,k} f^{(\delta)} \left(X\left(\frac{1}{p}\right) - \mathbf{y} \right) : x, \mathbf{y} \in \mathbb{R}^d, k \in A(2M) \right\} + C\delta^{\gamma_1-3(\epsilon_2+\epsilon_3)}T^2. \quad (2.64)$$

To estimate the first term in (2.64) we use the following.

Lemma 2.3.6 *Let p, q be as in (2.26). Then, there exist positive constants C_1, C_2 and C_3 such that for all $x, \mathbf{y} \in \mathbb{R}^d, k \in A(2M), j \in \{1, \dots, [pT]\}, \delta \in (0, 1]$ we have*

$$\mathfrak{Q}_{x,k} \left[\left| X\left(\frac{j}{p}\right) - \mathbf{y} \right| \leq \frac{5}{q} \right] \leq C_1 \left(\frac{p^{C_2}}{q^d} + e^{-C_3p} \right). \quad (2.65)$$

We leave the proof of this lemma outside these notes and finish the estimate of $R_{x,k}^{(\delta)}[A(\delta)]$. Using (2.65) we obtain that the expression in (2.64) can be estimated by

$$C_1 \left(\frac{p^{C_2}}{q^d} + e^{-C_3p} \right) + C\delta^{\gamma_1-3(\epsilon_2+\epsilon_3)}T^2 \leq C_1\delta^{(d-C_2)\epsilon_2+d\epsilon_3} + \exp\{-C_3\delta^{-\epsilon_2}\} + C\delta^{\gamma_1-3(\epsilon_2+\epsilon_3)}T^2.$$

Hence, from (2.60), we obtain that

$$\begin{aligned} R_{x,k}^{(\delta)}[A(\delta)] &\leq [Tq]^2 \left(C_1\delta^{(d-C_2)\epsilon_2+d\epsilon_3} + \exp\{-C_3\delta^{-\epsilon_2}\} + C\delta^{\gamma_1-3(\epsilon_2+\epsilon_3)}T^2 \right) \\ &\leq CT^2 \left(\delta^{(d-2-C_1)\epsilon_2+(d-2)\epsilon_3} + \delta^{-2(\epsilon_2+\epsilon_3)} \exp\{-C_3\delta^{-\epsilon_2}\} + \delta^{\gamma_1-5(\epsilon_2+\epsilon_3)}T^2 \right) \leq C\delta^{\gamma_2}T^4 \end{aligned} \quad (2.66)$$

for $\gamma_2 := \min[(d-2-C_1)\epsilon_2 + (d-2)\epsilon_3, \gamma_1 - 5(\epsilon_2 + \epsilon_3)] > 0$, provided that $\epsilon_2 + \epsilon_3 < \gamma_1/5$ and $\epsilon_2 \in (0, (d-2)\epsilon_3/(C_1+2-d))$. Here with no loss of generality we have assumed that $C_1+2 > d$. Recall also that $d \geq 3$. Now suppose that $\gamma_3 \in (0, \gamma_2)$. Consider two cases: $T^3 < \delta^{-\gamma_3}$ and $T^3 \geq \delta^{-\gamma_3}$. In the first one, the utmost right hand side of (2.66) can be bound from above by $C\delta^{\gamma_2-\gamma_3}T$. In the second we have a trivial bound of the left side by $\delta^{\gamma_3/3}T$. We have proved therefore that

$$R_{x,k}^{(\delta)}[A(\delta)] \leq C\delta^{\gamma}T \quad (2.67)$$

for some $C, \gamma > 0$ independent of δ and T .

An estimate of $R_{x,k}^{(\delta)}[B(\delta)]$

We start with a simple observation concerning the Hölder regularity of the K component of any path $\pi \in B(\delta)$. Let us denote $\rho := 2M_\delta^{-1}N^{-1/2}$ and

$$D := \left[\pi \in \mathcal{C}(T, \delta) : |K(t) - K(s)| \geq \rho \text{ for some } k \text{ s.t. } t_k^{(p)} \leq T \text{ and } t_{k-1}^{(p)} \leq s \leq t_k^{(p)} \leq t \leq t_{k+1}^{(p)} \right],$$

where M_δ has been defined in (2.9) and N in (2.26). Suppose that $\pi \in B(\delta)$, then we can find $t \in [t_k^{(p)}, t_{k+1}^{(p)}]$, $s \in [t_{k-1}^{(p)}, t_k^{(p)}]$ for which $\hat{K}(t) \cdot \hat{K}(s) \leq 1 - 1/N$. This, however, implies that

$$|K(t) - K(s)|^2 \geq \frac{1}{M_\delta^2} |\hat{K}(t) - \hat{K}(s)|^2 \geq \frac{2}{M_\delta^2 N},$$

thus $\pi \in D$. Hence the desired estimate of $R_{x,k}^{(\delta)}[B(\delta)]$ follows from the following lemma.

Lemma 2.3.7 *Under the assumptions of Theorem 2.3.5 there exist $C, \gamma > 0$ such that*

$$R_{x,k}^{(\delta)}[D] \leq CT\delta^\gamma, \quad \forall \delta \in (0, 1], T \geq 1, (x, k) \in \mathcal{A}(M). \quad (2.68)$$

The proof of this lemma is technical and we do not present it here. It relies heavily on Proposition 2.3.3.

2.3.2 The estimation of the convergence rate. The proof of Theorem 2.2.1.

Recall that $\phi_\delta, \bar{\phi}$ satisfy (2.10), (2.12), respectively, with the initial condition ϕ_0 . We start with the following lemma.

Lemma 2.3.8 *Assume that ϕ_0 satisfies the hypotheses formulated in Section 2.2.5. Then,*

$$\|\bar{\phi}\|_{0,0,0}^{[0,T]} \leq \|\phi_0\|_{0,0}, \quad \sum_{i=1}^d \|\partial_{x_i} \bar{\phi}\|_{0,0,0}^{[0,T]} \leq \|\phi_0\|_{1,0}. \quad (2.69)$$

Furthermore, there exists a constant $C > 0$ such that for all $T \geq 1$

$$\|\partial_t \bar{\phi}\|_{0,0,0}^{[0,T]} \leq C\|\phi_0\|_{1,2}. \quad (2.70)$$

In addition, for any nonnegative integer valued multi-index $\gamma = (\alpha_1, \alpha_2, \alpha_3)$ satisfying $|\gamma| \leq 3$ we have

$$\sum_{i_1, i_2, i_3=1}^d \|\partial_{k_{i_1}, k_{i_2}, k_{i_3}}^\gamma \bar{\phi}\|_{0,0,0}^{[0,T]} \leq CT^{|\gamma|} \|\phi_0\|_{1,4}, \quad (2.71)$$

The proof is quite straightforward and we leave it out of these notes.

Proof of Theorem 2.2.1. Let $u \in [\delta^{\gamma'_0}, T]$, where we assume that γ'_0 (as in the statement of Lemma 2.3.4) belongs to the interval $(1/2, 1)$. Substituting for $G(t, x, k) := \bar{\phi}(u - t, x, k)$, $\zeta \equiv 1$ into (2.45) we obtain (taking $v = u$, $t = \delta^{\gamma'_0}$)

$$\left| \tilde{E}_{x,k}^{(\delta)} \left[\phi_0(X(u), K(u)) - \bar{\phi}(u - \delta^{\gamma'_0}, X(\delta^{\gamma'_0}), K(\delta^{\gamma'_0})) - \int_{\delta^{\gamma'_0}}^u (\partial_\varrho + \hat{\mathcal{L}}_\varrho) G(\varrho, X(\varrho), K(\varrho)) d\varrho \right] \right| \leq C \|G\|_{1,1,3}^{[0,T]} \delta^{\gamma_1} T^2, \quad \forall \delta \in (0, 1]. \quad (2.72)$$

Using the fact that $|X(\delta\gamma'_0) - x| \leq C\delta\gamma'_0$, $|K(\delta\gamma'_0) - k| \leq C\delta\gamma'_0^{-1/2}$, $\tilde{Q}_{x,k}^{(\delta)}$ -a.s. for some deterministic constant $C > 0$, cf. (2.35), and Lemma 2.3.8 we obtain that there exist constants $C, \gamma > 0$ such that

$$\left| \tilde{E}_{x,k}^{(\delta)} \left[\phi_0(X(u), K(u)) - \bar{\phi}(u, x, k) - \int_0^u (\partial_\varrho + \hat{\mathcal{L}}_\varrho) G(\varrho, X(\varrho), K(\varrho)) d\varrho \right] \right| \leq C \|G\|_{1,1,3}^{[0,T]} \delta^\gamma T^2, \quad \delta \in (0, 1], T \geq 1, u \in [0, T]. \quad (2.73)$$

We have however

$$\left| E_{x,k}^{(\delta)} [\phi_0(X(u), K(u)) - \bar{\phi}(u, x, k), \tau_\delta \geq T] \right| = \left| \tilde{E}_{x,k}^{(\delta)} [\phi_0(X(u), K(u)) - \bar{\phi}(u, x, k), \tau_\delta \geq T] \right| \stackrel{(2.73)}{\leq} C \|G\|_{1,1,3}^{[0,T]} \delta^\gamma T^2 + \left(2\|\phi_0\|_{0,0} + T \|G\|_{1,1,2}^{[0,T]} \right) \tilde{Q}_{x,k}^{(\delta)}[\tau_\delta < T]. \quad (2.74)$$

Using $\mathcal{M}^{\tau_\delta}$ measurability of the event $[\tau_\delta < T]$ we obtain that $\tilde{Q}_{x,k}^{(\delta)}[\tau_\delta < T] = R_{x,k}^{(\delta)}[\tau_\delta < T]$ and by virtue of Theorem 2.3.5 we can estimate the right hand side of (2.74) by

$$C \|G\|_{1,1,3}^{[0,T]} \delta^\gamma T^2 + C \delta^\gamma T \left(2\|\phi_0\|_{0,0} + T \|G\|_{1,1,2}^{[0,T]} \right) \stackrel{\text{Lemma 2.3.8}}{\leq} C \delta^\gamma T^5.$$

On the other hand, the expression under the absolute value on the utmost left hand side of (2.74) equals

$$E_{x,k}^{(\delta)} [\phi_0(X(u), K(u)) - \bar{\phi}(u, x, k)] - E_{x,k}^{(\delta)} [\phi_0(X(u), K(u)) - \bar{\phi}(u, x, k), \tau_\delta < T].$$

The second term can be estimated by

$$2\|\phi_0\|_{0,0} R_{x,k}^{(\delta)}[\tau_\delta < T] \stackrel{(2.57)}{\leq} C \delta^\gamma \|\phi_0\|_{0,0} T,$$

by virtue of Theorem 2.3.5. Since

$$\mathbb{E} \phi_\delta \left(\frac{u}{\delta}, \frac{x}{\delta}, k \right) = \mathbb{E} \phi_0(z^{(\delta)}(u; x, k), \mathbf{m}^{(\delta)}(u; x, k)) = E_{x,k}^{(\delta)} \phi_0(X(u), K(u))$$

we conclude from the above that the left hand side of (2.13) can be estimated by $C \delta^\gamma \|\phi_0\|_{1,4} T^5$ for some constants $C, \gamma > 0$ independent of $\delta > 0, T \geq 1$. The bound appearing on the right hand side of (2.13) can be now concluded by the same argument as the one used after (2.66). \square

2.4 The spatial diffusion of wave energy

In this section we consider an application of the previous results to the random geometrical optics regime of propagation of acoustic waves. We show that when the wave length is much shorter than the correlation length of the random medium, there exist temporal and spatial scales where the energy density of the wave undergoes the spatial diffusion. We start with the wave equation in dimension $d \geq 3$

$$\frac{1}{c^2(x)} \frac{\partial^2 \phi}{\partial t^2} - \Delta \phi = 0 \quad (2.75)$$

and assume that the wave speed has the form $c(x) = c_0 + \sqrt{\delta}c_1(x)$. Here $c_0 > 0$ is the constant sound speed of the uniform background medium, while the small parameter $\delta \ll 1$ measures the strength of the mean zero random perturbation c_1 . Rescaling the spatial and temporal variables $x = x'/\delta$ and $t = t'/\delta$ we obtain (after dropping the primes) equation (2.75) with a rapidly fluctuating wave speed

$$c_\delta(x) = c_0 + \sqrt{\delta}c_1\left(\frac{x}{\delta}\right). \quad (2.76)$$

It is convenient to rewrite (2.75) as the system of acoustic equations for the “pressure” $p = \phi_t/c$ and “acoustic velocity” $\mathbf{u} = -\nabla\phi$:

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + \nabla(c_\delta(x)p) &= 0 \\ \frac{\partial p}{\partial t} + c_\delta(x)\nabla \cdot \mathbf{u} &= 0. \end{aligned} \quad (2.77)$$

We will denote for brevity $\mathbf{v} = (\mathbf{u}, p) \in \mathbb{R}^{d+1}$ and write (2.77) in the more general form of a first order linear symmetric hyperbolic system. To do so we introduce symmetric matrices A_δ and D^j defined by

$$A_\delta(x) = \text{diag}(1, 1, 1, c_\delta(x)), \quad \text{and} \quad D^j = \mathbf{e}_j \otimes \mathbf{e}_{d+1} + \mathbf{e}_{d+1} \otimes \mathbf{e}_j, \quad j = 1, \dots, d. \quad (2.78)$$

Here $\mathbf{e}_m \in \mathbb{R}^{d+1}$ is the standard orthonormal basis: $(\mathbf{e}_m)_k = \delta_{mk}$.

We consider the initial data for (2.77) as a mixture of states. Let S be a measure space equipped with a non-negative finite measure μ . A typical example is that the initial data is random, S is the state space and μ is the corresponding probability measure. We assume that for each parameter $\zeta \in S$ and $\varepsilon, \delta > 0$ the initial data is given by $\mathbf{v}_\varepsilon^\delta(0, x; \zeta) := (-\varepsilon \nabla \phi_0^\varepsilon(x), 1/c_\delta(x)\phi_0^\varepsilon(x))$ and $\mathbf{v}_\varepsilon^\delta(t, x; \zeta)$ solves the system of equations

$$\frac{\partial \mathbf{v}_\varepsilon^\delta}{\partial t} + \sum_{j=1}^d A_\delta(x) D^j \frac{\partial}{\partial x^j} (A_\delta(x) \mathbf{v}_\varepsilon^\delta(x)) = 0. \quad (2.79)$$

The set of initial data is assumed to form an ε -oscillatory and compact at infinity family [31] as $\varepsilon \rightarrow 0$. By the above we mean that for any continuous, compactly supported function $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ we have

$$\lim_{R \rightarrow +\infty} \limsup_{\varepsilon \rightarrow 0^+} \int_{|k| \geq R/\varepsilon} |\widehat{\varphi \mathbf{v}_\varepsilon^\delta}|^2 dk \rightarrow 0 \quad \text{and} \quad \lim_{R \rightarrow +\infty} \limsup_{\varepsilon \rightarrow 0^+} \int_{|x| \geq R} |\mathbf{v}_\varepsilon^\delta|^2 dx \rightarrow 0$$

for a fixed realization $\zeta \in S$ of the initial data and each $\delta > 0$. In the regime of geometric acoustics the scale ε of oscillations of the initial data is much smaller than the correlation length δ of the medium: $\varepsilon \ll \delta \ll 1$.

The dispersion matrix for (2.79) is

$$P_0^\delta(x, k) = i \sum_{j=1}^d A_\delta(x) k_j D^j A_\delta(x) = i \sum_{j=1}^d c_\delta(x) k_j D^j = i c_\delta(x) (\tilde{k} \otimes \mathbf{e}_{d+1} + \mathbf{e}_{d+1} \otimes \tilde{k}), \quad (2.80)$$

where $\tilde{k} = \sum_{j=1}^d k_j \mathbf{e}_j$. The self-adjoint matrix $(-iP_0^\delta)$ has an eigenvalue $H_0 = 0$ of the multiplicity $d - 1$, and two simple eigenvalues

$$H_\pm^\delta(x, k) = \pm c_\delta(x) |k|. \quad (2.81)$$

Its eigenvectors are

$$b_m^0 = (k_m^\perp, 0), \quad m = 1, \dots, d-1; \quad b_\pm = \frac{1}{\sqrt{2}} \left(\frac{\tilde{k}}{|k|} \pm \mathbf{e}_{d+1} \right), \quad (2.82)$$

where $k_m^\perp \in \mathbb{R}^d$ is the orthonormal basis of vectors orthogonal to k .

The $(d+1) \times (d+1)$ Wigner matrix of a mixture of solutions of (2.79) is defined by

$$W_\varepsilon^\delta(t, x, k) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_S e^{ik \cdot y} \mathbf{v}_\varepsilon^\delta(t, x - \frac{\varepsilon y}{2}; \zeta) \mathbf{v}_\varepsilon^{\delta*}(t, x + \frac{\varepsilon y}{2}; \zeta) dy d\zeta. \quad (2.83)$$

It is well-known, see [31, 49, 53], that for each fixed $\delta > 0$ (and even without introduction of a mixture of states) when $W_\varepsilon^\delta(t=0)$ converges weakly in $\mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$, as $\varepsilon \rightarrow 0$, to

$$W_0(x, k) = u_+^0(x, k) b_+(k) \otimes b_+(k) + u_-^0(x, k) b_-(k) \otimes b_-(k). \quad (2.84)$$

then $W_\varepsilon^\delta(t)$ converges weakly in $\mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ to

$$U^\delta(t, x, k) = u_+^\delta(t, x, k) b_+(k) \otimes b_+(k) + u_-^\delta(t, x, k) b_-(k) \otimes b_-(k).$$

The scalar amplitudes $u_\pm^{(\delta)}$ satisfy the Liouville equations:

$$\begin{cases} \partial_t u_\pm^\delta + \nabla_k H_\pm^\delta \cdot \nabla_x u_\pm^\delta - \nabla_x H_\pm^\delta \cdot \nabla_k u_\pm^\delta = 0, \\ u_\pm^\delta(0, x, k) = u_\pm^0(x, k). \end{cases} \quad (2.85)$$

These equations are of the form (2.10), written in the macroscopic variables, with the Hamiltonian given by (2.81).

One may obtain an L^2 -error estimate for this convergence when a mixture of states is introduced, as in (2.83). In order to make the scale separation $\varepsilon \ll \delta \ll 1$ precise we define the set

$$\mathcal{K}_\mu := \left\{ (\varepsilon, \delta) : |\ln \varepsilon|^{-2/3+\mu} \leq \delta \leq 1 \right\}.$$

The parameter μ is a fixed number in the interval $(0, 2/3)$. The following proposition has been proved in Theorem 3.2 of [6], using straightforward if tedious asymptotic expansions.

Proposition 2.4.1 *Let the acoustic speed $c_\delta(x)$ be of the form (2.76) and such that the Hamiltonian $H_\delta(x)$ given by (2.81) satisfies assumptions (2.6). We assume that the Wigner transform W_ε^δ satisfies*

$$W_\varepsilon^\delta(0, x, k) \rightarrow W_0(x, k) \text{ strongly in } L^2(\mathbb{R}^d \times \mathbb{R}^d) \text{ as } \mathcal{K}_\mu \ni (\varepsilon, \delta) \rightarrow 0. \quad (2.86)$$

We also assume that the limit $W_0 \in C_c^2(\mathbb{R}^d \times \mathbb{R}^d)$ with a support that satisfies

$$\text{supp } W_0(x, k) \subseteq \mathcal{A}(M) \quad (2.87)$$

for some $M > 0$. Moreover, we assume that the initial limit Wigner transform W_0 is of the form

$$W_0(x, k) = \sum_{q=\pm} u_q^0(x, k) \Pi_q(k), \quad \Pi_q(k) = b_q(k) \otimes b_q(k). \quad (2.88)$$

Let $U^\delta(t, x, k) = \sum_{p=\pm} u_p^\delta(t, x, k) \Pi_p(k)$, where the functions u_p^δ satisfy the Liouville equations (2.85). Then there exists a constant $C_1 > 0$ that is independent of δ so that

$$\|W_\varepsilon^\delta(t, x, k) - U^\delta(t, x, k)\|_2 \leq C(\delta) \left(\varepsilon \|W_0\|_{H^2} e^{C_1 t / \delta^{3/2}} + \varepsilon^2 \|W_0\|_{H^3} e^{C_1 t / \delta^{3/2}} \right) + \|W_\varepsilon^\delta(0) - W_0\|_2, \quad (2.89)$$

where $C(\delta)$ is a rational function of δ with deterministic coefficients that may depend on the constant $M > 0$ in the bound (2.87) on the support of W_0 .

The Liouville equations (2.85) are of the form (2.10). Therefore, one may pass to the limit $\delta \rightarrow 0$ in (2.85) using Theorem 2.2.1 and conclude that $\mathbb{E} u_\pm^\delta$ converge to the respective solutions of

$$\frac{\partial \bar{u}_\pm}{\partial t} = \sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left(|k|^2 D_{mn}(\hat{k}) \frac{\partial \bar{u}_\pm}{\partial k_n} \right) \pm c_0 \hat{k} \cdot \nabla_x \bar{u}_\pm \quad (2.90)$$

with the initial conditions as in (2.85). Here the diffusion matrix $D(\hat{k}) = [D_{mn}(\hat{k})]$ is given by

$$D_{mn}(\hat{k}) = -\frac{1}{2} \int_{-\infty}^{\infty} \frac{\partial^2 R(c_0 s \hat{k})}{\partial x_n \partial x_m} ds, \quad (2.91)$$

where $R(x)$ is the correlation function of the random field $c_1(x)$: $\mathbb{E}[c_1(z)c_1(x+z)] = R(x)$. Furthermore, it follows from Theorem 2.2.3 that there exists $\alpha_0 > 0$ so that solutions of (2.85) with the initial data of the form $u_\pm^\delta(0, x, k) = u_\pm^0(\delta^\alpha x, k)$ with $0 < \alpha < \alpha_0$, converge in the long time limit to the solutions of the spatial diffusion equation. More precisely, in that case the function $\bar{u}^\delta(t, x, k) = u_+^\delta(t/\delta^{2\alpha}, x/\delta^\alpha, k)$ (and similarly for u_-^δ) converges as $\delta \rightarrow 0$ to $w(t, x, k)$ – the solution of the spatial diffusion equation

$$\begin{aligned} \frac{\partial w}{\partial t} &= \sum_{m,n=1}^d a_{mn}(k) \frac{\partial^2 w}{\partial x_n \partial x_m}, \\ w(0, x, k) &= \bar{u}_+^0(x; k) := \frac{1}{\Gamma_{d-1}} \int_{\mathbb{S}^{d-1}} u_+^0(x, k) d\Omega(\hat{k}). \end{aligned} \quad (2.92)$$

with the diffusion matrix a_{mn} given by:

$$a_{nm}(k) = \frac{c_0}{\Gamma_{d-1}} \int_{\mathbb{S}^{d-1}} \hat{k}_n \chi_m(k) d\Omega(\hat{k}), \quad (2.93)$$

and the functions χ_j above are the mean-zero solutions of

$$\sum_{m,n=1}^d \frac{\partial}{\partial k_m} \left(k^2 D_{mn}(\hat{k}) \frac{\partial \chi_j}{\partial k_n} \right) = -c_0 \hat{k}_j. \quad (2.94)$$

Theorems 2.2.1, 2.2.2 and 2.2.3 allow us to make the passage to the limit $\varepsilon, \delta, \gamma \rightarrow 0$ rigorous. The assumption that $\varepsilon \ll \delta \ll \gamma$ is formalized as follows. We let

$$\mathcal{K}_{\mu, \rho} := \left\{ (\varepsilon, \delta, \gamma) : \delta \geq |\ln \varepsilon|^{-2/3+\mu} \text{ and } \gamma \geq \delta^\rho \right\},$$

with $0 < \mu < 2/3$, $\rho \in (0, 1)$. Suppose also that $u_0^\pm \in C_c^3(\mathbb{R}_*^{2d})$ and $\text{supp } u_0^\pm \subseteq \mathcal{A}(M)$. Let

$$W^0(x, k) := u_+^0(x, k) b_+(k) \otimes b_+(k) + u_-^0(x, k) b_-(k) \otimes b_-(k), \quad (2.95)$$

and

$$W(t, x, k) := w_+(t, x; k) b_+(k) \otimes b_+(k) + w_-(t, x; k) b_-(k) \otimes b_-(k). \quad (2.96)$$

Our main result regarding the diffusion of wave energy can be stated as follows.

Theorem 2.4.2 *Assume that the dimension $d \geq 3$ and $M \geq 1$ are fixed. Suppose for some $0 < \mu < 2/3$, $\rho \in (0, 1)$ we have, with W^0 as in (2.95) and W_ε^δ defined by (2.83)*

$$\int_{\mathbb{R}^{2d}} \left| \mathbb{E} W_\varepsilon^\delta \left(0, \frac{x}{\gamma}, k \right) - W^0(x, k) \right|^2 dx dk \rightarrow 0, \text{ as } (\varepsilon, \delta, \gamma) \rightarrow 0 \text{ and } (\varepsilon, \delta, \gamma) \in \mathcal{K}_{\mu, \rho}.$$

Then, there exists $\rho_1 \in (0, \rho]$ such that for any $T > T_ > 0$ we have*

$$\sup_{t \in [T_*, T]} \int \left| \mathbb{E} W_\varepsilon^\delta \left(\frac{t}{\gamma^2}, \frac{x}{\gamma}, k \right) - W(t, x, k) \right|^2 dx dk \rightarrow 0, \text{ as } (\varepsilon, \delta, \gamma) \rightarrow 0 \text{ and } (\varepsilon, \delta, \gamma) \in \mathcal{K}_{\mu, \rho_1}.$$

Here $W(t, x, k)$ is of the form (2.96) with the functions w_\pm that satisfy (2.92) with the initial data $w_\pm(0, x, k) = u_\pm^0(x, k)$.

The proof follows immediately from Theorems 2.2.1, 2.2.2 and 2.2.3 as well as Proposition 2.4.1.

Chapter 3

Applications to Time-Reversal Experiments

This chapter is based on [4, 5] but the concept of one-step time reversal emerged during early discussions with Knut Solna.

3.1 Time-reversal experiments

In time reversal experiments, acoustic waves are emitted from a localized source, recorded in time by an array of receivers-transducers, time reversed, and re-transmitted into the medium, so that the signals recorded first are re-emitted last and vice versa [20, 21, 28, 35, 42, 46]: a schematic description of the time reversal procedure is depicted in Fig. 3.1.

Early experiments in time reversal acoustics are described in [20]; see also the more recent papers [26, 27, 28] – this list is by no means exhaustive and the literature on the subject is by now vast. The re-transmitted signal refocuses at the location of the original source with a modified shape that depends on the array of receivers. The salient feature of these time reversal experiments is that refocusing is much better when wave propagation occurs in complicated environments than in homogeneous media. Time reversal techniques with improved refocusing in heterogeneous medium have found important applications in medicine, non-destructive testing, underwater acoustics, and wireless communications (see the above references). It has been also applied to imaging in weakly random media [10, 14, 28] and led to a recent concept of coherent interferometric imaging (CINT) of Borcea, Papanicolaou and Tsogka [15, 16, 17].

A very qualitative explanation for the better refocusing observed in heterogeneous media is based on *multipathing*. Since waves can scatter off a larger number of heterogeneities, more paths coming from the source reach the recording array, thus more is known about the source by the transducers than in a homogeneous medium. The heterogeneous medium plays the role of a lens that widens the aperture through which the array of receivers sees the source. Refocusing is also qualitatively justified by ray theory (geometrical optics). The phase shift caused by multiple scattering is exactly compensated when the time reversed signal follows the same path back to the source location. This phase cancellation happens only at the source location. The phase shift along paths leading to other points in space is essentially random. The interference of multiple paths will thus be constructive at the source location and destructive anywhere else. This explains why refocusing at the source location is improved when the number of scatterers is large.

As convincing as they are, the above explanations remain qualitative and do not allow us

Figure 3.1: The Time Reversal Procedure. Top: Propagation of signal and measurements in time. Bottom: Time reversal of recorded signals and back-propagation into the medium.

to quantify how the refocused signal is modified by the time reversal procedure. Quantitative justifications require to analyze wave propagation more carefully. The first quantitative description of time reversal was obtained in [18] in the framework of one-dimensional random media. That paper provides the first mathematical explanation of two of the most prominent features of time reversal: heterogeneities improve refocusing and refocusing occurs for almost every realization of the random medium. Various extensions and generalizations to the three-dimensional layered case, including nonlinear effects, have been done in the work by Garnier, Fouque, Nachbin, Papanicolaou and Solna, and are described in detail in the recent excellent book [30]. The first multi-dimensional quantitative description of time reversal was obtained in [13] for the parabolic approximation, i.e., for waves that propagate in a privileged direction with no backscattering (see also [54] for further analysis of time reversal in this regime). That paper shows that the random medium indeed plays the role of a lens. The back-propagated signal behaves as if the initial array were replaced by another one with a much bigger effective aperture. In a slightly different context, time reversal in ergodic cavities was analyzed in [8]. There, wave mixing is created by reflection at the boundary of a chaotic cavity, which plays a similar role to the heterogeneities in a heterogeneous medium.

In this chapter we consider the theory of time-reversal experiments for general classical waves propagating in weakly fluctuating random media. It is convenient to understand refocusing in time reversal experiments in the following three-step general framework:

- (i) A signal propagating from a localized source is recorded at a single time $T > 0$ by an array of receivers.
- (ii) The recorded signal is processed at the array location.
- (iii) The processed signal is emitted from the array and propagates in the *same* medium during the same amount of time T .

As we will see, this formulation allows us to reduce the mathematical problem of the description of the refocused signal to the question of the passage from the wave equations to the kinetic models. While the latter problem is also difficult, we may apply whatever is known in that area to the time-reversal problems. Accordingly, the mathematical rigor of our statements on time-reversal experiments below depends on the regime of consideration – for instance, they are mostly formal in the radiative transfer regime but are rigorous in the random geometric optics regime (see [6] for the precise statements). To keep the presentation uniform we will concentrate here solely on the transport regime.

The first main result of this chapter is that the repropagated signal will refocus at the location of the original source for a large class of waves and a large class of processings. The experiments described above correspond to the specific processing of acoustic waves in which pressure is kept unchanged and the sign of the velocity field is reversed.

The second main result is a quantitative description of the re-transmitted signal. We show that the re-propagated signal $\mathbf{u}^B(\xi)$ at a point ξ near the source location can be written in the high frequency limit as the following convolution of the original source \mathbf{S}

$$\mathbf{u}^B(\xi) = (F * \mathbf{S})(\xi). \quad (3.1)$$

The kernel F depends on the location of the recording array and on the signal processing. The quality of the refocusing depends on the spatial decay of F . It turns out that it can be expressed in terms of the Wigner transform [53] of two wave fields. The decay properties of F depend on the smoothness of the Wigner transform in the phase space and it is here that the kinetic theories becomes useful. Here we consider the high frequency regime when the

wavelength of the initial signal is small compared to the distance of propagation. In addition we assume that the wavelength is comparable to the correlation length of the medium. This is the radiative transport regime. It has been extensively studied mathematically for the Schrödinger equation [23, 56] and formally using perturbation expansions for the classical waves [7, 53]. In this regime the Wigner transform satisfies a radiative transport equation, which is used to describe the evolution of the energy density of waves in random media [36, 53, 55, 56]. The transport equations possess a smoothing effect so that the Wigner distribution becomes less singular in random media, which implies a stronger decay of the convolution kernel F and a better refocusing. The diffusion approximation to the radiative transport equations provides simple reconstruction formulas that can be used to quantify the refocusing quality of the back-propagated signal. This construction applies to a large class of classical waves: acoustic, electromagnetic, elastic, and others, and allows for a large class of signal processings at the recording array.

3.2 Classical Time Reversal and One-Step Time Reversal

Propagation of acoustic waves is described by a system of equations for the pressure $p(t, x)$ and acoustic velocity $\mathbf{v}(t, x)$:

$$\begin{aligned}\rho(x) \frac{\partial \mathbf{v}}{\partial t} + \nabla p &= 0 \\ \kappa(x) \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{v} &= 0,\end{aligned}\tag{3.2}$$

with suitable initial conditions and where $\rho(x)$ and $\kappa(x)$ are density and compressibility of the underlying medium, respectively. These equations can be recast as the following linear hyperbolic system

$$A(x) \frac{\partial \mathbf{u}}{\partial t} + D^j \frac{\partial \mathbf{u}}{\partial x^j} = 0, \quad x \in \mathbb{R}^3\tag{3.3}$$

with the vector $\mathbf{u} = (\mathbf{v}, p) \in \mathbb{C}^4$. The matrix $A = \text{Diag}(\rho, \rho, \rho, \kappa)$ is positive definite. The 4×4 matrices D^j , $j = 1, 2, 3$, are symmetric and given by $D_{mn}^j = \delta_{m4} \delta_{nj} + \delta_{n4} \delta_{mj}$. We use the Einstein convention of summation over repeated indices.

The time reversal experiments in [20] consist of two steps. First, the direct problem

$$\begin{aligned}A(x) \frac{\partial \mathbf{u}}{\partial t} + D^j \frac{\partial \mathbf{u}}{\partial x^j} &= 0, \quad 0 \leq t \leq T \\ \mathbf{u}(0, x) &= \mathbf{S}(x)\end{aligned}\tag{3.4}$$

with a localized source \mathbf{S} centered at a point x_0 is solved. The signal is recorded during the period of time $0 \leq t \leq T$ by an array of receivers located at $\Omega \subset \mathbb{R}^3$. Second, the signal is time reversed and re-emitted into the medium. Time reversal is described by multiplying $\mathbf{u} = (\mathbf{v}, p)$ by the matrix $\Gamma = \text{Diag}(-1, -1, -1, 1)$. The back-propagated signal solves

$$\begin{aligned}\frac{\partial \mathbf{u}}{\partial t} + A^{-1}(x) D^j \frac{\partial \mathbf{u}}{\partial x^j} &= \frac{1}{T} \mathbf{R}(2T - t, x), \quad T \leq t \leq 2T \\ \mathbf{u}(T, x) &= 0\end{aligned}\tag{3.5}$$

with the source term

$$\mathbf{R}(t, x) = \Gamma \mathbf{u}(t, x) \chi(x).\tag{3.6}$$

The function $\chi(x)$ is either the characteristic function of the set where the recording array is located, or some other function that allows for possibly space-dependent amplification of the re-transmitted signal.

The back-propagated signal is then given by $\mathbf{u}(2T, x)$. We can decompose it as

$$\mathbf{u}(2T, x) = \frac{1}{T} \int_0^T ds \mathbf{w}(s, x; s), \quad (3.7)$$

where the vector-valued function $\mathbf{w}(t, x; s)$ solves the initial value problem

$$\begin{aligned} A(x) \frac{\partial \mathbf{w}(t, x; s)}{\partial t} + D^j \frac{\partial \mathbf{w}(t, x; s)}{\partial x^j} &= 0, \quad 0 \leq t \leq s \\ \mathbf{w}(0, x; s) &= \mathbf{R}(s, x). \end{aligned}$$

We deduce from (3.7) that it is sufficient to analyze the refocusing properties of $\mathbf{w}(s, x; s)$ for $0 \leq s \leq T$ to obtain those of $\mathbf{u}(2T, x)$. For a fixed value of s , we call the construction of $\mathbf{w}(s, x; s)$ one-step time reversal.

We define one-step time reversal more generally as follows. The direct problem (3.4) is solved until time $t = T$ to yield $\mathbf{u}(T^-, x)$. At time T , the signal is recorded and processed. The processing is modeled by an amplification function $\chi(x)$, a blurring kernel $f(x)$, and a (possibly spatially varying) time reversal matrix Γ . After processing, we have

$$\mathbf{u}(T^+, x) = \Gamma(f * (\chi \mathbf{u}))(T^-, x) \chi(x). \quad (3.8)$$

The processed signal then propagates for the same amount of time T :

$$\begin{aligned} A(x) \frac{\partial \mathbf{u}}{\partial t} + D^j \frac{\partial \mathbf{u}}{\partial x^j} &= 0, \quad T \leq t \leq 2T \\ \mathbf{u}(T^+, x) &= \Gamma(f * (\chi \mathbf{u}))(T^-, x) \chi(x). \end{aligned} \quad (3.9)$$

The main question is whether $\mathbf{u}(2T, x)$ refocuses at the location of the original source $\mathbf{S}(x)$ and how the original signal has been modified by the time reversal procedure. Notice that in the case of full ($\Omega = \mathbb{R}^3$) and exact ($f(x) = \delta(x)$) measurements with $\Gamma = \text{Diag}(-1, -1, -1, 1)$, the time-reversibility of first-order hyperbolic systems implies that $\mathbf{u}(2T, x) = \Gamma \mathbf{S}(x)$, which corresponds to exact refocusing. When only partial measurements are available we shall see in the following sections that $\mathbf{u}(2T, x)$ is closer to $\Gamma \mathbf{S}(x)$ when propagation occurs in a heterogeneous medium than in a homogeneous medium.

The pressure field $p(t, x)$ satisfies the following scalar wave equation

$$\frac{\partial^2 p}{\partial t^2} - \frac{1}{\kappa(x)} \nabla \cdot \left(\frac{1}{\rho(x)} \nabla p \right) = 0. \quad (3.10)$$

A schematic description of the one-step procedure for the wave equation is presented in Fig. 3.2. A numerical experiment for the one-step time reversal procedure is shown in Fig. 3.3. In the numerical simulations, there is no blurring, $f(x) = \delta(x)$, and the array of receivers is the domain $\Omega = (-1/6, 1/6)^2$ ($\chi(x)$ is the characteristic function of Ω). Note that the truncated signal does not retain any information about the ballistic part of the original wave (the part that propagates without scattering with the underlying medium). In a homogeneous medium, the truncated signal would then be nearly identically zero (not quite zero since the numerics are done in two dimensions) and no refocusing would be observed. The interesting aspect of time reversal is that a coherent signal emerges at time $2T$ out of a signal at time T^+ that seems to have no useful information.

3.3 Theory of Time Reversal in Random Media

Our objective is now to present a theory that explains in a quantitative manner the refocusing properties described in the preceding sections. We consider here the one-step time reversal for acoustic wave. Generalizations to other types of waves and more general processings in (3.9) are given in Section 3.4.

Figure 3.2: The One-Step Time Reversal Procedure. Here, p_t denotes $\frac{\partial p}{\partial t}$.

Figure 3.3: Numerical experiment using the one-step time reversal procedure. Top Left: initial condition $p(0, x)$, a peaked Gaussian of maximal amplitude equal to 1. Top Right: forward solution $p(T^-, x)$, of maximal amplitude 0.04. Bottom Right: recorded solution $p(T^+, x)$, of maximal amplitude 0.015 on the domain $\Omega = (-1/6, 1/6)^2$. Bottom Left: back-propagated solution $p(2T, x)$, of maximal amplitude 0.07.

3.3.1 Refocused Signal

We recall that the one-step time reversal procedure consists of letting an initial pulse $\mathbf{S}(x)$ propagate according to (3.4) until time T ,

$$\mathbf{u}(T^-, x) = \int_{\mathbb{R}^3} G(T, x; z) \mathbf{S}(z) dz,$$

where $G(T, x; z)$ is the Green's matrix solution of

$$\begin{aligned} A(x) \frac{\partial G(t, x; y)}{\partial t} + D^j \frac{\partial G(t, x; y)}{\partial x^j} &= 0, \quad 0 \leq t \leq T \\ G(0, x; y) &= I \delta(x - y). \end{aligned} \quad (3.11)$$

At time T , the “intelligent” array reverses the signal. For acoustic pulses, this means keeping pressure unchanged and reversing the sign of the velocity field. The array of receivers is located in $\Omega \subset \mathbb{R}^3$. The amplification function $\chi(x)$ is an arbitrary bounded function supported in Ω , such as its characteristic function ($\chi(x) = 1$ for $x \in \Omega$ and $\chi(x) = 0$ otherwise) when all transducers have the same amplification factor. We also allow for some blurring of the recorded data modeled by a convolution with a function $f(x)$. The case $f(x) = \delta(x)$ corresponds to exact measurements. Finally, the signal is time reversed, that is, the direction of the acoustic velocity is reversed. Here, the operator Γ in (3.8) is simply multiplication by the matrix

$$\Gamma = \text{Diag}(-1, -1, -1, 1). \quad (3.12)$$

The signal at time T^+ after time reversal takes then the form

$$\mathbf{u}(T^+, x) = \int_{\mathbb{R}^6} \Gamma G(T, y'; z) \chi(x) \chi(y') f(x - y') \mathbf{S}(z) dz dy'. \quad (3.13)$$

The last step (3.9) consists of letting the time reversed field propagate through the random medium until time $2T$. To compare this signal with the initial pulse \mathbf{S} , we need to reverse the acoustic velocity once again, and define

$$\mathbf{u}^B(x) = \Gamma \mathbf{u}(2T, x) = \int_{\mathbb{R}^9} \Gamma G(T, x; y) \Gamma G(T, y'; z) \chi(y) \chi(y') f(y - y') \mathbf{S}(z) dy dy' dz. \quad (3.14)$$

The time reversibility of first-order hyperbolic systems implies that $\mathbf{u}^B(x) = \mathbf{S}(x)$ when $\Omega = \mathbb{R}^d$, $\chi \equiv 1$, and $f(x) = \delta(x)$, that is, when full and non-distorted measurements are available. It remains to understand which features of \mathbf{S} are retained by $\mathbf{u}^B(x)$ when only partial measurement is available.

3.3.2 Localized Source and Scaling

We consider an asymptotic solution of the time reversal problem (3.4), (3.9) when the support λ of the initial pulse $\mathbf{S}(x)$ is much smaller than the distance L of propagation between the source and the recording array: $\varepsilon = \lambda/L \ll 1$. We also take the size a of the array comparable to L : $a/L = O(1)$. We assume that the time T between the emission of the original signal and recording is of order L/c_0 , where c_0 is a typical speed of propagation of the acoustic wave. We consequently consider the initial pulse to be of the form

$$\mathbf{u}(0, x) = \mathbf{S}\left(\frac{x - x_0}{\varepsilon}\right)$$

in non-dimensionalized variables $x' = x/L$ and $t' = t/(L/c_0)$. We drop primes to simplify notation. Here x_0 is the location of the source. The transducers obviously have to be capable of capturing signals of frequency ε^{-1} and blurring should happen on the scale of the source, so we replace $f(x)$ by $\varepsilon^{-d}f(\varepsilon^{-1}x)$. Finally, we are interested in the refocusing properties of $\mathbf{u}^B(x)$ in the vicinity of x_0 . We therefore introduce the scaling $x = x_0 + \varepsilon\xi$. With these changes of variables, expression (3.14) is recast as

$$\begin{aligned}\mathbf{u}^B(\xi; x_0) &= \Gamma \mathbf{u}(2T, x_0 + \varepsilon\xi) \\ &= \int_{\mathbb{R}^9} \Gamma G(T, x_0 + \varepsilon\xi; y) \Gamma G(T, y'; x_0 + \varepsilon z) \chi(y, y') \mathbf{S}(z) dy dy' dz,\end{aligned}\tag{3.15}$$

where

$$\chi(y, y') = \chi(y) \chi(y') f\left(\frac{y - y'}{\varepsilon}\right).\tag{3.16}$$

In the sequel we will also allow the medium to vary on a scale comparable to the source scale ε . Thus the Green's function G and the matrix A depend on ε . We do not make this dependence explicit to simplify notation. We are interested in the limit of $\mathbf{u}^B(\xi; x_0)$ as $\varepsilon \rightarrow 0$.

3.3.3 Adjoint Green's Function

The analysis of the re-propagated signal relies on the study of the two point correlation at nearby points of the Green's matrix in (3.15). There are two undesirable features in (3.15). First, the two nearby points $x_0 + \varepsilon\xi$ and $x_0 + \varepsilon z$ are terminal and initial points in their respective Green's matrices. Second, one would like the matrix Γ between the two Green's matrices to be outside of their product. However, Γ and G do not commute. For these reasons, we introduce the *adjoint* Green's matrix, solution of

$$\begin{aligned}\frac{\partial G_*(t, x; y)}{\partial t} A(x) + \frac{\partial G_*(t, x; y)}{\partial x^j} D^j &= 0 \\ G_*(0, x; y) &= A^{-1}(x) \delta(x - y).\end{aligned}\tag{3.17}$$

We now prove that

$$G_*(t, x; y) = \Gamma G(t, y; x) A^{-1}(x) \Gamma.\tag{3.18}$$

Note that for all initial data $\mathbf{S}(x)$, the solution $\mathbf{u}(t, x)$ of (3.4) satisfies

$$\mathbf{u}(t, x) = \int_{\mathbb{R}^d} G(t - s, x; y) \mathbf{u}(s, y) dy$$

for all $0 \leq s \leq t \leq T$ since the coefficients in (3.4) are time-independent. Differentiating the above with respect to s and using (3.4) yields

$$0 = \int_{\mathbb{R}^d} \left(-\frac{\partial G(t - s, x; y)}{\partial t} \mathbf{u}(s, y) - G(t - s, x; y) A^{-1}(y) D^j \frac{\partial \mathbf{u}(s, y)}{\partial y^j} \right) dy$$

Upon integrating by parts and letting $s = 0$, we get

$$0 = \int_{\mathbb{R}^d} \left(-\frac{\partial G(t, x; y)}{\partial t} + \frac{\partial}{\partial y^j} [G(t, x; y) A^{-1}(y) D^j] \right) \mathbf{S}(y) dy.$$

Since the above relation holds for all test functions $\mathbf{S}(y)$, we deduce that

$$\frac{\partial G(t, x; y)}{\partial t} - \frac{\partial}{\partial y^j} [G(t, x; y) A^{-1}(y) D^j] = 0.\tag{3.19}$$

Interchanging x and y in the above equation and multiplying it on the left and the right by Γ , we obtain that

$$\frac{\partial}{\partial t} [\Gamma G(t, y; x) A^{-1}(x)] A(x) \Gamma - \frac{\partial}{\partial x^j} [\Gamma G(t, y; x) A^{-1}(x)] D^j \Gamma = 0. \quad (3.20)$$

We remark that

$$\Gamma D^j = -D^j \Gamma \quad \text{and} \quad \Gamma A(x) = A(x) \Gamma, \quad (3.21)$$

so that

$$\frac{\partial}{\partial t} [\Gamma G(t, y; x) A^{-1}(x) \Gamma] A(x) + \frac{\partial}{\partial x^j} [\Gamma G(t, y; x) A^{-1}(x) \Gamma] D^j = 0$$

with $\Gamma G(0, y; x) A^{-1}(x) \Gamma = A^{-1}(x) \delta(x - y)$. Thus (3.18) follows from the uniqueness of the solution to the above hyperbolic system with given initial conditions. We can now recast (3.15) as

$$\begin{aligned} \mathbf{u}^B(\xi; x_0) &= \int_{\mathbb{R}^9} \Gamma G(T, x_0 + \varepsilon \xi; y) G_*(T, x_0 + \varepsilon z; y') \Gamma \\ &\quad \times \chi(y) \chi(y') f\left(\frac{y - y'}{\varepsilon}\right) A(x_0 + \varepsilon z) \mathbf{S}(z) dy dy' dz. \end{aligned} \quad (3.22)$$

One may further simplify (3.22) with the help of the auxiliary matrix-valued functions $Q(t, x; q)$ and $Q_*(t, x, q)$ defined by

$$\begin{aligned} Q(T, x; q) &= \int_{\mathbb{R}^d} G(T, x; y) \chi(y) e^{iq \cdot y / \varepsilon} dy, \\ Q_*(T, x; q) &= \int_{\mathbb{R}^3} G_*(T, x; y) \chi(y) e^{-iq \cdot y / \varepsilon} dy. \end{aligned} \quad (3.23)$$

They solve the hyperbolic systems of equations (3.4) and (3.17) with initial conditions given by $Q(0, x; q) = \chi(x) e^{iq \cdot x / \varepsilon} I$ and $Q_*(0, x; q) = A^{-1}(x) \chi(x) e^{-iq \cdot x / \varepsilon}$, respectively. Thus (3.22) becomes

$$\mathbf{u}^B(\xi; x_0) = \int_{\mathbb{R}^6} \Gamma Q(T, x_0 + \varepsilon \xi; q) Q_*(T, x_0 + \varepsilon z; q) \Gamma A(x_0 + \varepsilon z) \mathbf{S}(z) \hat{f}(q) \frac{dq dz}{(2\pi)^3}, \quad (3.24)$$

where $\hat{f}(q) = \int_{\mathbb{R}^d} e^{-iq \cdot x} f(x) dx$ is the Fourier transform of $f(x)$.

3.3.4 Wigner Transform

The back-propagated signal in (3.24) now has the suitable form to be analyzed in the Wigner transform formalism [31, 53]. We define

$$W_\varepsilon(t, x, k) = \int_{\mathbb{R}^d} \hat{f}(q) U_\varepsilon(t, x, k; q) dq, \quad (3.25)$$

where

$$U_\varepsilon(t, x, k; q) = \int_{\mathbb{R}^d} e^{ik \cdot y} Q(t, x - \frac{\varepsilon y}{2}; q) Q_*(t, x + \frac{\varepsilon y}{2}; q) \frac{dy}{(2\pi)^3}. \quad (3.26)$$

Taking the inverse Fourier transform we verify that

$$Q(t, x; q) Q_*(t, y; q) = \int_{\mathbb{R}^3} e^{-ik \cdot (y-x)/\varepsilon} U_\varepsilon(t, \frac{x+y}{2}, k; q) dk,$$

hence

$$\mathbf{u}^B(\xi; x_0) = \int_{\mathbb{R}^6} e^{ik \cdot (\xi - z)} \Gamma W_\varepsilon(T, x_0 + \varepsilon \frac{z + \xi}{2}, k) \Gamma A(x_0 + \varepsilon z) \mathbf{S}(z) \frac{dz dk}{(2\pi)^3}. \quad (3.27)$$

We have thus reduced the analysis of $\mathbf{u}(\xi; x_0)$ as $\varepsilon \rightarrow 0$ to that of the asymptotic properties of the Wigner transform W_ε . The Wigner transform has been used extensively in the study of wave propagation in random media, especially in the derivation of radiative transport equations modeling the propagation of high frequency waves. We refer to [31, 49, 53]. Note that in the usual definition of the Wigner transform, one has the adjoint matrix Q^* in place of Q_* in (3.26). This difference is not essential since Q_* and Q^* satisfy the same evolution equation, though with different initial data.

The main reason for using the Wigner transform in (3.27) is that W_ε has a weak limit W as $\varepsilon \rightarrow 0$. Its existence follows from simple a priori bounds for $W_\varepsilon(t, x, k)$. Let us introduce the space \mathcal{A} of matrix-valued functions $\phi(x, k)$ bounded in the norm $\|\cdot\|_{\mathcal{A}}$ defined by

$$\|\phi\|_{\mathcal{A}} = \int_{\mathbb{R}^3} \sup_x \|\tilde{\phi}(x, y)\| dy, \quad \text{where} \quad \tilde{\phi}(x, y) = \int_{\mathbb{R}^3} e^{-ik \cdot y} \phi(x, k) dk.$$

We denote by \mathcal{A}' its dual space, which is a space of distributions large enough to contain matrix-valued bounded measures, for instance. We then have the following result:

Lemma 3.3.1 *Let $\chi(x) \in L^2(\mathbb{R}^3)$ and $\hat{f}(q) \in L^1(\mathbb{R}^3)$. Then there is a constant $C > 0$ independent of $\varepsilon > 0$ and $t \in [0, \infty)$ such that for all $t \in [0, \infty)$, we have $\|W_\varepsilon(t, x, k)\|_{\mathcal{A}'} < C$.*

The proof of this lemma is essentially contained in [31, 49], see also [4]. One may actually get L^2 -bounds for W_ε in our setting because of the regularizing effect of \hat{f} in (3.25) but this is not essential for the purposes of this chapter as we are working on a formal level. However, this setting is one example when the mixture of states arises naturally. This is also crucial for the rigorous justification of the analog of the results of this chapter in the geometric optics regime in [6].

We therefore obtain the existence of a subsequence $\varepsilon_k \rightarrow 0$ such that W_{ε_k} converges weakly to a distribution $W \in \mathcal{A}'$. Moreover, an easy calculation shows that at time $t = 0$, we have

$$W(0, x_0, k) = |\chi(x_0)|^2 A_0^{-1}(x_0) \hat{f}(k). \quad (3.28)$$

Here, $A_0 = A$ when A is independent of ε , and $A_0 = \lim_{\varepsilon \rightarrow 0} A_\varepsilon$ if we assume that the family of matrices $A_\varepsilon(x)$ is uniformly bounded and continuous with the limit A_0 in $\mathcal{C}(\mathbb{R}^d)$. These assumptions on A_ε are sufficient to deal with the radiative transport regime we will consider in section 3.3.7. Under the same assumptions on A_ε , we have the following result.

Proposition 3.3.2 *The back-propagated signal $\mathbf{u}^B(\xi; x_0)$ given by (3.27) converges weakly in $\mathcal{S}'(\mathbb{R}^3 \times \mathbb{R}^3)$ as $\varepsilon \rightarrow 0$ to the limit*

$$\mathbf{u}^B(\xi; x_0) = \int_{\mathbb{R}^6} e^{ik \cdot (\xi - z)} \Gamma W(T, x_0, k) \Gamma A_0(x_0) \mathbf{S}(z) \frac{dz dk}{(2\pi)^3}. \quad (3.29)$$

The proof of this proposition is based on taking the duality product of $\mathbf{u}^B(\xi; x_0)$ with a vector-valued test function $\phi(\xi; x_0)$ in $\mathcal{S}(\mathbb{R}^3 \times \mathbb{R}^3)$. After a change of variables we obtain $\langle \mathbf{u}^B, \phi \rangle = \langle W_\varepsilon, Z_\varepsilon \rangle$. Here the duality product for matrices is given by the trace $\langle A, B \rangle = \sum_{i,k} \langle A_{ik}, B_{ik} \rangle$, and

$$Z_\varepsilon(x_0, k) = \int_{\mathbb{R}^6} e^{ik \cdot (z - \xi)} \Gamma \phi(\xi, x_0 - \varepsilon \frac{z + \xi}{2}) \mathbf{S}^*(z) A_\varepsilon(x_0 + \varepsilon \frac{z - \xi}{2}) \Gamma \frac{dz d\xi}{(2\pi)^3}. \quad (3.30)$$

Defining Z as the limit of Z_ε as $\varepsilon \rightarrow 0$ by replacing formally ε by 0 in the above expression, (3.29) follows from showing that $\|Z_\varepsilon - Z\|_{\mathcal{A}} \rightarrow 0$ as $\varepsilon \rightarrow 0$. This is straightforward and we omit the details.

The above proposition tells us how to reconstruct the back-propagated solution in the high frequency limit from the limit Wigner matrix W . Notice that we have made almost no assumptions on the medium described by the matrix $A_\varepsilon(x)$. At this level, the medium can be either homogeneous or heterogeneous, and the particular scale of oscillations is not important as long as $A_\varepsilon(x)$ strongly converge to A_0 . Without any further assumptions, we can also obtain some information about the matrix W . Let us define the dispersion matrix for the system (3.4) as [53]

$$L(x, k) = A_0^{-1}(x)k_j D^j. \quad (3.31)$$

It is given explicitly by

$$L(x, k) = \begin{pmatrix} 0 & 0 & 0 & k_1/\rho(x) \\ 0 & 0 & 0 & k_2/\rho(x) \\ 0 & 0 & 0 & k_3/\rho(x) \\ k_1/\kappa(x) & k_2/\kappa(x) & k_3/\kappa(x) & 0 \end{pmatrix}.$$

The matrix L has a double eigenvalue $\omega_0 = 0$ and two simple eigenvalues $\omega_\pm(x, k) = \pm c(x)|k|$, where $c(x) = 1/\sqrt{\rho(x)\kappa(x)}$ is the speed of sound. The eigenvalues ω_\pm are associated with eigenvectors $\mathbf{b}_\pm(x, k)$ and the eigenvalue $\omega_0 = 0$ is associated with the eigenvectors $\mathbf{b}_j(x, k)$, $j = 1, 2$. They are given by

$$\mathbf{b}_\pm(x, k) = \begin{pmatrix} \pm \frac{\hat{k}}{\sqrt{2\rho(x)}} \\ 1 \\ \sqrt{2\kappa(x)} \end{pmatrix}, \quad \mathbf{b}_j(x, k) = \begin{pmatrix} \frac{\mathbf{z}^j(k)}{\sqrt{\rho(x)}} \\ 0 \end{pmatrix}, \quad (3.32)$$

where $\hat{k} = k/|k|$ and $z^1(k)$ and $z^2(k)$ are chosen so that the triple $(\hat{k}, z^1(k), z^2(k))$ forms an orthonormal basis. The eigenvectors are normalized so that

$$(A_0(x)\mathbf{b}_j(x, k) \cdot \mathbf{b}_k(x, k)) = \delta_{jk}, \quad (3.33)$$

for all $j, k \in J = \{+, -, 1, 2\}$. The space of 4×4 matrices is clearly spanned by the basis $\mathbf{b}_j \otimes \mathbf{b}_k$. We then have the following result:

Proposition 3.3.3 *There exist scalar distributions a_\pm and a_0^{mn} , $m, n = 1, 2$ so that the limit Wigner distribution matrix can be decomposed as*

$$W(t, x, k) = \sum_{j,m=1}^2 a_0^{jm}(t, x, k) \mathbf{b}_j(x, k) \otimes \mathbf{b}_m(x, k) + a_+(t, x, k) \mathbf{b}_+(x, k) \otimes \mathbf{b}_+(x, k) + a_-(t, x, k) \mathbf{b}_-(x, k) \otimes \mathbf{b}_-(x, k). \quad (3.34)$$

The main result of this proposition is that the cross terms $\mathbf{b}_j \otimes \mathbf{b}_k$ with $\omega_j \neq \omega_k$ do not contribute to the limit W . The proof of this proposition can be found in [31] and a formal derivation in [53].

The initial conditions for the amplitudes a_j are calculated using the identity

$$A_0^{-1}(x) = \sum_{j \in J} \mathbf{b}_j(x, k) \otimes \mathbf{b}_j(x, k).$$

Then (3.28) implies that $a_0^{12}(0, x, k) = a_0^{21}(0, x, k) = 0$ and

$$a_0^{jj}(0, x, k) = a_\pm(0, x, k) = |\chi(x)|^2 f(k), \quad j = 1, 2. \quad (3.35)$$

3.3.5 Mode Decomposition and Refocusing

We can use the above result to recast (3.29) as

$$\mathbf{u}^B(\xi; x_0) = (F(T, \cdot; x_0) * \mathbf{S})(\xi), \quad (3.36)$$

where

$$\begin{aligned} F(T, \xi; x_0) = & \sum_{m,n=1}^2 \int_{\mathbb{R}^3} e^{ik \cdot \xi} a_0^{mn}(T, x_0; k) \Gamma \mathbf{b}_m(x_0, k) \otimes \mathbf{b}_n(x_0, k) A_0(x_0) \Gamma \frac{dk}{(2\pi)^3} \\ & + \int_{\mathbb{R}^3} e^{ik \cdot \xi} a_+(T, x_0; k) \Gamma \mathbf{b}_+(x_0, k) \otimes \mathbf{b}_+(x_0, k) A_0(x_0) \Gamma \frac{dk}{(2\pi)^3} \\ & + \int_{\mathbb{R}^3} e^{ik \cdot \xi} a_-(T, x_0; k) \Gamma \mathbf{b}_-(x_0, k) \otimes \mathbf{b}_-(x_0, k) A_0(x_0) \Gamma \frac{dk}{(2\pi)^3}. \end{aligned} \quad (3.37)$$

This expression can be used to assess the quality of the refocusing. When $F(T, \xi; x_0)$ has a narrow support in ξ , refocusing is good. When its support in ξ grows larger, its quality degrades. The spatial decay of the kernel $F(t, \xi; x_0)$ in ξ is directly related to the smoothness in k of its Fourier transform in ξ :

$$\begin{aligned} \hat{F}(T, k; x_0) = & \sum_{m,n=1}^2 a_0^{mn}(T, x_0; k) \Gamma \mathbf{b}_m(x_0, k) \otimes \mathbf{b}_n(x_0, k) A_0(x_0) \Gamma \frac{dk}{(2\pi)^3} \\ & + \Gamma [a_+(T, x_0; k) \mathbf{b}_+(x_0, k) \otimes \mathbf{b}_+(x_0, k) + a_-(T, x_0; k) \mathbf{b}_-(x_0, k) \otimes \mathbf{b}_-(x_0, k)] A_0(x_0) \Gamma. \end{aligned}$$

Namely, for F to decay in ξ , one needs $\hat{F}(k)$ to be smooth in k . However, the eigenvectors b_j are singular at $k = 0$ as can be seen from the explicit expressions (3.32). Therefore, a priori \hat{F} is not smooth at $k = 0$. This means that in order to obtain good refocusing one needs the original signal to have no low frequencies: $\hat{S}(k) = 0$ near $k = 0$. Low frequencies in the initial data will not refocus well.

We can further simplify (3.36)-(3.37) if we assume that the initial condition is irrotational. Taking Fourier transform of both sides in (3.36), we obtain that

$$\hat{\mathbf{u}}^B(k; x_0) = \sum_{j,n \in J} a_j(T, x_0, k) \hat{S}_n(k) (A_0(x_0) \Gamma b_n(x_0, k) \cdot b_j(x_0, k)) \Gamma b_j(x_0, k) \quad (3.38)$$

where we have defined

$$\hat{\mathbf{S}}(k) = \sum_{n \in J} \hat{S}_n(k) \mathbf{b}_n(x_0, k). \quad (3.39)$$

Irrotationality of the initial condition means that \hat{S}_1 and \hat{S}_2 identically vanish, or equivalently that

$$\mathbf{S}(x) = \begin{pmatrix} \nabla \phi(x) \\ p(x) \end{pmatrix} \quad (3.40)$$

for some pressure $p(x)$ and potential $\phi(x)$. Remarking that $\Gamma \mathbf{b}_{\pm} = -\mathbf{b}_{\mp}$ and by irrotationality that $(A_0(x_0) \hat{\mathbf{S}}(k) \cdot \mathbf{b}_{1,2}(k)) = 0$, we use (3.33) to recast (3.38) as

$$\hat{\mathbf{u}}^B(k; x_0) = a_-(T, x_0, k) \hat{S}_+(k) \mathbf{b}_+(x_0, k) + a_+(T, x_0, k) \hat{S}_-(k) \mathbf{b}_-(x_0, k). \quad (3.41)$$

Decomposing the initial condition $\mathbf{S}(x)$ as

$$\mathbf{S}(x) = \mathbf{S}_+(x) + \mathbf{S}_-(x), \quad \text{such that} \quad \hat{\mathbf{S}}_{\pm}(k) = \hat{S}_{\pm}(k) \mathbf{b}_{\pm}(x_0, k),$$

the back-propagated signal takes the form

$$\mathbf{u}^B(\xi; x_0) = (\hat{a}_-(T, x_0, \cdot) * \mathbf{S}_+(\cdot))(\xi) + (\hat{a}_+(T, x_0, \cdot) * \mathbf{S}_-(\cdot))(\xi) \quad (3.42)$$

where \hat{a}_\pm is the Fourier of a_\pm in k . This form is much more tractable than (3.36)-(3.37). It is also almost as general. Indeed, rotational modes do not propagate in the high frequency regime. Therefore, they are exactly back-propagated when $\chi(x_0) = 1$ and $f(x) = \delta(x)$, and not back-propagated at all when $\chi(x_0) = 0$. All the refocusing properties are thus captured by the amplitudes $a_\pm(T, x_0, k)$. Their evolution equation characterizes how waves propagate in the medium and their initial conditions characterize the recording array.

3.3.6 Homogeneous Media

In homogeneous media with $c(x) = c_0$ the amplitudes $a_\pm(T, x, k)$ satisfy the free transport equation [31, 53]

$$\frac{\partial a_\pm}{\partial t} \pm c_0 \hat{k} \cdot \nabla_x a_\pm = 0 \quad (3.43)$$

with initial data $a_\pm(0, x, k) = |\chi(x)|^2 f(k)$ as in (3.35). They are therefore given by

$$a_\pm(t, x_0, k) = |\chi(x_0 \mp c_0 \hat{k} t)|^2 \hat{f}(k). \quad (3.44)$$

These amplitudes become more and more singular in k as time grows since their gradient in k grows linearly with time. The corresponding kernel $F = F_H$ decays therefore more slowly in ξ as time grows. This implies that the quality of the refocusing degrades with time. For sufficiently large times, all the energy has left the domain Ω (assumed to be bounded), and the coefficients $a_\pm(t, x_0, k)$ vanish. Therefore the back-propagated signal $\mathbf{u}^B(\xi; x_0)$ also vanishes, which means that there is no refocusing at all. The same conclusions could also be drawn by analyzing (3.14) directly in a homogeneous medium. This is the situation in the numerical experiment presented in Fig. 3.3: in a homogeneous medium, the back-propagated signal would vanish.

3.3.7 Heterogeneous Media and Radiative Transport Regime

The results of the preceding sections show how the back-propagated signal $\mathbf{u}^B(\xi; x_0)$ is related to the propagating modes $a_\pm(T, x_0, k)$ of the Wigner matrix $W(T, x_0, k)$. The form assumed by the modes $a_\pm(T, x_0, k)$, and in particular their smoothness in k , will depend on the hypotheses we make on the underlying medium; i.e., on the density $\rho(x)$ and compressibility $\kappa(x)$ that appear in the matrix $A(x)$. We have seen that partial measurements in homogeneous media yield poor refocusing properties. We now show that refocusing is much better in random media.

We consider here the radiative transport regime, also known as weak coupling limit. There, the fluctuations in the physical parameters are weak and vary on a scale comparable to the scale of the initial condition. Density and compressibility assume the form

$$\rho(x) = \rho_0 + \sqrt{\varepsilon} \rho_1\left(\frac{x}{\varepsilon}\right) \quad \text{and} \quad \kappa(x) = \kappa_0 + \sqrt{\varepsilon} \kappa_1\left(\frac{x}{\varepsilon}\right). \quad (3.45)$$

The functions ρ_1 and κ_1 are assumed to be mean-zero spatially homogeneous processes. The average (with respect to realizations of the medium) of the propagating amplitudes a_\pm , denoted by \bar{a}_\pm , satisfy in the high frequency limit $\varepsilon \rightarrow 0$ a radiative transfer equation (RTE), which is a linear Boltzmann equation of the form

$$\begin{aligned} \frac{\partial \bar{a}_\pm}{\partial t} \pm c_0 \hat{k} \cdot \nabla_x \bar{a}_\pm &= \int_{\mathbb{R}^3} \sigma(k, p) (\bar{a}_\pm(t, x, p) - \bar{a}_\pm(t, x, k)) \delta(c_0(|k| - |p|)) dp \\ \bar{a}_\pm(0, x, k) &= |\chi(x)|^2 \hat{f}(k). \end{aligned} \quad (3.46)$$

The scattering coefficient $\sigma(k, p)$ depends on the power spectra of ρ_1 and κ_1 . We refer to [53] for the details of the derivation and explicit form of $\sigma(k, p)$. The above result remains formal for the wave equation and requires averaging over the realizations of the random medium although this is not necessary in the physical and numerical time reversal experiments. A rigorous derivation of the linear Boltzmann equation (which also requires averaging over realizations) has only been obtained for the Schrödinger equation; see [23, 56]. Nevertheless, the above result formally characterizes the filter $F(T, \xi; x_0)$ introduced in (3.37) and (3.42).

The transport equation (3.46) has a smoothing effect best seen in its integral formulation. Let us define the total scattering coefficient $\Sigma(k) = \int_{\mathbb{R}^3} \sigma(k, p) \delta(c_0(|k| - |p|)) dp$. Then the transport equation (3.46) may be rewritten as

$$\begin{aligned} \bar{a}_{\pm}(t, x, k) &= \bar{a}_{\pm}(0, x \mp c_0 \hat{k}t, k) e^{-\Sigma(k)t} \\ &+ \frac{|k|^2}{c_0} \int_0^t ds \int_{S^2} \sigma(k, |k| \hat{\mathbf{p}}) \bar{a}_{\pm}(s, x \mp c_0(t-s) \hat{k}, |k| \hat{\mathbf{p}}) e^{-\Sigma(k)(t-s)} d\Omega(\hat{\mathbf{p}}). \end{aligned} \quad (3.47)$$

Here $\hat{\mathbf{p}} = p/|p|$ is the unit vector in direction of p and $d\Omega(\hat{\mathbf{p}})$ is the surface element on the sphere S^2 . The first term in (3.47) is the ballistic part that undergoes no scattering. It has no smoothing effect, and, moreover, if $a(0, x, k)$ is not smooth in x , as may be the case for (3.35), the discontinuities in x translate into discontinuities in k at later times as in (3.44) in a homogeneous medium. However, in contrast to the homogeneous medium case, the ballistic term decays exponentially in time, and does not affect the refocused signal for sufficiently long times $t \gg 1/\Sigma$. The second term in (3.47) exhibits a smoothing effect. Namely the operator $\mathcal{L}g$ defined by

$$\mathcal{L}g(t, x, k) = \frac{|k|^2}{c_0} \int_0^t ds \int_{S^2} \sigma(k, |k| \hat{\mathbf{p}}) g(s, x \mp c_0(t-s) \hat{k}, |k| \hat{\mathbf{p}}) e^{-\Sigma(k)(t-s)} d\Omega(\hat{\mathbf{p}})$$

is regularizing, in the sense that the function $\tilde{g} = \mathcal{L}g$ has at least $1/2$ -more derivatives than g (in some Sobolev scale). The precise formulation of this smoothing property is given by the averaging lemmas [34, 51] and will not be dwelt upon here. Iterating (3.47) n times we obtain

$$\bar{a}_{\pm}(t, x, k) = a_{\pm}^0(t, x, k) + a_{\pm}^1(t, x, k) + \cdots + a_{\pm}^n(t, x, k) + \mathcal{L}^{n+1} \bar{a}_{\pm}(t, x, k). \quad (3.48)$$

The terms $a_{\pm}^0, \dots, a_{\pm}^n$ are given by

$$a_{\pm}^0(t, x, k) = \bar{a}_{\pm}(0, x \mp c_0 \hat{k}t, k) e^{-\Sigma(k)t}, \quad a_{\pm}^j(t, x, k) = \mathcal{L} a_{\pm}^{j-1}(t, x, k).$$

They describe, respectively, the contributions from waves that do not scatter, scatter once, twice, \dots . It is straightforward to verify that all these terms decay exponentially in time and are negligible for times $t \gg 1/\Sigma$. The last term in (3.48) has at least $n/2$ more derivatives than the initial data a_0 , or the solution (3.44) of the homogeneous transport equation. This leads to a faster decay in ξ of the Fourier transforms $\hat{a}_{\pm}(T, x_0, \xi)$ of $a_{\pm}(T, x_0, k)$ in k . This gives a qualitative explanation as to why refocusing is better in heterogeneous media than in homogeneous media. A more quantitative answer requires to solve the transport equation (3.46).

3.3.8 Diffusion Regime

It is known for times t much longer than the scattering mean free time $\tau_{sc} = 1/\Sigma$ and distances of propagation L very large compared to $l_{sc} = c_0 \tau_{sc}$ that solutions to the radiative transport equation (3.46) can be approximated by solutions to a diffusion equation, provided that $c(x) =$

c_0 is independent of x [19, 48]. More precisely, we let $\delta = l_{sc}/L \ll 1$ be a small parameter and rescale time and space variables as $t \rightarrow t/\delta^2$ and $x \rightarrow x/\delta$. In this limit, the wave direction is completely randomized so that

$$\bar{a}_+(t, x, k) \approx \bar{a}_-(t, x, k) \approx a(t, x, |k|),$$

where a solves

$$\begin{aligned} \frac{\partial a(t, x, |k|)}{\partial t} - D(|k|)\Delta_x a(t, x, |k|) &= 0, \\ a(0, x, |k|) &= |\chi(x)|^2 \frac{1}{4\pi|k|^2} \int_{\mathbb{R}^3} \hat{f}(q) \delta(|q| - |k|) dq. \end{aligned} \quad (3.49)$$

The diffusion coefficient $D(|k|)$ may be expressed explicitly in terms of the scattering coefficient $\sigma(k, p)$ and hence related to the power spectra of ρ_1 and κ_1 . We refer to [53] for the details. For instance, let us assume for simplicity that the density is not fluctuating, $\rho_1 \equiv 0$, and that the compressibility fluctuations are delta-correlated, so that $\mathbb{E}\{\hat{\kappa}_1(p)\hat{\kappa}_1(q)\} = \kappa_0^2 \hat{R}_0 \delta(p + q)$. Then we have

$$\sigma(k, p) = \frac{\pi c_0^2 |k|^2 \hat{R}_0}{2}, \quad \Sigma(|k|) = 2\pi^2 c_0 |k|^4 \hat{R}_0 \quad (3.50)$$

and

$$D(|k|) = \frac{c_0^2}{3\Sigma(|k|)} = \frac{c_0}{6\pi^2 |k|^4 \hat{R}_0} \quad (3.51)$$

Let us assume that there are no initial rotational modes, so that the source $\mathbf{S}(x)$ is decomposed as in (3.40). Using (3.41), we obtain that

$$\hat{\mathbf{u}}^B(k; x_0) = a(T, x_0, |k|) \hat{\mathbf{S}}(k). \quad (3.52)$$

When $f(x)$ is isotropic so that $\hat{f}(k) = \hat{f}(|k|)$, and the diffusion coefficient is given by (3.51), the solution of (3.49) takes the form

$$a(T, x_0, |k|) = \hat{f}(|k|) \left(\frac{3\pi |k|^4 \hat{R}_0}{2c_0 T} \right)^{3/2} \int_{\mathbb{R}^3} \exp\left(-\frac{3\pi^2 |k|^4 \hat{R}_0 |x_0 - y|^2}{2c_0 T}\right) |\chi(y)|^2 dy. \quad (3.53)$$

When $f(x) = \delta(x)$, and $\Omega = \mathbb{R}^3$, so that $\chi(x) \equiv 1$, we retrieve $a(T, x_0, k) \equiv 1$, hence the refocusing is perfect. When only partial measurement is available, the above formula indicates how the frequencies of the initial pulse are filtered by the one-step time reversal process. Notice that both the low and high frequencies are damped. The reason is that low frequencies scatter little from the underlying medium so that it takes a long time for them to be randomized. High frequencies strongly scatter with the underlying medium and consequently propagate little so that the signal that reaches the recording array Ω is small unless recorders are also located at the source point: $x_0 \in \Omega$. In the latter case they are very well measured and back-propagated although this situation is not the most interesting physically. Expression (3.53) may be generalized to other power spectra of medium fluctuations in a straightforward manner using the formula for the diffusion coefficient in [53].

3.3.9 Numerical Results

The numerical results in Fig. 3.3 show that some signal refocuses at the location of the initial source after the time reversal procedure. Based on the above theory however, we do not expect the refocused signal to have exactly the same shape as the original one. Since the location of the initial source belongs to the recording array ($\chi(x_0) = 1$) in our simulations, we expect from our theory that high frequencies will refocus well but that low frequencies will not. This

Figure 3.4: Zoom of the initial source and the refocused signal for the numerical experiment of Fig. 3.3.

is confirmed by the numerical results in Fig. 3.4, where a zoom in the vicinity of $x_0 = \mathbf{0}$ of the initial source and refocused signal are represented. Notice that the numerical simulations are presented here only to help in the understanding of the refocusing theory and do not aim at reproducing the theory in a quantitative manner. The random fluctuations are quite strong in our numerical simulations and it is unlikely that the diffusive regime will be valid. The refocused signal on the right figure looks however like a high-pass filter of the signal in the left figure, as expected from theory.

3.4 Refocusing of Classical Waves

The theory presented in section 3.3 provides a quantitative explanation for the results observed in time reversal physical and numerical experiments. However, the time reversal procedure is by no means necessary to obtain refocusing. Time reversal is associated with the specific choice (3.12) for the matrix Γ in the preceding section, which reverses the direction of the acoustic velocity and keeps pressure unchanged. Other choices for Γ are however possible. When nothing is done at time T , i.e., when we choose $\Gamma = I$, no refocusing occurs as one might expect. It turns out that $\Gamma = I$ is more or less the only choice of a matrix that prevents some sort of refocusing. Section 3.4.1 presents the theory of refocusing for acoustic waves, which is corroborated by numerical results presented in Section 3.4.2. Sections 3.4.3 and 3.4.4 generalize the theory to other linear hyperbolic systems.

3.4.1 General Refocusing of Acoustic Waves

In one-step time reversal, the action of the “intelligent” array is captured by the choice of the signal processing matrix Γ in (3.13). Time reversal is characterized by Γ given in (3.12). A passive array is characterized by $\Gamma = I$. This section analyzes the role of other choices for Γ , which we let depend on the receiver location so that each receiver may perform its own kind of signal processing.

The signal after time reversal is still given by (3.13), where $\Gamma(y')$ is now arbitrary. At time $2T$, after back-propagation, we are free to multiply the signal by an arbitrary invertible matrix to analyze the signal. It is convenient to multiply the back-propagated signal by the matrix $\Gamma_0 = \text{Diag}(-1, -1, -1, 1)$ as in classical time reversal. The reconstruction formula (3.15) in

the localized source limit is then replaced by

$$\mathbf{u}^B(\xi; x_0) = \int_{\mathbb{R}^9} \Gamma_0 G(T, x_0 + \varepsilon \xi; y) \Gamma(y') G(T, y'; x_0 + \varepsilon z) \chi(y, y') \mathbf{S}(z) dy dy' dz \quad (3.54)$$

with $\chi(y, y')$ defined by (3.16). To generalize the results of section 3.3, we need to define an appropriate adjoint Green's matrix G_* . As before, this will allow us to remove the matrix Γ between the two Green's matrices in (3.54) and to interchange the order of points in the second Green's matrix. We define the new adjoint Green's function $G_*(t, x; y)$ as the solution to

$$\frac{\partial G_*(t, x; y)}{\partial t} A(x) + \frac{\partial G_*(t, x; y)}{\partial x^j} D^j = 0 \quad (3.55)$$

$$G_*(0, x; y) = \Gamma(x) \Gamma_0 A^{-1}(x) \delta(x - y).$$

Following the steps of section 3.3.3, we show that

$$G_*(t, x, y) = \Gamma(y) G(t, y; x) A^{-1}(x) \Gamma_0. \quad (3.56)$$

The only modification compared to the corresponding derivation of (3.18) is to multiply (3.19) on the left by $\Gamma(x)$ and on the right by Γ_0 so that $\Gamma(y)$ appears on the left in (3.20). The re-transmitted signal may now be recast as

$$\mathbf{u}^B(\xi; x_0) = \int_{\mathbb{R}^9} \Gamma_0 G(T, x_0 + \varepsilon \xi; y) G_*(T, x_0 + \varepsilon z; y') \Gamma_0^{-1} A(x_0 + \varepsilon z) \chi(y, y') \mathbf{S}(z) dy dy' dz. \quad (3.57)$$

Therefore the only modification in the expression for the re-transmitted signal compared to the time reversed signal (3.22) is in the initial data for (3.55), which is the only place where the matrix $\Gamma(x)$ appears.

The analysis in Sections 3.3.3-3.3.7 requires only minor changes, which we now outline. The back-propagated signal may still be expressed in term of the Wigner distribution (compare to (3.27))

$$\mathbf{u}^B(\xi; x_0) = \int_{\mathbb{R}^6} e^{ik \cdot (\xi - z)} \Gamma_0 W_\varepsilon(T, x_0 + \varepsilon \frac{z + \xi}{2}, k) \Gamma_0 A(x_0 + \varepsilon z) \mathbf{S}(z) \frac{dz dk}{(2\pi)^3}. \quad (3.58)$$

The Wigner distribution is defined as before by (3.25) and (3.26). The function Q is defined as before as the solution of (3.4) with initial data $Q(0, x; q) = \chi(x) e^{iq \cdot x / \varepsilon} I$, while Q_* solves (3.17) with the initial data $Q_*(0, x; q) = \Gamma(x) \Gamma_0 A^{-1}(x) \chi(x) e^{-iq \cdot x / \varepsilon}$. The initial Wigner distribution is now given by

$$W(0, x, k) = |\chi(x)|^2 \Gamma(x) \Gamma_0 A^{-1}(x) \hat{f}(k). \quad (3.59)$$

Lemma 3.3.1 and Proposition 3.3.2 also hold, and we obtain the analog of (3.29)

$$\mathbf{u}(\xi; x_0) = \int_{\mathbb{R}^6} e^{ik \cdot (\xi - z)} \Gamma_0 W(T, x_0, k) \Gamma_0 A_0(x_0) \mathbf{S}(z) dz dk. \quad (3.60)$$

The limit Wigner distribution $W(T, x_0, k)$ admits the mode decomposition (3.34) as before. If we assume that the source $\mathbf{S}(x)$ has the form (3.40) so that no rotational modes are present initially, we recover the refocusing formula (3.41):

$$\hat{\mathbf{u}}^B(k; x_0) = a_-(T, x_0, k) \hat{S}_+(k) b_+(x_0, k) + a_+(T, x_0, k) \hat{S}_-(k) b_-(x_0, k). \quad (3.61)$$

The initial conditions for the amplitudes a_{\pm} are replaced by

$$\begin{aligned} a_{\pm}(0, x, k) &= \text{Tr} [A_0(x)W(0, x, k)A_0(x)b_{\pm}(x_0, k)b_{\pm}^*(x_0, k)] \\ &= |\chi(x)|^2 \hat{f}(k)(A_0(x)\Gamma(x)b_{\mp}(x, k) \cdot b_{\pm}(x, k)). \end{aligned} \quad (3.62)$$

Observe that when $\Gamma(x) = \Gamma_0$, we get back the results of Section 3.3.7. When the signal is not changed at the array, so that $\Gamma = I$, the coefficients $a_{\pm}(0, x, k) \equiv 0$ by orthogonality (3.33) of the eigenvectors \mathbf{b}_j . We thus obtain that no refocusing occurs when the “intelligent” array is replaced by a passive array, as expected physically.

Another interesting example is when only pressure p is measured, so that the matrix $\Gamma = \text{Diag}(0, 0, 0, 1)$. Then the initial data is

$$a_{\pm}(0, x, k) = \frac{1}{2}|\chi(x)|^2 \hat{f}(k),$$

which differs by a factor $1/2$ from the full time reversal case (3.35). Therefore the re-transmitted signal \mathbf{u}^B also differs only by a factor $1/2$ from the latter case, and the quality of refocusing as well as the shape of the re-propagated signal are exactly the same. The same observation applies to the measurement and reversal of the acoustic velocity only, which corresponds to the matrix $\Gamma = \text{Diag}(-1, -1, -1, 0)$. The factor $1/2$ comes from the fact that only the potential energy or the kinetic energy is measured in the first and second cases, respectively. For high frequency acoustic waves, the potential and kinetic energies are equal, hence the factor $1/2$. We can also verify that when only the first component of the velocity field is measured so that $\Gamma = \text{Diag}(-1, 0, 0, 0)$, the initial data is

$$a_{\pm}(0, x, k) = |\chi(x)|^2 \hat{f}(k) \frac{k_1^2}{2|k|^2}. \quad (3.63)$$

As in the time reversal setting of Section 3.3, the quality of the refocusing is related to the smoothness of the amplitudes a_{\pm} in k . In a homogeneous medium they satisfy the free transport equation (3.43), and are given by

$$a_{\pm}(t, x, k) = |\chi(x - c_0 \hat{k}t)|^2 \hat{f}(k)(A_0(x - c_0 \hat{k}t)\Gamma(x - c_0 \hat{k}t)b_{\mp}(x - c_0 \hat{k}t, k) \cdot b_{\pm}(x - c_0 \hat{k}t, k)).$$

Once again, we observe that in a uniform medium a_{\pm} become less regular in k as time grows, thus refocusing is poor.

The considerations of Section 3.3.7 show that in the radiative transport regime the amplitudes a_{\pm} become smoother in k also with initial data given by (3.62). This leads to a better refocusing as explained in Section 3.3.5. Let us assume that the diffusion regime of Section 3.3.8 is valid and that the kernel f is isotropic $\hat{f}(k) = \hat{f}(|k|)$. This requires in particular that $A_0(x)$ be independent of x . We obtain that $a_{\pm}(T, x_0, k) = \tilde{a}(T, x_0, |k|)$, thus the refocusing formula (3.61) reduces to

$$\hat{\mathbf{u}}^B(k; x_0) = \tilde{a}(T, x_0, |k|)\hat{\mathbf{S}}(k). \quad (3.64)$$

The difference with the case treated in Section 3.3.8 is that $\tilde{a}(T, x, |k|)$ solves the diffusion equation (3.49) with new initial conditions given by

$$\begin{aligned} \tilde{a}(0, x, |k|) &= \frac{|\chi(x)|^2}{4\pi|k|^2} \int_{\mathbb{R}^3} \hat{f}(|q|)(A_0\Gamma(x)\mathbf{b}_{-}(q) \cdot \mathbf{b}_{+}(q))\delta(|q| - |k|)dq \\ &= \frac{|\chi(x)|^2}{4\pi|k|^2} \int_{\mathbb{R}^3} \hat{f}(|q|)(A_0\Gamma(x)\mathbf{b}_{+}(q) \cdot \mathbf{b}_{-}(q))\delta(|q| - |k|)dq. \end{aligned} \quad (3.65)$$

When only the first component of the velocity field is measured, as in (3.63), the initial data for \tilde{a} is

$$\tilde{a}(0, x, |k|) = \frac{1}{6} |\chi(x)|^2 \hat{f}(|k|).$$

Therefore even time reversing only one component of the acoustic velocity field produces a re-propagated signal that is equal to the full re-propagated field up to a constant factor.

More generally, we deduce from (3.65) that a detector at x will contribute some refocusing for waves with wavenumber $|k|$ provided that

$$\int_{S^2} \hat{f}(|k|\hat{q})(A_0\Gamma(x)\mathbf{b}_{\mp}(\hat{q}) \cdot \mathbf{b}_{\pm}(\hat{q}))d\Omega(\hat{q}) \neq 0.$$

When $f(x) = f(|x|)$ is radial, this property becomes independent of the wavenumber $|k|$ and reduces to $\int_{S^2} (A_0\Gamma(x)\mathbf{b}_{\mp}(\hat{q}) \cdot \mathbf{b}_{\pm}(\hat{q}))d\Omega(\hat{q}) \neq 0$.

3.4.2 Numerical Results

Let us come back to the numerical results presented in Fig. 3.3 and 3.4. We now consider two different processings at the recording array. The first array is passive, corresponding to $\Gamma = I$, and the second array only measures pressure so that $\Gamma = \text{Diag}(0, 0, 0, 1)$. The zoom in the vicinity of $x_0 = \mathbf{0}$ of the “refocused” signals is given in Fig. 3.5. The left figure shows

Figure 3.5: Zoom of the refocused signals for the numerical experiment of Fig. 3.3 with processing $\Gamma = I$ (left), with a maximal amplitude of roughly $4 \cdot 10^{-3}$ and $\Gamma = \text{Diag}(0, 0, 0, 1)$ (right), with a maximal amplitude of roughly 0.035.

no refocusing, in accordance with physical intuition and theory. The right figure shows that refocusing indeed occurs when only pressure is recorded (and its time derivative is set to 0 in the solution of the wave equation presented in the appendix). Notice also that the refocused signal is roughly one half the one obtained in Fig. 3.4 as predicted by theory.

3.4.3 Refocusing of Other Classical Waves

The preceding sections deal with the refocusing of acoustic waves. The theory can however be extended to more complicated linear hyperbolic systems of the form (3.4) with $A(x)$ a positive definite matrix, D^j symmetric matrices, and $\mathbf{u} \in \mathbb{C}^m$. These include electromagnetic and

elastic waves. Their explicit representation in the form (3.4) and expressions for the matrices $A(x)$ and D^j in these cases may be found in [53]. For instance, the Maxwell equations

$$\begin{aligned}\frac{\partial E}{\partial t} &= \frac{1}{\epsilon(x)} \operatorname{curl} H \\ \frac{\partial H}{\partial t} &= -\frac{1}{\mu(x)} \operatorname{curl} E\end{aligned}$$

may be written in the form (3.4) with $\mathbf{u} = (E, H) \in \mathbb{C}^6$ and the matrix

$$A(x) = \operatorname{Diag}(\epsilon(x), \epsilon(x), \epsilon(x), \mu(x), \mu(x), \mu(x)).$$

Here $\epsilon(x)$ is the dielectric constant (not to be confused with the small parameter ε), and $\mu(x)$ is the magnetic permeability. The 6×6 dispersion matrix $L(x, k)$ for the Maxwell equations is given by

$$L(x, k) = - \begin{pmatrix} 0 & 0 & 0 & 0 & -k_3/\epsilon(x) & k_2/\epsilon(x) \\ 0 & 0 & 0 & k_3/\epsilon(x) & 0 & -k_1/\epsilon(x) \\ 0 & 0 & 0 & -k_2/\epsilon(x) & k_1/\epsilon(x) & 0 \\ 0 & k_3/\mu(x) & -k_2/\mu(x) & 0 & 0 & 0 \\ -k_3/\mu(x) & 0 & k_1/\mu(x) & 0 & 0 & 0 \\ k_2/\mu(x) & -k_1/\mu(x) & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Generalization of our results for acoustic waves to such general systems is quite straightforward so we concentrate only on the modifications that need be made. The time reversal procedure is exactly the same as before: a signal propagates from a localized source, is recorded, processed as in (3.13) with a general matrix $\Gamma(y')$, and re-emitted into the medium. The re-transmitted signal is given by (3.54). Furthermore, the equation for the adjoint Green's matrix (3.55), the definition of the Wigner transform in Section 3.3.4, and the expression (3.60) for the re-propagated signal still hold.

The analysis of the re-propagated signal is reduced to the study of the Wigner distribution, which is now modified. The mode decomposition must be generalized. We recall that

$$L(x, k) = A_0^{-1}(x) k_j D^j$$

is the $m \times m$ dispersion matrix associated with the hyperbolic system (3.4). Since $L(x, k)$ is symmetric with respect to the inner product $\langle \mathbf{u}, \mathbf{v} \rangle_{A_0} = (A_0 \mathbf{u} \cdot \mathbf{v})$, its eigenvalues are real and its eigenvectors form a basis. We assume the existence of a time reversal matrix Γ_0 such that (3.21) holds with $\Gamma = \Gamma_0$ and such that $\Gamma_0^2 = I$. For example, for electromagnetic waves $\Gamma_0 = \operatorname{Diag}(1, 1, 1, -1, -1, -1)$. Then the spectrum of L is symmetric about zero and the eigenvalues $\pm \omega^\alpha$ have the same multiplicity. We assume in addition that L is isotropic so that its eigenvalues have the form $\omega_\pm^\alpha(x, k) = \pm c^\alpha(x) |k|$, where $c_\alpha(x)$ is the speed of mode α . We denote by r_α their respective multiplicities, assumed to be independent of x and k for $k \neq 0$. The matrix L has a basis of eigenvectors $\mathbf{b}_\pm^{\alpha, j}(x, k)$ such that

$$L(x, k) \mathbf{b}_\pm^{\alpha, j}(x, k) = \pm \omega^\alpha(x, k) \mathbf{b}_\pm^{\alpha, j}(x, k), \quad j = 1, \dots, r_\alpha,$$

and $b_\pm^{\alpha, j}$ form an orthonormal set with respect to the inner product $\langle \cdot, \cdot \rangle_{A_0}$. The different ω_α correspond to different types of waves (modes). Various indices $1 \leq j \leq r_\alpha$ refer to different polarizations of a given mode. The eigenvectors $b_+^{\alpha, j}$ and $b_-^{\alpha, j}$ are related by

$$\Gamma_0 \mathbf{b}_+^{\alpha, j}(x, k) = \mathbf{b}_-^{\alpha, j}(x, k), \quad \Gamma_0 \mathbf{b}_-^{\alpha, j}(x, k) = \mathbf{b}_+^{\alpha, j}(x, k). \quad (3.66)$$

Proposition 3.3.3 is then generalized as follows [31, 53]:

Proposition 3.4.1 *There exist scalar functions $a_{\pm}^{\alpha,jm}(t, x, k)$ such that*

$$W(t, x, k) = \sum_{\pm, \alpha, j, m} a_{\pm}^{\alpha,jm}(t, x, k) \mathbf{b}_{\pm}^{\alpha,j}(x, k) \otimes \mathbf{b}_{\pm}^{\alpha,m}(x, k). \quad (3.67)$$

Here the sum runs over all possible values of \pm , α , and $1 \leq j, m \leq r_{\alpha}$.

The main content of this proposition is again that the cross terms $\mathbf{b}_{\pm}^{\alpha,j}(x, k) \otimes \mathbf{b}_{\mp}^{\beta,m}(x, k)$ do not contribute, as well as the terms $\mathbf{b}_{\pm}^{\alpha,j}(x, k) \otimes \mathbf{b}_{\pm}^{\alpha',m}(x, k)$ when $\alpha \neq \alpha'$. This is because modes propagating with different speeds do not interfere constructively in the high frequency limit.

We may now insert expression (3.67) into (3.60) and obtain the following generalization of (3.61)

$$\begin{aligned} \hat{\mathbf{u}}^B(k; x_0) = \sum_{\alpha, j, m} & \left[a_{-}^{\alpha,mj}(T, x_0, k) \hat{S}_{+}^{\alpha,j}(x_0, k) \mathbf{b}_{+}^{\alpha,m}(x_0, k) \right. \\ & \left. + a_{+}^{\alpha,mj}(T, x_0, k) \hat{S}_{-}^{\alpha,j}(x_0, k) \mathbf{b}_{-}^{\alpha,m}(x_0, k) \right], \end{aligned} \quad (3.68)$$

where $\hat{S}_{\pm}^{\alpha,j}(k) = (A(x_0) \hat{\mathbf{S}}(k) \cdot \mathbf{b}_{\pm}^{\alpha,j}(x_0, k))$. This formula tells us that only the modes that are present in the initial source ($\hat{S}_{\pm}^{\alpha,j}(k) \neq 0$) will be present in the back-propagated signal but possibly with a different polarization, that is, $j \neq m$.

The initial conditions for the modes $a_{\pm}^{\alpha,jm}$ are given by

$$a_{\pm}^{\alpha,jm}(0, x, k) = |\chi(x)|^2 \hat{f}(k) (A(x) \Gamma(x) \mathbf{b}_{\mp}^{\alpha,m}(x, k) \cdot \mathbf{b}_{\pm}^{\alpha,j}(x, k)), \quad (3.69)$$

which generalizes (3.62). When $\Gamma(x) \equiv I$, we again obtain that $a_{\pm}^{\alpha,jm}(0, x, k) \equiv 0$, i.e., there is no refocusing as physically expected. When $\Gamma(x) \equiv \Gamma_0$, we have for all α that

$$a_{\pm}^{\alpha,jm}(0, x, k) = |\chi(x)|^2 \hat{f}(k) \delta_{jm}.$$

In a uniform medium the amplitudes $a_{\pm}^{\alpha,jm}$ satisfy an uncoupled system of free transport equations (3.43):

$$\frac{\partial a_{\pm}^{\alpha,jm}}{\partial t} \pm c_{\alpha} \hat{k} \cdot \nabla_x a_{\pm}^{\alpha,jm} = 0, \quad (3.70)$$

which have no smoothing effect, and hence refocusing in a homogeneous medium is still poor. When $f(x) = \delta(x)$ and $\Omega = \mathbb{R}^3$, so that $\chi(x) \equiv 1$, we still have that $a_{\pm}^{\alpha,jm}(T, x_0, k) = \delta_{jm}$ and refocusing is again perfect, that is, $\mathbf{u}^B(\xi; x_0) = \mathbf{S}(\xi)$, as may be seen from (3.68).

3.4.4 The diffusive regime

The radiative transport regime holds when the matrices $A(x)$ have the form

$$A(x) = A_0(x) + \sqrt{\varepsilon} A_1\left(\frac{x}{\varepsilon}\right),$$

as in (3.45). Then the $r_{\alpha} \times r_{\alpha}$ coherence matrices w_{\pm}^{α} with entries $w_{\pm, jm}^{\alpha} = a_{\pm}^{\alpha,jm}$ satisfy a system of matrix-valued radiative transport equations (see [53] for the details) similar to (3.46). The matrix transport equations simplify considerably in the diffusive regime, such as the one considered in Section 3.3.8 when waves propagate over large distances and long times.

We assume for simplicity that $A_0 = A_0(x)$ and $\Gamma = \Gamma(x)$ are independent of x . Polarization is lost in this regime, that is, $a^{\alpha,jm}(t, x, k) = 0$ for $j \neq m$ and wave energy is equidistributed over all directions. This implies that

$$a_+^{\alpha,jj}(t, x, k) = a_-^{\alpha,jj}(t, x, k) = a_\alpha(t, x, |k|)$$

so that $a^{\alpha,jj}$ is independent of $j = 1, \dots, r_\alpha$ and of the direction $\hat{k} = k/|k|$. Furthermore, because of multiple scattering, a universal equipartition regime takes place so that

$$a_\alpha(t, x_0, |k|) = \phi(t, x_0, c_\alpha |k|), \quad (3.71)$$

where $\phi(t, x, \omega)$ solves a diffusion equation in x like (3.49) (see [53]). The diffusion coefficient $D(\omega)$ may be expressed explicitly in terms of the power spectra of the medium fluctuations [53]. Using (3.69) and (3.71), we obtain when f is isotropic the following initial data for the function ϕ

$$\phi(0, x, \omega) = \frac{1}{4\pi} |\chi(x)|^2 \int_{S^2} \frac{2}{|\alpha|} \sum_{j, \omega_\alpha > 0} \hat{f}\left(\frac{\omega}{c_\alpha}\right) (A_0 \Gamma \mathbf{b}_-^{\alpha,j}(\hat{k}), \mathbf{b}_+^{\alpha,j}(\hat{k})) d\Omega(\hat{k}), \quad (3.72)$$

where $|\alpha|$ is the number of non-vanishing eigenvalues of $L(x, k)$, and $d\Omega(\hat{k})$ is the Lebesgue measure on the unit sphere S^2 .

Let us assume that non-propagating modes are absent in the initial source $\mathbf{S}(x)$, that is, $\hat{S}_0^j(k) = 0$ with the subscript zero referring to modes corresponding to $\omega_0 = 0$. Then (3.68) becomes

$$\hat{\mathbf{u}}(k; x_0) = \sum_{\alpha, j} \phi(T, x_0, c_\alpha |k|) \left[\hat{S}_+^{\alpha,j}(k) \mathbf{b}_+^{\alpha,j}(x_0, k) + \hat{S}_-^{\alpha,j}(k) \mathbf{b}_-^{\alpha,j}(x_0, k) \right]. \quad (3.73)$$

This is an explicit expression for the re-propagated signal in the diffusive regime, where ϕ solves the diffusion equation (3.49) with initial conditions (3.72).

3.5 Conclusions

This chapter presents a theory that quantitatively describes the refocusing phenomena in time reversal acoustics as well as for more general processings of acoustic and other classical waves. We show that the back-propagated signal may be expressed as the convolution (3.1) of the original source \mathbf{S} with a filter F . The quality of the refocusing is therefore determined by the spatial decay of the kernel F . For acoustic waves, the explicit expression (3.37) relates F to the Wigner distribution of certain solutions of the wave equation. The decay of F is related to the smoothness in the phase space of the amplitudes $a_j(t, x, k)$ defined in Proposition 3.3.3. The latter satisfy free transport equations in homogeneous media, which sharpens the gradients of a_j and leads to poor refocusing. In contrast, the amplitudes a_j satisfy the radiative transport equation (3.46) in heterogeneous media, which has a smoothing effect. This leads to a rapid spatial decay of the filter F and a better refocusing. For longer times, a_j satisfies a diffusion equation. This allows for an explicit expression (3.52)-(3.53) of the time reversed signal. The same theory holds for more general waves and more general processing procedures at the recording array, which allows us to describe the refocusing of electromagnetic waves when only one component of the electric field is measured, for instance.

Bibliography

- [1] F. Bailly, J.F. Clouet and J.P. Fouque, Parabolic and white noise approximation for waves in random media, *SIAM Jour. Appl. Math.* **56**, 1996, 1445–1470.
- [2] G.Bal, G. Papanicolaou and L. Ryzhik, Radiative transport limit for the random Schrödinger equation, *Nonlinearity*, **15**, 2002, 513–529.
- [3] G. Bal, On the self-averaging of wave energy in random media, *SIAM Multiscale Model. Simul.* , **2**, 2004, 398–420.
- [4] G. Bal and L. Ryzhik, Time reversal for classical waves in random media, *Comptes rendus de l’Académie des sciences - Série I - Mathématique*, **333**, 2001, 1041–1046.
- [5] G.Bal and L. Ryzhik, Time reversal and refocusing in random media, *SIAM Jour. Appl. Math.*, **63**, 2003, 1475–1498.
- [6] G. Bal, T. Komorowski and L. Ryzhik, Self-averaging of the Wigner transform in random media, *Comm. Math. Phys.*, 2003, **242**, 81–135.
- [7] G. Bal, Kinetics of scalar wave fields in random media, to appear in *Wave Motion*, 2005.
- [8] C. Bardos and M. Fink.
- [9] C. Bardos and M. Fink, Mathematical foundations of the time reversal mirror, *Asymptotic Analysis*, **29**, 2002, 157–182.
- [10] J. Berryman, L. Borcea, G. Papanicolaou, and C. Tsogka. Imaging and time reversal in random media, *J. Acoust. Soc. Am.* , **112**, 2002, 1509–1522.
- [11] P. Billingsley, *Convergence of Probability Measures*, Wiley, 1999.
- [12] G. Blankenship and G. Papanicolaou, Stability and control of stochastic systems with wide-band noise disturbances, *SIAM Jour. Appl. Math.*, **34**, 1978, 437–476.
- [13] P. Blomgren, G. Papanicolaou, and H. Zhao. Super-Resolution in Time-Reversal Acoustics. *to appear in J. Acoust. Soc. Am.*, 2001.
- [14] L. Borcea, G. Papanicolaou, C. Tsogka and J. Berryman, Imaging and time reversal in random media, *Inverse Problems*, **18**, 2002, 1247–1279.
- [15] L. Borcea, G. Papanicolaou and C. Tsogka, Theory and applications of time reversal and interferometric imaging, *Inverse Problems*, **19**, 2003, S139–S164.
- [16] L. Borcea, G. Papanicolaou and C. Tsogka, Interferometric array imaging in clutter, *Inverse Problems*, **21**, 2005, 1419–1460.

- [17] L. Borcea, G. Papanicolaou and C. Tsogka, Coherent interferometry in finely layered random media, *SIAM J. on Multiscale Model. Simul.*, in press, 2006.
- [18] J. F. Clouet and J. P. Fouque. A time-reversal method for an acoustical pulse propagating in randomly layered media. *Wave Motion*, 25:361–368, 1997.
- [19] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Vol.6.* Springer Verlag, Berlin, 1993.
- [20] A. Derode, P. Roux and M. Fink, Robust Acoustic Time-Reversal With High-Order Multiple-Scattering, *Phys. Rev. Lett.*, **75**, 1995, 4206–4209.
- [21] D. R. Dowling, and D.R. Jackson, Narrow-Band performance of phase-conjugate arrays in dynamic random media, *J. Acoust. Soc. Am.*, **91**, 1992, 3257–3277.
- [22] 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.
- [23] L. Erdős and H. T. Yau. Linear Boltzmann equation as the weak coupling limit of a random Schrödinger Equation. *Comm. Pure Appl. Math.*, 53(6):667–735, 2000.
- [24] L. Erdős, M. Salmhofer and H.T. Yau, Quantum diffusion of the random Schrodinger evolution in the scaling limit I. The non-recollision diagrams Preprint, 2005.
- [25] L. Erdős, M. Salmhofer and H.T. Yau, Quantum diffusion of the random Schrodinger evolution in the scaling limit II. The recollision diagrams, Preprint, 2005.
- [26] M. Fink. Time reversed acoustics. *Physics Today*, 50(3):34–40, 1997.
- [27] M. Fink. Chaos and time-reversed acoustics. *Physica Scripta*, 90:268–277, 2001.
- [28] M. Fink and C. Prada. Acoustic time-reversal mirrors. *Inverse Problems*, **17**, 2001, R1–R38.
- [29] J.-P. Fouque, La convergence en loi pour les processus à valeurs dans un espace nucléaire, *Ann. Inst. H. Poincaré Probab. Statist.* **20**, 1984, 225–245.
- [30] J. Garnier, J.P. Fouque, G. Papanicolaou and K. Solna, *Wave Propagation and Time Reversal in Randomly Layered Media*, to appear.
- [31] P. Gérard, P. A. Markowich, N. J. Mauser, and F. Poupaud, Homogenization limits and Wigner transforms, *Comm. Pure Appl. Math.*, **50**, 1997, 323–380.
- [32] I.I. Gikhman and A.V. Skorochod, *Theory of stochastic processes, vol. 3*, Springer Verlag, 1974, Berlin.
- [33] D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, Springer Verlag, 1998, Berlin.
- [34] F. Golse, P.-L. Lions, B. Perthame, and R. Sentis. Regularity of the moments of the solution of a transport equation. *Journal of Functional Analysis* **76**, pages 110–125, 1988.
- [35] W. Hodgkiss, H. Song, W. Kuperman, T. Akal, C. Ferla, and D. Jackson. A long-range and variable focus phase-conjugation experiment in a shallow water. *J. Acoust. Soc. Am.*, 105:1597–1604, 1999.

- [36] A. Ishimaru. *Wave Propagation and Scattering in Random Media*. New York: Academics, 1978.
- [37] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, New York, 1991.
- [38] J.B. Keller, Stochastic equations and wave propagation in random media, Proc. Sympos. Appl. Math. AMS, Vol. XVI, 1964, 145–170.
- [39] H. Kesten and G. Papanicolaou, A limit theorem for turbulent diffusion, Comm. Math. Phys., **65**, 1979, 97-128.
- [40] H. Kesten, G. C. Papanicolaou, A Limit Theorem for Stochastic Acceleration, Comm. Math. Phys., **78**, 1980, 19–63.
- [41] R. Khasminskii, A limit theorem for solutions of differential equations with a random right hand side, Theory Probab. Appl. **11**, 390-406.
- [42] S. R. Khosla and D. R. Dowling. Time-reversing array retrofocusing in noisy environments. *J. Acous. Soc. Am.*, 109(2):538–546, 2001.
- [43] T. Komorowski, Diffusion approximation for the advection of particles in a a strongly turbulent random environment, Ann. Probab., **24**, 1996, 346–376.
- [44] T. Komorowski and L. Ryzhik, Diffusion in a weakly random Hamiltonian flow, Comm. Math. Phys., **262**, 2006, 277–323.
- [45] T. Komorowski and L. Ryzhik, The stochastic acceleration problem in two dimensions, to appear in Israel Journal of Mathematics, 2006.
- [46] W. Kuperman, W. Hodgkiss, H. Song, T. Akal, C. Ferla, and D. Jackson. Phase-conjugation in the ocean. *J. Acoust. Soc. Am.*, 102:1–16, 1997.
- [47] S. Kusuoka and D. Stroock, Applications of the Malliavin calculus, Part II, J. Fac. Sci. Univ. Tokyo, Sect. IA, Math, **32**, 1985, 1–76.
- [48] E. W. Larsen and J. B. Keller. Asymptotic solution of neutron transport problems for small mean free paths. *J. Math. Phys.*, 15:75–81, 1974.
- [49] P.-L. Lions and T. Paul. Sur les mesures de Wigner. *Rev. Mat. Iberoamericana*, 9:553–618, 1993.
- [50] I. Mitoma, On the sample continuity of \mathcal{S}' -processes. J. Math. Soc. Japan, **35**, 1983, 629–636.
- [51] M. Mokhtar-Kharroubi. *Mathematical Topics in Neutron Transport Theory*. World Scientific, Singapore, 1997.
- [52] G. Papanicolaou and W. Kohler, Asymptotic theory of mixing stochastic ordinary differential equations, Comm. Pure Appl. Math., **27**, 1974, 641-668.
- [53] L. Ryzhik, G. Papanicolaou, and J. B. Keller, Transport equations for elastic and other waves in random media, Wave Motion, **24**, 1996, 327–370.
- [54] G. Papanicolaou, L. Ryzhik, and K. Solna, Statistical stability in time reversal, SIAM Jour. Appl. Math., **64**, 2004, 1133–1155.

- [55] P. Sheng. *Introduction to Wave Scattering, Localization and Mesoscopic Phenomena*. Academic Press, New York, 1995.
- [56] H. Spohn. Derivation of the transport equation for electrons moving through random impurities. *Jour. Stat. Phys.*, 17:385–412, 1977.
- [57] D. Strook, *An Introduction to the Analysis of Paths on a Riemannian Manifold*, Math. Surv. and Monographs v. **74**, 2000.
- [58] D. Strook, S. R. S. Varadhan *Multidimensional Diffusion Processes*, Berlin, Heidelberg, New York: Springer-Verlag, 1979.
- [59] F. Tappert, The parabolic approximation method, Lecture notes in physics, vol. 70, *Wave propagation and underwater acoustics*, Springer-Verlag, 1977
- [60] M. Taylor, *Partial differential equations, vol. 2*, New York, Springer-Verlag, 1996.
- [61] B.S. White, The stochastic caustic, *SIAM J. Appl. Math.* **44**, 1984, 127–149.