Waves in weakly random media: lecture notes for the Vienna Inverse Problems school

Lenya Ryzhik*

May 6, 2014

1 Introduction

Wave propagation in complex media is an ubiquitous phenomenon – applications include light propagation through the atmosphere, underwater acoustic, tomography, and innumerable other ones. These problems may have very different specific details but possess one unifying feature: the precise knowledge of the medium between the wave source and the receiver is not and can not be known. This makes the numerical computation of the solutions of the exact wave equations (whether acoustic, electromagnetic or elastic) not only beyond the reach of even the fastest modern computers but also pointless – as we do not know the details of the medium, there is nothing to plug into the computer as the coefficients to solve the wave equations. Fortunately, the microscopic details of the medium often do not matter for quantities of interest. An obvious situation when that is true is if the medium is essentially uniform, so that the fluctuations have a very small effect on the wave evolution. However, even very small fluctuations will have a non-trivial effect after a sufficiently long time, and propagation over long distances, a regime often encountered in practice. A surprising phenomenon is that while such small fluctuations will eventually have a large effect, the macroscopic features of the wave will nevertheless not depend on the fine details on the microstructure. That is, the wave will be very far from what it would be in a uniform medium (both on the microscopic and the macroscopic levels) but its macroscopic features can be captured by models that do not need the knowledge of the microstructure. An introduction to such models and some of the ways to obtain them are the subject of these notes.

As we can not know the details of the medium, it is convenient to model the media parameters (sound speed, elastic parameters, the dielectric constant and so on) as random fields. The main interest will be then in finding the features of the solutions of the wave equations with random coefficients that would not depend on the particular realization of the random medium but rather on its statistics which may be encoded in a few parameters. This is particularly important in the inverse problems – we can not afford to use unstable data (in the statistical sense) for (usually) ill-posed inverse problems, so it is imperative to be sure that the data used for inverse problems is as stable (non-random) as possible.

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

The main regimes of wave propagation in heterogenous media

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

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

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

$$\frac{\partial W(t,x,\xi)}{\partial t} + \nabla_\xi \omega(\xi) \cdot \nabla_x W(t,x,\xi) = \mathcal{L}_{sc} W(t,x,\xi). \tag{1.1}$$

Here, $\omega(\xi)$ is the dispersion relation of the wave and depends on the particular type of the wave. The left side of (1.1) has nothing to do with the inhomogeneities of the medium[1] and represents the free transport of the wave energy along the characteristics $\dot{X} = \nabla_\xi \omega(\xi)$ (which are straight lines). On the other hand, the scattering operator $\mathcal{L}_{sc}$ incorporates the macroscopic effects of the small scale inhomogeneities, and involves the possibility for waves to scatter in different directions at a given point. Its exact form depends on the physical regime of the problem, and the task of modeling is typically two-fold: to find the relation of the phase space energy density $W(t, x, \xi)$ to the underlying wave field that can be directly

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

measured (pressure, electric and magnetic fields, elastic displacements, and so on, depending on the problem), and to identify the scattering operator $\mathcal{L}_{sc}$ for a particular physical problem. Next, we describe some of the possible macroscopic models.

Random geometric optics. The geometric optics regime arises when the wave length of the signal is much smaller than the typical scale of variations in the medium, whether the latter are random or not. Then the wave propagation is described in terms of the rays (this description goes back to Fermat and Huygens') that are straight lines in a uniform medium but are curved if the sound speed is varying. In our terminology, this corresponds to the relative sizes

$$\lambda \ll l_c \ll L,$$

and the problem has three scales: on the microscopic level (scale λ) one considers the precise wave evolution, on the intermediate scale l_c the problem is described in terms of rays in a random medium, and, finally, the macroscopic description (on the scale L we will need to find in terms of λ, l_c and ε) will be in terms of the Fokker-Planck equation:

$$\frac{\partial W(t,x,\xi)}{\partial t} + \nabla\omega(\xi)\cdot\nabla_x W(t,x,\xi) = \sum_{i,j=1}^{n} \frac{\partial}{\partial \xi_i}\big(D_{ij}(\xi)\frac{\partial W(t,x,\xi)}{\partial \xi_j}\big). \tag{1.2}$$

That is, the scattering operator in (1.1) in this regime is a momentum diffusion:

$$\mathcal{L}_{sc} f(\xi) = \sum_{i,j=1}^{n} \frac{\partial}{\partial \xi_i}\big(D_{ij}(\xi)\frac{\partial f(\xi)}{\partial \xi_j}\big). \tag{1.3}$$

This means that multiple scattering makes the ray direction diffuse over long distances – this is the macroscopic effect of the small scale heterogeneities, and it is encoded in the deterministic effective diffusion matrix $D_{ij}(\xi)$. The role of the fluctuations is to create a non-trivial diffusion, in the absence of fluctuations, we have $D = 0$, of course.

Radiative transport regime. The radiative transport regime arises when the correlation length of the medium is comparable to the wave length of the probing signal: $\lambda \approx l_c \ll L$. This is a two-scale problem, the microscopic scale is $\lambda = l_c$, and the macroscopic scale is L (that will, once again, be determined by $\lambda = l_c$ and ε), and on the microscopic level the interactions between the inhomogeneities and the wave are of a different nature than in the geometric optics regime, leading to a different macroscopic limit. The effective kinetic equation is of the radiative transfer type:

$$\frac{\partial W(t,x,\xi)}{\partial t} + \nabla\omega(\xi)\cdot\nabla_x W(t,x,\xi) = \int \sigma(\xi,p)(W(t,x,p) - W(t,x,\xi))dp. \tag{1.4}$$

The scattering operator is now of the form

$$\mathcal{L}_{sc} f(\xi) = \int \sigma(\xi,p)(f(p) - f(\xi))dp. \tag{1.5}$$

The (deterministic) differential scattering cross-section $\sigma(\xi,p)$ encodes the macroscopic effect of the small scale inhomogeneities, as the diffusion coefficient $D_{ij}(\xi)$ did it the random geometric optics regime.

The homogenization regime. The homogenization regime corresponds to probing signals with $\lambda \gg l_c$, so that from the point of view the wave, the inhomogeneities are small scale. In that case, the phase of the wave is affected in a non-trivial case before the wave amplitude, and the kinetic equation description does not capture this phase modulation.
The spatial diffusion regime. We would be remiss not to mention that a typical situation in a weakly random medium is that, no matter what exactly the scattering operator $\mathcal{L}_{sc}$ in the kinetic equation (1.1) is, the multiple scattering will lead to equilibration of energy in all directions: $W(t,x,\xi) = W(t,x,|\xi|)$ is uniformly distributed in momenta after "very long" times, and the energy density satisfies the spatial diffusion equation:

$$\frac{\partial W(t,x,|\xi|)}{\partial t} = D(|\xi|\Delta_x W(t,x,|\xi)). \tag{1.6}$$

In this ultimate regime, the only input of the random medium is in the diffusion coefficient $D(|\xi|)$. This model, is extremely simple, and by virtue of its simplicity, is very popular in practice.

When do things happen in a weakly random medium?

We finish this introduction with an illustration of when one can expect a weakly random medium to have a non-trivial effect. Probably, the simplest such situation is evolution of a particle in a random time-dependent velocity field:

$$\frac{dX(t)}{dt} = \varepsilon V(t), \quad X(0) = 0, \tag{1.7}$$

that is,

$$X(t) = \varepsilon \int_0^t V(s)ds. \tag{1.8}$$

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

$$R_{ij}(t) = \langle V_i(0)V_j(t)\rangle,$$

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

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

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

Going back to the particle trajectory (1.8), we see that its average position is

$$\bar{X}(t) = \langle X(t) \rangle = \varepsilon \bar{V} t,$$

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

$$\begin{aligned}
\langle X_i(t) X_j(t) \rangle &= \varepsilon^2 \int_0^t ds_1 \int_0^t ds_2 \mathbb{E}(V_i(s_1) V_j(s_2)) = \varepsilon^2 \int_0^t ds_1 \int_0^t ds_2 R_{ij}(s_1 - s_2) \\
&= \varepsilon^2 \int_0^t ds_1 \int_0^{s_1} ds_2 R_{ij}(s_1 - s_2) + \varepsilon^2 \int_0^t ds_1 \int_{s_1}^t ds_2 R_{ij}(s_1 - s_2) \\
&= \varepsilon^2 \int_0^t ds_1 \int_0^{s_1} ds_2 R_{ij}(s_2) + \varepsilon^2 \int_0^t ds_1 \int_0^{t-s_1} ds_2 R_{ij}(-s_2) \qquad (1.9) \\
&= \varepsilon^2 \int_0^t (t - s_2)[R_{ij}(s_2) + R_{ij}(-s_2)] ds_2 = \varepsilon^2 [D_{ij} t + O(1)], \quad \text{as } t \to +\infty.
\end{aligned}$$

with the diffusivity matrix

$$D_{ij} = \int_{-\infty}^{\infty} R_{ij}(s) ds = \hat{R}_{ij}(0). \qquad (1.10)$$

Expression (1.9) tells us (at least) two things: first, we should expect a non-trivial behavior for the particle at times of the order $t \sim \varepsilon^{-2}$, and, second, that the particle behavior at this time scale should be a Brownian motion $B_D(t)$ with the correlation matrix D_{ij}. Strictly speaking, we have only computed that its variance agrees with that of $B_D(t)$ but it is not difficult to make this rigorous. That is, we have the following result: if $X(t)$ solves (1.7) with a mean-zero statistically time homogeneous random field $V(t)$ then the process $X_\varepsilon(t) = X(t/\varepsilon^2)$ converges, as $t \to +\infty$, to a Brownian motion with the covariance matrix D_{ij}. The main observation here is that "mean-zero randomness of size ε has a non-trivial effect on the time scales of the order ε^{-2}" – something that any probabilist knows very well from the classical central limit theorem, going at least as far back as de Moivre and 1733.

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

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

Organization of the notes

The goal of the present notes is to present some of the mathematical results on the aforementioned kinetic models. Ideally, one would like to do that for the true wave equation

$$\frac{1}{c^2(x)}\phi_{tt} - \Delta\phi = 0, \tag{1.11}$$

with a weakly random velocity profile $c(x)$, and occasionally we will be able to do this. However, we should mention two models that are much simpler mathematically but rich enough to appreciate the difficulties and the diversity of the possible regimes. The first is simply a first-order advection equation

$$\phi_t + v(t,x)\cdot\nabla\phi(x) = 0, \tag{1.12}$$

with a weakly random velocity $v(t,x)$. Its advantage is that the method of characteristics allows us to obtain various results about the solutions of the PDE (1.12) using the particle methods of the probability theory. The simple advection equation captures some (but by no means all) of the common features of the solutions of the first order hyperbolic systems (such as the acoustic, electromagnetic and elastic wave equations) surprisingly well. The second, on which we will mostly focus, is the Schrödinger equation

$$i\phi_t + \frac{1}{2}\Delta\phi - \varepsilon V(t,x)\phi = 0, \tag{1.13}$$

with a weakly random potential $V(t,x)$. This equation appears not only in the quantum mechanics but also as the paraxial approximation for the propagation of a time-harmonic narrow beam - then, the "time" t is the coordinate in the direction of the beam, and the "spatial" variables x correspond to the true spatial variables in the directions perpendicular to the beam.

2 The geometric optics via the Wigner transform

In this section we introduce a useful tool, the Wigner transform, for the passage from the oscillatory solutions of a linear hyperbolic or dispersive non-dissipative PDE to the characteristic in the phase space. In order to keep the presentation manageable we will focus solely on the solutions of the Schrödinger equation

$$i\phi_t + \frac{1}{2}\Delta\phi - V(t,x)\phi = 0, \tag{2.1}$$

with a real potential $V(t,x)$. The generalization of the methods and results we describe below to the wave equations is usually (but not always) reasonably straightforward though it typically involves rather lengthy calculations that we will try to avoid here, to the extent possible.

2.1 The Wigner transform and its properties

The unscaled Wigner transform

The Schrödinger equation (2.1) preserves the total energy of the solution (or the total number of particles depending on the point of view or physical application):

$$\mathcal{E}(t) = \int |\phi(t,x)|^2 dx = \mathcal{E}(0),$$

as may be verified by a straightforward time differentiation. However, often one is interested not only in the conservation of the total energy $\mathcal{E}(t)$ but also in its local spatial distribution – that is, where the energy is concentrated. This requires understanding the local energy density $E(t,x) = |\phi(t,x)|^2$. Note that even if $\phi(t,x)$ is oscillatory the function $E(t,x)$ may vary slowly in space – this happens if the phase of $\phi(t,x)$ oscillates much faster than its amplitude, as in the geometric optics regime. Unfortunately, while all the information about the "relatively simple" function $E(t,x)$ may be extracted from a "complicated" function $\phi(t,x)$, the energy density $E(t,x)$ itself does not satisfy a closed equation. Rather, its evolution is described by a conservation law

$$\frac{\partial E}{\partial t} + \nabla \cdot F = 0,$$

with the flux

$$F(t,x) = \frac{1}{2i}\left(\bar{\phi}\nabla\phi - \phi\nabla\bar{\phi}\right).$$

A remedy for this lack of equation for $E(t,x)$ when the potential $V = 0$ was proposed by Wigner in his 1932 paper [12] (where he credits Leo Szilard for this discovery). Wigner introduced the following object:

$$W(t,x,k) = \int \phi\left(t, x - \frac{y}{2}\right)\bar{\phi}\left(t, x + \frac{y}{2}\right) e^{ik\cdot y}\frac{dy}{(2\pi)^n}. \tag{2.2}$$

It is immediate to check that

$$\int W(t,x,k)dk = |\phi(t,x)|^2 = E(t,x), \tag{2.3}$$

so that in some sense $W(t,x,k)$ is "a local energy density resolved over momenta". In addition, the "average momentum" is

$$\begin{aligned}\int kW(t,x,k)dk &= \frac{1}{i}\int ik\phi\left(t, x - \frac{y}{2}\right)\bar{\phi}\left(t, x + \frac{y}{2}\right) e^{ik\cdot y}\frac{dydk}{(2\pi)^n} \\ &= -\frac{1}{i}\int \nabla_y\left[\phi\left(t, x - \frac{y}{2}\right)\bar{\phi}\left(t, x + \frac{y}{2}\right)\right] e^{ik\cdot y}\frac{dydk}{(2\pi)^n} \\ &= \frac{1}{2i}\left[\bar{\phi}(t,x)\nabla\phi(t,x) - \phi(t,x)\nabla\bar{\phi}(t,x)\right].\end{aligned}$$

Therefore, the flux can be expressed in terms of the Wigner transform as

$$F(t,x) = \int kW(t,x,k)dk,$$

re-enforcing the interpretation of $W(t,x,k)$ as a phase space energy density. It is also immediate to observe that $W(t,x,k)$ is real-valued.

A remarkable observation is that if $V=0$, the function $W(t,x,k)$ satisfies an evolution equation:

$$W_t + k\cdot\nabla_x W = 0. \tag{2.4}$$

Therefore, one may describe the energy density evolution for the Schrödinger equation with zero potential as follows: compute the initial data $W(0,x,k)$, solve the kinetic equation (2.4) and find $|\phi(t,x)|^2$ using relation (2.3). However, there is one drawback in the interpretation of $W(t,x,k)$ as the energy density resolved over positions and momenta – there is no reason for $W(t,x,k)$ to be non-negative!

The Schrödinger equation (2.1) with a potential $V\not\equiv 0$ leads to the following evolution equation for $W(t,x,k)$:

$$\frac{\partial W}{\partial t} + k\cdot\nabla_x W = \frac{1}{i}\int e^{ip\cdot x}\hat{V}(p)\left[W\left(k-\frac{p}{2}\right) - W\left(k-\frac{p}{2}\right)\right]\frac{dp}{(2\pi)^n}. \tag{2.5}$$

While the uniform kinetic equation (2.4) posseses some nice properties – in particular, it preserves positivity of the initial data and has a particle interpretation: it describes density evolution of particles moving along the trajectories $\dot{X}=K$, $\dot{K}=0$, the Wigner equation (2.5) has very few attractions. In particular, it does not preserve positivity of the initial data. Probably, for that reason the Wigner transform ideas did not evolve mathematically (at least they did not spread widely in the mathematics community though they were used by physicists and engineers) until the work of P. Gérard and L. Tartar in the late eighties. They have realized that the Wigner transforms become a useful tool in the analysis of the semiclassical asymptotics, that is, in the study of the oscillatory solutions of the Schrödinger equation (as well as in other oscillatory problems).

The semiclassical Wigner transform

The definition of the Wigner transform for oscillatory functions has to be modified: to see this, consider a simple oscillating plane wave $\phi_\varepsilon(x) = e^{ik_0\cdot x/\varepsilon}$ with a fixed $k_0\in\mathbb{R}^n$. Then its Wigner transform as defined by (2.2) is

$$W(x,k) = \int e^{ik\cdot y}e^{ik_0\cdot(x-y/2)/\varepsilon - ik_0\cdot(x+y/2)/\varepsilon}\frac{dy}{(2\pi)^n} = \delta\left(k-\frac{k_0}{\varepsilon}\right).$$

We see that $W(x,k)$ does not have a nice limit as $\varepsilon\to 0$. On the other hand, its rescaled version $W_\varepsilon(x,k) = \varepsilon^{-d}W(x,k/\varepsilon)$ does converge (actually, equals to) to $\delta(k-k_0)$. This motivates the following defininition of the (rescaled) Wigner transform of a family of functions $\phi_\varepsilon(x)$:

$$W_\varepsilon(x,k) = \frac{1}{\varepsilon^d}\int \phi_\varepsilon\left(x-\frac{y}{2}\right)\bar{\phi}_\varepsilon\left(x+\frac{y}{2}\right)e^{ik\cdot y/\varepsilon}\frac{dy}{(2\pi)^n},$$

that may be more conveniently re-written as

Definition 2.1 *The Wigner transform (or the Wigner distribution) of a family of functions $\phi_\varepsilon(x)$ is a distribution $W_\varepsilon(x,k)\in\mathcal{S}'(\mathbb{R}^n\times\mathbb{R}^n)$ defined by*

$$W_\varepsilon(t,x,k) = \int \phi_\varepsilon\left(x-\frac{\varepsilon y}{2}\right)\bar{\phi}_\varepsilon\left(x+\frac{\varepsilon y}{2}\right)e^{ik\cdot y}\frac{dy}{(2\pi)^n}. \tag{2.6}$$

Expression (2.6) shows that $W_\varepsilon(x,k)$ is well suited to study functions oscillalting on the scale $\varepsilon \ll 1$ – in that case the difference of the arguments εy is chosen so that the function ϕ_ε changes by $O(1)$.

The Wigner transform is mostly used for families of solutions of non-dissipative evolution equations that conserve the L^2-norm (or a weighted L^2-norm), simply because the scaling in (2.6) is particularly well suited for families of functions $\phi_\varepsilon(x)$ that are uniformly (in the parameter $\varepsilon \in (0,1)$) bounded in $L^2(\mathbb{R}^n)$. Let us define the space of test functions

$$\mathcal{A} = \left\{ \lambda(x,k) \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n) : \int \sup_x \left[\left| \tilde{\lambda}(x,y) \right| \right] dy < +\infty \right\}$$

with the norm

$$\|\lambda\|_{\mathcal{A}} = \int \sup_x \left[\left| \tilde{\lambda}(x,y) \right| \right] dy.$$

We have the following proposition.

Proposition 2.2 *Let the family of functions $\phi_\varepsilon(x)$ be uniformly bounded in $L^2(\mathbb{R}^n)$. Then the corresponding family of Wigner transforms $W_\varepsilon(x,k)$ is uniformly bounded in $\mathcal{A}'(\mathbb{R}^n \times \mathbb{R}^n)$.*

The following is an immediate corollary of the above proposition and the Banach-Alaoglu theorem.

Corollary 2.3 *Let the family of functions $\psi_\varepsilon(x)$ be uniformly bounded in $L^2(\mathbb{R}^n)$. Then the corresponding family of Wigner transforms $W_\varepsilon(x,k)$ has a weak-$\star$ converging subsequence in the space $\mathcal{A}'(\mathbb{R}^n \times \mathbb{R}^n)$.*

The limit is a non-negative measure of a bounded total mass.

Proposition 2.4 *Let $\phi_\varepsilon(x)$ be a uniformly bounded family of functions in $L^2(\mathbb{R}^n)$, and let $W(x,k) \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$ be a limit point of the corresponding family $W_\varepsilon(x,k)$. Then we have $W(x,k) \geq 0$ and the total mass*

$$\int_{\mathbb{R}^{2n}} W(dxdk) < +\infty.$$

We summarize Corollary 2.3 and Proposition 2.4 into the following theorem.

Theorem 2.5 *Let the family ϕ_ε be uniformly bounded in $L^2(\mathbb{R}^n)$. Then the Wigner transform W_ε converges weakly along a subsequence $\varepsilon_k \to 0$ to a distribution $W(x,k) \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$. Any such limit point $W(x,k)$ is a non-negative measure of bounded total mass.*

Can the weak convergence of the Wigner transforms become strong? This is possible in principle – for instance, the Wigner transforms of $\psi_\varepsilon(x) = e^{ik_0 \cdot x/\varepsilon}$ is independent of ε – $W_\varepsilon(x,k) = \delta(k-k_0)$. However, this is impossible in $L^2(\mathbb{R}^n \times \mathbb{R}^n)$ as the L^2-norm of W_ε is unbounded unless $\phi_\varepsilon(x)$ converges strongly to zero:

$$\begin{aligned}
&\int |W_\varepsilon(x,k)|^2 dxdk \\
&= \int e^{ik\cdot y - ik\cdot y'} \phi_\varepsilon\left(x - \frac{\varepsilon y}{2}\right) \bar{\phi}_\varepsilon\left(x + \frac{\varepsilon y}{2}\right) \bar{\phi}_\varepsilon\left(x - \frac{\varepsilon y'}{2}\right) \phi_\varepsilon\left(x + \frac{\varepsilon y'}{2}\right) \frac{dydy'dxdk}{(2\pi)^{2n}} \\
&= \int \left| \phi_\varepsilon\left(x - \frac{\varepsilon y}{2}\right) \phi_\varepsilon\left(x + \frac{\varepsilon y}{2}\right) \right|^2 \frac{dydx}{(2\pi)^n} = \frac{1}{(2\pi\varepsilon)^n} \|\phi_\varepsilon\|^4_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

Therefore, it is impossible to expect even weak convergence of W_ε in $L^2(\mathbb{R}^n \times \mathbb{R}^n)$ unless the family ϕ_ε converges strongly to zero. In that case, however, $W_\varepsilon = 0$, which is not a very interesting case.

Examples of the Wigner measures

We now give some examples of the Wigner measures, leaving the computations leading to the limit measures as an exercise to the reader.

A strongly converging sequence. Let $\phi_\varepsilon(x)$ converge strongly in $L^2(\mathbb{R}^n)$ to a limit $\phi(x)$. Then the limit Wigner measure is $W(x,k) = |\phi(x)|^2\delta(k)$. This means that for non-oscillatory families the limit Wigner measure is supported at $k = 0$.

The localized case. The Wigner transform of the family $f_\varepsilon(x) = \varepsilon^{-n/2}\phi(x/\varepsilon)$ with a compactly supported function $\phi(x)$ is $W(x,k) = (2\pi)^{-n}|\hat{\phi}(k)|^2\delta(x)$.

The WKB case. The Wigner measure of the family $\phi_\varepsilon(x) = A(x)\exp\{iS(x)/\varepsilon\}$ with a smooth amplitude $A(x)$ and phase function $S(x)$, is $W(x,k) = |A(x)|^2\delta(k - \nabla S(x))$ since

$$\begin{aligned} W^\varepsilon(x,k) &= \int e^{ik\cdot y} e^{iS(x-\frac{\varepsilon y}{2})/\varepsilon} A(x - \frac{\varepsilon y}{2}) e^{-iS(x+\frac{\varepsilon y}{2})/\varepsilon} \bar{A}(x + \frac{\varepsilon y}{2}) \frac{dy}{(2\pi)^n} \\ &= \int e^{ik\cdot y} e^{-i\nabla S(x)\cdot y} |A(x)|^2 \frac{dy}{(2\pi)^n} + O(\varepsilon) = |A(x)|^2\delta(k - \nabla S) + O(\varepsilon). \end{aligned}$$

Coherent states. The WKB and concentrated cases can be combined – this is a coherent state

$$\phi_\varepsilon(x) = \frac{1}{\varepsilon^{n/2}}\phi\left(\frac{x - x_0}{\varepsilon}\right) e^{ik_0\cdot x}.$$

The Wigner measure of this family is

$$W(x,k) = \frac{1}{(2\pi)^n}\delta(x - x_0)|\hat{\phi}(k - k_0)|^2.$$

Scale mismatch. The Wigner transform captures oscillations on a scale ε but not on a different scale. To see this, consider a WKB family $\phi_\varepsilon(x) = A(x)e^{ik_0\cdot x/\varepsilon^\alpha}$ – we have treated the case $\alpha = 1$ but now we look at $0 \le \alpha < 1$ or $\alpha > 1$. First, if $\alpha \in (0,1)$ then we have

$$\begin{aligned} W^\varepsilon(x,k) &= \int e^{ik\cdot y} e^{ik_0\cdot(x-\frac{\varepsilon y}{2})/\varepsilon^\alpha} A(x - \frac{\varepsilon y}{2}) e^{-ik_0\cdot(x+\frac{\varepsilon y}{2})/\varepsilon^\alpha} \bar{A}(x + \frac{\varepsilon y}{2}) \frac{dy}{(2\pi)^n} \\ &= \int e^{i(k-\varepsilon^{1-\alpha}k_0)\cdot y} |A(x)|^2 \frac{dy}{(2\pi)^n} + O(\varepsilon) = |A(x)|^2\delta(k) + o(1). \end{aligned}$$

Therefore, if $0 \le \alpha < 1$ then W_ε has the limit $W(x,k) = |A(x)|^2\delta(k)$ as in the case $\alpha = 0$ – the limit does not capture the oscillations at all. On the other hand, if $\alpha > 1$ then

$$\begin{aligned} \langle a, W_\varepsilon\rangle &= \int e^{ik\cdot y} e^{ik_0\cdot(x-\frac{\varepsilon y}{2})/\varepsilon^\alpha} a(x,k) A(x - \frac{\varepsilon y}{2}) e^{-ik_0\cdot(x+\frac{\varepsilon y}{2})/\varepsilon^\alpha} \bar{A}(x + \frac{\varepsilon y}{2}) \frac{dydxdk}{(2\pi)^n} \\ &= \int e^{-ik_0\cdot y/\varepsilon^{1-\alpha}} \tilde{a}(x,y) A(x - \frac{\varepsilon y}{2}) \bar{A}(x + \frac{\varepsilon y}{2}) \frac{dxdy}{(2\pi)^n} \to 0 \end{aligned}$$

as $\varepsilon \to 0$. We see that when the family oscillates on a scale much smaller than ε the limit Wigner measure computed with respect to a "too large" scale ε vanishes and does not capture the oscillations correctly. This is a mixed blessing of the Wigner measures – they are very useful but only as long they are computed with respect to a correct scale. We will make this statement precise in the next section.

Basic properties of the Wigner measures

It turns out that though the definition of the Wigner transform $W_\varepsilon(x,k)$ involves integration over the whole space, the limit Wigner measure is a local notion in space (on the macroscopic scale). We say that a family of functions $\phi_\varepsilon(x)$ is pure if the Wigner transforms W_ε converge as $\varepsilon \to 0$ to the limit $W(x,k)$ – that is, we do not need to pass to a subsequence $\varepsilon_k \to 0$ and the limit is unique.

Proposition 2.6 (Localization) *Let $\phi_\varepsilon(x)$ be a pure family of uniformly bounded functions in L^2 and let $\mu(x,k)$ be the unique limit Wigner measure of this family. Let $\theta(x)$ be a smooth function. Then the family $\psi_\varepsilon(x) = \theta(x)\phi_\varepsilon(x)$ is also pure, and the Wigner transforms $W_\varepsilon[\psi_\varepsilon]$ of the family $\psi_\varepsilon(x)$ converge to $|\theta(x)|^2\mu(x,k)$ as $\varepsilon \to 0$. Moreover, let ϕ_ε be a uniformly bounded pure family of L^2 functions, and let ψ_ε coincide with ϕ_ε in an open neighbourhood of a point x_0. Then the the limit Wigner measures $\mu[\phi]$ and $\mu[\psi]$ coincide in this neighborhood.*

Another useful and intuitively clear property is that the Wigner measure of waves going in different directions is the sum of the individual Wigner measures.

Lemma 2.7 (Orthogonality) *Let ϕ_ε, ψ_ε be two pure families of functions with Wigner measures μ and ν, respectively,which are mutually singular. Then the Wigner measure of the sum $\phi_\varepsilon + \psi_\varepsilon$ is $\mu + \nu$.*

The above properties: positivity, orthogonality and localization show that the Wigner measure may be indeed reasonably interpreted as the phase space energy density. However, the following pair of examples shows that the limit may not capture the energy correctly. The first "bad" example is the family

$$\phi_\varepsilon(x) = A(x)e^{ik\cdot x/\varepsilon^2}.$$

Then the limit Wigner transform is $W = 0$ while the spatial energy density

$$E_\varepsilon(x) = |\phi_\varepsilon(x)|^2 \equiv |A(x)|^2$$

does not vanish in the limit $\varepsilon \to 0$. The second "misbehavior" is more classical, and can be seen on standard "escape to infinity" example

$$\phi_\varepsilon(x) = \theta\left(x - \frac{1}{\varepsilon}\right), \quad x \in \mathbb{R}, \tag{2.7}$$

with $\theta(x) \in C_c^\infty(\mathbb{R})$. Then the limit Wigner measure $W(x,k) = 0$ and the local energy density $|\phi_\varepsilon(x)|^2$ converges weakly to zero as well. However, the total mass $\|\phi_\varepsilon\|_{L^2} \equiv \|\theta\|_{L^2}$ is not captured correctly by the limit.

It turns out that the above two examples exhaust the possibilities for the Wigner measure to fail to capture the energy correctly and it is well suited for families of functions that depend on a small parameter in an oscillatory manner, the ε-oscillatory families of [5]. The ε-oscillatory property guarantees that the functions ϕ_ε oscillate on a scale which is not smaller than $O(\varepsilon)$, and is characterized by the following definition.

Definition 2.8 *A family of functions ϕ_ε that is bounded in L^2_{loc} is said to be ε-oscillatory if for every smooth and compactly supported function $\theta(x)$*

$$\limsup_{\varepsilon\to 0}\int_{|\xi|\geq R/\varepsilon}|\widehat{\theta\phi_\varepsilon}(\xi)|^2 d\xi \to 0 \quad \text{as} \quad R\to+\infty. \tag{2.8}$$

A simple and intuitive sufficient condition for (2.8) is that there exist a positive integer j and a constant C independent of ε such that

$$\varepsilon^j\left|\left|\frac{\partial^j\phi_\varepsilon}{\partial x^j}\right|\right|_{L^2_{loc}}\leq C. \tag{2.9}$$

Indeed, if (2.9) is satisfied then

$$\int_{\mathbb{R}^n}|\xi|^j|\widehat{\theta\phi_\varepsilon}|^2 d\xi\leq\frac{C}{\varepsilon^j}$$

and therefore

$$\int_{|\xi|\geq R/\varepsilon}|\widehat{\theta\phi_\varepsilon}(\xi)|^2 d\xi\leq\left(\frac{\varepsilon}{R}\right)^j\int_{|\xi|\geq R/\varepsilon}|\xi|^j|\widehat{\theta\phi_\varepsilon}(\xi)|^2 d\xi\leq\frac{C}{\varepsilon^j}\left(\frac{\varepsilon}{R}\right)^j=\frac{C}{R^j}\to 0 \text{ as } R\to+\infty$$

so that (2.8) holds. Condition (2.9) is satisfied, for instance, for high frequency plane waves $\phi_\varepsilon(x)=Ae^{i\xi\cdot x/\varepsilon}$ with wave vector ξ/ε, $\xi\in\mathbb{R}^n$ but not by a similar family with a wave vector ξ/ε^2: $\psi_\varepsilon(x)=Ae^{i\xi\cdot x/\varepsilon^2}$. Another natural example of ε-oscillatory functions is $g_\varepsilon(x)=g(x/\varepsilon)$, where $g(x)$ is a periodic function with a bounded gradient.

In order to curtail the ability of a family of functions to "run away to infinity" (as it happens with the family (2.7)), we introduce the following definition.

Definition 2.9 *A bounded family $\phi_\varepsilon(x)\in L^2(\mathbb{R}^n)$ is said to be compact at infinity if*

$$\limsup_{\varepsilon\to 0}\int_{|x|\geq R}|\phi_\varepsilon(x)|^2 dx\to 0 \text{ as } R\to+\infty. \tag{2.10}$$

The main reason for introducing ε-oscillatory and compact at infinity families of functions is the following theorem concerning weak convergence of energy, i.e. of the integral of the square of the wave function.

Theorem 2.10 *Let ϕ_ε be a pure, uniformly bounded family in L^2_{loc} with the limit Wigner measure $\mu(x,k)$. Then, if $|\phi_\varepsilon(x)|^2$ converges to a measure ν on $\mathbb{R}^n$, we have*

$$\int_{\mathbb{R}^n}\mu(\cdot,dk)\leq\nu \tag{2.11}$$

with equality if and only if ϕ_ε is an ε-oscillatory family. Moreover, we also have

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \mu(dx, dk) \leq \limsup_{\varepsilon\to 0} \int_{\mathbb{R}^n} |\phi_\varepsilon(x)|^2 dx \tag{2.12}$$

with equality holding if and only if ϕ_ε is ε-oscillatory and compact at infinity. In this case lim sup *can be replaced by* lim *in the right side of (2.12).*

With this theorem and the positivity property we can interpret $\mu(x, k)$ as the limit phase space energy density of the family ϕ_ε, that is, the energy density resolved over directions and wave numbers.

The evolution of the Wigner transform

We will now derive the evolution equation for the Wigner measure of a family of functions $\phi_\varepsilon(t, x)$ that satisfy the semiclassical Schrödinger equation

$$i\varepsilon\frac{\partial\phi_\varepsilon}{\partial t} + \frac{\varepsilon^2}{2}\Delta\phi_\varepsilon - V(x)\phi_\varepsilon(x) = 0 \tag{2.13}$$

with a smooth potential $V(x)$. The initial data $\phi_\varepsilon(0, x) = \phi^0_\varepsilon(x)$ forms an ε-oscillatory and compact at infinity family of functions uniformly bounded in $L^2(\mathbb{R}^n)$. Physically, we are in the regime where the potential varies on the scale much larger than the initial data. In particular, if $V(t, x)$ is a random potential, we should be thinking of the regime $\lambda \ll l_c$ in the terminology of the introduction. As (2.13) preserves the L^2-norm of solutions, the family $\phi_\varepsilon(t, x)$ is bounded in $L^2(\mathbb{R}^n)$ for each $t \geq 0$ and it makes sense to define the Wigner transform

$$W_\varepsilon(t, x, k) = \int \phi_\varepsilon\left(t, x - \frac{\varepsilon y}{2}\right)\bar{\phi}_\varepsilon\left(t, x + \frac{\varepsilon y}{2}\right)e^{ik\cdot y}\frac{dy}{(2\pi)^n}. \tag{2.14}$$

We may obtain the equation for the limit Wigner transform by differentiating (2.14) with respect to time, and using (2.13). We arrive at the following equation for the Wigner transform

$$W^\varepsilon_t + k\cdot\nabla_x W^\varepsilon = \frac{1}{i\varepsilon}\int_{\mathbb{R}^n} e^{ip\cdot x}\hat{V}(p)\left[W^\varepsilon(x, k - \frac{\varepsilon p}{2}) - W^\varepsilon(x, k + \frac{\varepsilon p}{2})\right]\frac{dp}{(2\pi)^n}. \tag{2.15}$$

The limit Wigner measure $W(t, x, k)$ satisfies the Liouville equation in phase space

$$W_t + k\cdot\nabla_x W - \nabla V\cdot\nabla_k W = 0 \tag{2.16}$$

with the initial condition $W(0, x, k) = W_0(x, k)$. We have the following proposition.

Proposition 2.11 *Let the family $\phi^0_\varepsilon(x)$ be uniformly bounded in $L^2(\mathbb{R}^n)$ and pure and let $W_0(x, k)$ be its Wigner measure. Then the Wigner transforms $W_\varepsilon(t, x, k)$ converge uniformly on finite time intervals in $\mathcal{S}'$ to the solution of (2.16) with the initial data $W(0, x, k) = W_0(x, k)$.*

Let us now compare the information one may obtain from the Liouville equation (2.16) to the standard geometric optics approach. First, we derive the eikonal and transport equations for the semiclassical Schrödinger equation (2.13). We consider initial data of the form

$$\phi^\varepsilon(0, x) = e^{iS_0(x)/\varepsilon}A_0(x) \tag{2.17}$$

with a smooth, real valued initial phase function $S_0(x)$ and a smooth compactly supported complex valued initial amplitude $A_0(x)$. We then look for an asymptotic solution of (2.13) in the same form as the initial data (2.17), with an evolved phase and amplitude

$$\phi^\varepsilon(t,x) = e^{iS(t,x)/\varepsilon}(A(t,x) + \varepsilon A_1(t,x) + \dots). \tag{2.18}$$

Inserting this form into (2.13) and equating the powers of ε we get evolution equations for the phase and amplitude

$$S_t + \frac{1}{2}|\nabla S|^2 + V(x) = 0, \qquad S(0,x) = S_0(x) \tag{2.19}$$

and

$$(|A|^2)_t + \nabla \cdot (|A|^2 \nabla S) = 0, \qquad |A(0,x)|^2 = |A_0(x)|^2. \tag{2.20}$$

The phase equation (2.19) is called the eikonal and the amplitude equation (2.20) the transport equation. The eikonal equation that evolves the phase is nonlinear and, in general, it will have a solution only up to some finite time t^* that depends on the initial phase.

How are the eikonal and transport equations related to the Liouville equation (2.16)? As we have computed before, for the WKB initial data (2.17) the initial Wigner distribution has the form

$$W_0(x,k) = |A_0(x)|^2 \delta(k - \nabla S_0(x)). \tag{2.21}$$

As long as the geometric optics approximation (2.18) remains valid we expect the solution of the Liouville equation (2.16) to have the same form:

$$W(t,x,k) = |A(t,x)|^2 \delta(k - \nabla S(t,x)). \tag{2.22}$$

We insert this ansatz into (2.16) :

$$\left(\frac{\partial}{\partial t} + k \cdot \nabla_x - \nabla V \cdot \nabla_k\right) \big(|A(t,x)|^2 \delta(k - \nabla S(t,x))\big) = 0. \tag{2.23}$$

or, equivalently,

$$\begin{aligned} &\delta(k - \nabla S) \left(\frac{\partial}{\partial t} + k \cdot \nabla_x - \nabla V \cdot \nabla_k\right) (|A(t,x)|^2) \\ &+|A(t,x)|^2 \sum_{m,p=1}^{n} \left(\frac{\partial^2 S}{\partial t \partial x_m} + k_p \frac{\partial^2 S}{\partial x_p \partial x_m} - \frac{\partial V}{\partial x_m}\right) D_m = 0, \end{aligned} \tag{2.24}$$

where

$$D_m = \delta(k_1 - S_{x_1}) \dots \delta(k_{m-1} - S_{x_{m-1}}) \delta'(k_m - S_{x_m}) \delta(k_{m+1} - S_{x_{m+1}}) \dots \delta(k_n - S_{x_n}).$$

Equating similar terms in (2.24) we obtain the transport equation (2.20) from the term in the first line, while the coefficient at D_m gives the eikonal equation (2.19) differentiated with respect to x_m. Expression (2.22) holds of course only until the time when the solution of the eikonal equation stops being smooth.

Let us see what happens with the Wigner measure when a caustic forms. Consider the Schrödinger equation (2.13) with $V = 0$ – the corresponding Liouville equation is

$$W_t + k \cdot \nabla_x W = 0, \quad W(0, x, k) = W_0(x, k). \tag{2.25}$$

Its solution is $W(t, x, k) = W_0(x - kt, k)$ and clearly exists for all time. If the initial phase $S_0(x) = -x^2/2$ with a smooth initial amplitude $A_0(x)$ then the Wigner transform at $t = 0$ is $W_0(x, k) = |A_0(x)|^2 \delta(k + x)$ so that solution of (2.25) is

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

This means that at the time $t = 1$ the Wigner measure

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

is no longer singular in wave vectors k but rather in space being concentrated at $x = 0$. This is the caustic point. On the other hand, solution of the eikonal equation (2.19) with the same initial phase and $V = 0$ is given by $S(t, x) = -x^2/(2(1 - t))$ – we see that the same caustic appears at $t = 1$. The transport equation becomes

$$(|A|^2)_t - \frac{x}{1 - t} \cdot \nabla(|A|^2)_t - \frac{n}{1 - t}|A|^2.$$

The corresponding trajectories satisfy

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

and are given by $X(t) = x(1 - t)$ – hence they all arrive to the point $x = 0$ at the time $t = 1$. At this time the geometric optics approximation breaks down and is no longer valid while the solution of the Liouville equation exists beyond this time.

We see that from the Wigner distribution we can recover the information contained in the leading order of the standard high frequency approximation. In addition, it provides flexibility to deal with initial data that is not of the form (2.21).

2.2 Random geometric optics: "short" times

We now assume that the potential $V(t, x)$ is random, weak and varies on the scale much larger than the initial data. More precisely, we consider the semiclassical Schrödinger equation

$$i\varepsilon\phi_t + \frac{\varepsilon^2}{2}\Delta\phi - \delta V(x)\phi = 0 \tag{2.26}$$

with the ε-oscillatory initial data $\phi(0, x) = \phi_0^\varepsilon(x)$. This equation is written on the scale of the variations of the random potential, and $\delta \ll 1$ is the parameter measuring its strength. Passing to the high frequency limit $\varepsilon \to 0$ we obtain the Liouville equation for the Wigner measure of the family $\phi_\varepsilon(t, x)$:

$$\frac{\partial W}{\partial t} + k \cdot \nabla_x W - \delta \nabla_x V(x) \cdot \nabla_k W = 0, \tag{2.27}$$

with the initial data $W(0,x,k)=W_0(x,k)$, the Wigner measure of the family $\phi_0^\varepsilon(x)$. As the parameter $\delta \ll 1$ is small, the effect of the randomness will be felt only after long times. We will build our analysis of (2.27) slowly, starting with relatively short times, and later for the long times. We will assume that $V(x)$ is a spatially homogeneous random process with mean zero and the correlation function $R(x)$:

$$\langle V(x)\rangle = 0, \quad R(x) = \langle V(y)V(x+y)\rangle. \tag{2.28}$$

It will be convenient for us to use the correlation matrix for the force ∇V:

$$\Big\langle \frac{\partial V(y)}{\partial y_i}\frac{\partial V(x+y)}{\partial y_j}\Big\rangle = -\frac{\partial^2 R(x)}{\partial x_i \partial x_j}. \tag{2.29}$$

The characteristics at short times

We begin with the very basic theory of characteristics in a weakly random medium – this material originated in the classical paper by J.B. Keller [7]. The characteristics for the Liouville equation (2.27) are

$$\frac{dX}{dt} = -K(t), \quad \frac{dK}{dt} = \delta \nabla V(X(t)), \quad X(0)=x, \quad K(0)=k. \tag{2.30}$$

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

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

We insert this expansion into the characteristics (2.30), and get in the leading order:

$$X_0(t) = x - k_0 t, \quad K_0(t) = k.$$

As expected, in the leading order the characteristics are straight lines. The first order correction in δ is

$$K_1(t) = \int_0^t \nabla V(X_0(s))ds = \int_0^t \nabla V(x-ks)ds, \tag{2.31}$$

and

$$X_1(t) = \int_0^t K_1(s)ds = \int_0^t (t-s)\nabla V(x-ks)ds. \tag{2.32}$$

Naively, in order to see how long this approximation should hold, we estimate that during a time T we would get $K_1(T) \sim T$, and $X_1(T)$ of the order T^2 meaning that we would need $\delta T^2 \ll 1$, or $T \ll \delta^{-1/2}$ for the spatial trajectory to stay close to the straight line. Let us now see how randomness affects this ballpark estimate – we have, as in (1.9):

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

with the diffusion coefficient

$$D = -\int_{-\infty}^{\infty} \Delta R(ks)ds. \tag{2.33}$$

With a little bit more work, one can show that an appropriate rescaling of $K_1(t)$ converges to a Brownian motion with the diffusion matrix

$$D_{ij} = -\int_{-\infty}^{\infty} \frac{\partial^2 R(ks)}{\partial x_i \partial x_j} ds. \tag{2.34}$$

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

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

and, once again, with a bit more work it can be shown that an appropriate rescaling of $X(t)$ converges, at large times to the time integral of the Brownian motion with the diffusion matrix D_{ij}. The above computations indicate that the simple perturbation expansion should hold for times T such that

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

that is, for times of the order $T \sim \delta^{-2/3}$, which is much longer than the "deterministic prediction" $T \sim \delta^{-1/2}$.

Formally, this means that for large times (but much smaller than $\delta^{-2/3}$), the expected value of the solutions of the Liouville equation (2.27) is well-approximated by the solutions of the Fokker-Planck kinetic equation

$$\frac{\partial \bar{W}}{\partial t} + k\cdot\nabla_x \bar{W} = \delta^2 \sum_{i,j=1}^{n} D_{ij}\frac{\partial^2 \bar{W}}{\partial k_i \partial k_j}, \tag{2.35}$$

that is, $\langle W(t,x,k)\rangle \approx \bar{W}(t,x,k)$. This is probably the simplest way to get to a kinetic description of waves in random media. Instead of trying to make this approximation result precise, for times $t \ll \delta^{-2/3}$, let us explain why such result, while providing a very nice "hooligan's derivation of the kinetic limit", can not "truly hold" for longer times, when the deviation of the characteristics from straight lines will be not small. The problem is that the original characteristics (2.30) preserve the classical Hamiltonian:

$$\omega(x,k) = \frac{k^2}{2} + V(x),$$

that is, $\omega(X(t),K(t)) = \omega(X(0),K(0))$. In particular, if, say, $V(x)$ is a bounded random potential, it is impossible for $K(t)$ to behave as a Brownian motion for large times. Nevertheless, the overall picture described above is not too wrong, and in the next step we will see how it can be naturally modified to see what happens at large times.

2.3 Random geometric optics: the long time limit

A particle in a random Hamiltonian

We will now study the "truly" long time asymptotics of geometric optics in a weakly random medium. This problem can be analyzed in the general setting of a particle in a weakly random

Hamiltonian field:

$$\frac{dX^\delta}{dt} = \nabla_k H_\delta, \quad \frac{dK^\delta}{dt} = -\nabla_x H_\delta, \quad X^\delta(0) = 0, \quad K^\delta(0) = k_0, \tag{2.36}$$

with a random Hamiltonian of the form $H_\delta(x,k) = H_0(k) + \delta H_1(x,k)$. Here $H_0(k)$ is the background Hamiltonian and $H_1(x,k)$ is a random perturbation, while the small parameter $\delta \ll 1$ measures the relative strength of random fluctuations. This was done in [1] and [9]. Here, we will resist the temptation to describe the general results, and restrict ourselves to the case at hand, with $H_0(k) = |k|^2/2$ and $H_1(x,k) = V(x)$, which simplifies some considerations. Thus, we are interested in the Liouville equations

$$\frac{\partial \phi}{\partial t} + k \cdot \nabla_x \phi - \delta \nabla V(x) \cdot \nabla_k \phi = 0, \tag{2.37}$$

and the corresponding characteristics

$$\frac{dX}{dt} = K, \quad \frac{dK}{dt} = -\delta \nabla_x V(X), \quad X(0) = 0, \quad K(0) = k_0, \tag{2.38}$$

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

$$\langle V(x) \rangle = 0, \quad \langle V(y) V(x+y) \rangle = R(x), \tag{2.39}$$

We have already seen that at relatively short times $t \ll \delta^{-2/3}$ the "boosted" deviation $(K(t) - k_0)/\delta$ behaves as a Brownian motion. At the longer times, we are interested not in the deviation from the original direction but in the particle momentum itself. An important simple observation is that (2.38) preserves the Hamiltonian

$$H(x,k) = \frac{k^2}{2} + \delta V(x). \tag{2.40}$$

Hence, the law of any possible limit for the process $K_\delta(t) = K(t/\delta^2)$, as $\delta \to 0$, has to be supported on the sphere $|K(t)| = |k_0|$ (and can not be a regular Brownian motion). Moreover, one would expect the law of the limit process to be isotropic – there is no preferred direction in the problem. One possibility is that $K_\delta(t)$ tends to a uniform distribution on the sphere $\{|k| = |k_0|\}$ – and this is, indeed, what happens at times $t \gg \delta^{-2}$. However, at an intermediate stage, at times of the order δ^{-2}, the process $K_\delta(t)$ converges to the Brownian motion $B_s(t)$ on the sphere (this is an isotropic diffusion such that $|B_s(t)| = 1$ for all t). This intuitive result has been first proved in [8] in dimensions higher than two, and later extended to two dimensions with the Poisson distribution of scatterers in [2], and in a general two-dimensional setting in [10]. The rescaled spatial component $X^\delta(t) = \delta^2 X(t/\delta^2)$ converges to the time integral of the Brownian motion on the sphere:

$$X(t) = \int_0^t B_s(\tau) d\tau.$$

In turn, the long time limit of a momentum diffusion is the standard spatial Brownian motion, and we will see that on the times longer than δ^{-2} the spatial component $X(t)$ converges to the Brownian motion, while $K(t)$ becomes uniformly distributed on the sphere $\{|k| = |k_0|\}$.

Let us mention that another important, (in the context of waves in random media) Hamiltonian

$$H_\delta(x,k) = (c_0 + \delta c_1(x))|k|, \tag{2.41}$$

arises in the geometrical optics limit of the wave equation. We will not address it directly here, but, as we have mentioned, the analysis of the classical Hamiltonian (2.40) can be generalized in a relatively straightforward way – see [9] for details. We stick here with (2.40) solely for the sake of simplicity of presentation.

The Fokker-Planck limit

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

$$\begin{aligned} &\frac{\partial \phi^\delta}{\partial t} + k \cdot \nabla_x \phi^\delta - \delta \nabla V(x) \cdot \nabla_k \phi^\delta = 0, \\ &\phi^\delta(0,x,k) = \phi_0(\delta^2 x, k). \end{aligned} \tag{2.42}$$

There are two assumptions implicitly made here: first is that the random potential is weak, and the second is that the initial data varies on the scale $1/\delta^2$ relative to the scale of the variations of the potential. In the terminology of the introduction, this means that $l_c/L = \delta^2$ – or, we choose the particular observation scale $L = l_c/\delta^2$. One may wonder also as to what happens on other observation scales – we will address this further below.

Let us define the diffusion matrix D_{mn} by

$$D_{ml}(k) = -\frac{1}{|k|}\int_{-\infty}^{\infty} \frac{\partial^2 R(s\hat{k})}{\partial x_n \partial x_m} ds, \quad m,l = 1,\ldots,n. \tag{2.43}$$

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

$$D_{ml}(k) = D(\delta_{mn} - \hat{k}_l \hat{k}_m), \quad D = -\frac{2}{|k|}\int_0^\infty \frac{R'(r)}{r} dr, \quad m,l = 1,\ldots,n. \tag{2.44}$$

We have the following result.

Theorem 2.12 *Let ϕ^δ be the solution of (2.42), with the initial data $\phi_0 \in C_c^\infty(\mathbb{R}^{2d})$, whose support is contained inside a spherical shell $\mathcal{A}(M) = \{(x,k) : M^{-1} < |k| < M\}$ for some positive $M > 0$, and let $\bar{\phi}$ satisfy*

$$\begin{aligned} &\frac{\partial \bar{\phi}}{\partial t} + k \cdot \nabla_x \bar{\phi} = \sum_{m,n=1}^{d} \frac{\partial}{\partial k_m}\left(D_{mn}(k) \frac{\partial \bar{\phi}}{\partial k_n}\right) \\ &\bar{\phi}(0,x,k) = \phi_0(x,k). \end{aligned} \tag{2.45}$$

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

$$\sup_{(t,x,k)\in[0,T]\times K} \left| \mathbb{E}\phi^\delta\left(\frac{t}{\delta^2}, \frac{x}{\delta^2}, k\right) - \bar{\phi}(t,x,k)\right| \leq CT(1 + \|\phi_0\|_{1,4})\delta^{\alpha_0} \tag{2.46}$$

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

Note that

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

and thus the K-process generated by (2.45) is indeed a diffusion process on a sphere $|k| =$ const, or, equivalently, equations (2.45) for different values of $|k|$ are decoupled. Another important point is that the assumption that the initial data does not concentrate close to $k = 0$ is important – if $|k|$ is very small, the particle moves very slowly, and does not have a sufficient time to sample enough of the random medium by the time δ^{-2}.

Beyond the Fokker-Planck limit

Let us now return to the question of what happens to the solutions of the Liouville equation with the initial data that varies on a scale much longer than δ^{-2} – in other words, the observation is taken on even larger scales than described by the Fokker-Planck limit. It is straightforward to see that solutions of the Fokker-Planck equation (2.45) themselves converge in the long time limit to the solutions of the spatial diffusion equation. More, precisely, we have the following result. Let $\bar{\phi}_\gamma(t,x,k) = \bar{\phi}(t/\gamma^2, x/\gamma, k)$, where $\bar{\phi}$ satisfies (2.45) with slowly varying initial data $\bar{\phi}_\gamma(0,t,x,k) = \phi_0(\gamma x, k)$. We also let $w(t,x,|k|)$ be the solution of the spatial diffusion equation:

$$\frac{\partial w}{\partial t} = \sum_{m,n=1}^{d} a_{mn}(|k|)\frac{\partial^2 w}{\partial x_n \partial x_m}, \tag{2.47}$$
$$w(0,x,|k|) = \bar{\phi}_0(x,|k|)$$

with the averaged initial data

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

Here, $d\Omega(\hat{k})$ is the surface measure on the unit sphere $\mathbb{S}^{n-1}$ and Γ_n is the area of an n-dimensional sphere. The diffusion matrix $A := [a_{nm}]$ in (2.47) is given explicitly as

$$a_{ij}(k) = \frac{|k|^2}{\Gamma_{n-1}}\int_{\mathbb{S}^{n-1}}\hat{k}_i\chi_j(k)d\Omega(\hat{k}). \tag{2.48}$$

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

$$\sum_{m,i=1}^{d}\frac{\partial}{\partial k_m}\left(D_{mi}(k)\frac{\partial \chi_j}{\partial k_i}\right) = -\hat{k}_j, \tag{2.49}$$

and when the correlation function $R(x)$ is isotropic, so that D_{mi} is given by (2.44), they are just multiples of $\hat{k}_j$: $a_j(k) = c(|k|)\hat{k}_j$, with an appropriate constant $c(|k|)$ that can be computed explicitly. In that case, the matrix a_{nm} is a multiple of identity, and (2.47) becomes the standard diffusion equation

$$\frac{\partial w}{\partial t} = \bar{a}(|k|)\Delta_x w, \tag{2.50}$$

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

Theorem 2.13 *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* (2.45) *converges as* $\gamma \to 0$ *in* $C([T_*, T]; L^\infty(\mathbb{R}^{2d}))$ *to* $w(t,x,k)$*. Moreover, there exists a constant* $C_0 > 0$*, so that we have*

$$\|w(t,\cdot) - \bar{\phi}_\gamma(t,\cdot)\|_{L^\infty} \le C_0\,(\gamma T + \sqrt{\gamma})\,\|\phi_0\|_{C^1}, \tag{2.51}$$

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

The proof of Theorem 2.13 is based on classical asymptotic expansions and is quite straightforward. As an immediate corollary of Theorems 2.12 and 2.13, we obtain the following result.

Theorem 2.14 *Let* ϕ_δ *be solution of* (2.42) *with the initial data* $\phi_\delta(0,x,k) = \phi_0(\delta^{2+\alpha}x, k)$ *and let* $\bar{w}(t,x)$ *be the solution of the diffusion equation* (2.47) *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* $0 \le \alpha < \alpha_0$ *and all* $0 < T_* \le T$ *we have for all compact sets* $K \subset \mathcal{A}(M)$*:*

$$\sup_{(t,x,k)\in[T_*,T]\times K} \left| w(t,x,k) - \mathbb{E}\bar{\phi}_\delta(t,x,k)\right| \le CT\delta^{\alpha_0-\alpha}, \tag{2.52}$$

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

Theorem 2.14 shows that if the initial data varies on a scale slightly larger than δ^{-2} then we observe spatial diffusion for the solution (and uniform distribution in k) on the appropriate time scale. The requirement that α is small is most likely technical and a constraint of a "perturbative" proof – the result should hold for any $\alpha > 0$.

To summarize: if the initial data for the random Liouville equation

$$\frac{\partial \phi}{\partial t} + k \cdot \nabla_x \phi - \delta V(x) \cdot \nabla_x \phi = 0, \tag{2.53}$$

varies on the scale δ^{-2}: $\phi(0,x) = \phi_0(\delta^2 x, k)$, then on the time scale $t \sim \delta^{-2}$ the expectation of the rescaled solution $\phi_\delta(t,x,k) = \phi(t/\delta^2, x/\delta^2, k)$ converges to the solution of the Fokker-Planck equation. On the other hand, if the initial data varies on an even larger scale: $\phi(0,x,k) = \phi(\delta^{2+\alpha}x, k)$ then on the time scale $t \sim \delta^{-2-2\alpha}$ the expectation of the rescaled field $\phi_\delta(t,x,k) = \phi(t/\delta^{2+2\alpha}, x/\delta^{2+\alpha}, k)$ converges to the solution of the spatial diffusion equation and is uniformly distributed in the directions $\hat{k}$ for each $|k|$ fixed. Thus, the appropriate kinetic limit depends on the scale of the probing signal, which, in turn, determines the proper time scale of the observations.

A formal derivation of the momentum diffusion

We now describe how the momentum diffusion operator in (2.45) can be derived in a quick formal way. We represent the solution of (2.42) as $\phi^\delta(t,x,k) = \psi^\delta(\delta^2 t, \delta^2 x, k)$ and write an asymptotic multiple scale expansion for ψ^δ

$$\psi^\delta(t,x,k) = \bar{\phi}(t,x,k) + \delta\phi_1\left(t, x, \frac{x}{\delta^2}, k\right) + \delta^2\phi_2\left(t, x, \frac{x}{\delta^2}, k\right) + \dots \tag{2.54}$$

We assume formally that the leading order term $\bar{\phi}$ is deterministic and independent of the fast variable $z = x/\delta^2$. We insert this expansion into (2.42) and obtain in the order $O(\delta^{-1})$:

$$\nabla V(z) \cdot \nabla_k \bar{\phi} - k \cdot \nabla_z \phi_1 = 0. \tag{2.55}$$

Let $\theta \ll 1$ be a small positive regularization parameter that will be later sent to zero, and consider a regularized version of (2.55):

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

Its solution is

$$\phi_1(z,k) = -\frac{1}{|k|} \int_0^\infty \sum_{m=1}^d \frac{\partial V(z + s\hat{k})}{\partial z_m} \frac{\partial \bar{\phi}(t,x,k)}{\partial k_m} e^{-\theta s} ds, \tag{2.56}$$

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

$$\frac{\partial \bar{\phi}}{\partial t} + k \cdot \nabla_x \bar{\phi} = \langle \nabla V(z) \cdot \nabla_k \phi_1 \rangle. \tag{2.57}$$

The term in the right side above may be written using expression (2.56) for ϕ_1:

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

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

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

We insert the above expression into (2.57) and obtain

$$\frac{\partial \bar{\phi}}{\partial t} = \sum_{m,n=1}^d \frac{\partial}{\partial k_n} \left(D_{nm}(k) \frac{\partial \bar{\phi}}{\partial k_m} \right) + k \cdot \nabla_x \bar{\phi} \tag{2.58}$$

with the diffusion matrix $D(\hat{k}, k)$ as in (2.43). Observe that (2.58) is nothing but (2.45). However, the naive asymptotic expansion (2.54) may not be justified directly, to the best of my knowledge. The rigorous proof is based on a completely different method.

3 Passive sensor imaging using noisy signals

In this section, we describe how imaging can be done using cross-correlation of signals in the presence of random noise sources. That is, we have the following setup: two (or more) sensors are located at the positions x_1 and x_2, and record the time-dependent wave fields $u(t, x_1)$ and $u(t, x_2)$ that come from a noisy distribution of sources. Our goal is to estimate the travel time from x_1 to x_2, as well as to find any reflectors present in the medium. We will be following the paper [3] by J. Garnier and G. Papanicolaou where a detailed list of references can be found, as well as a much deeper discussion of the problem. The main miracle is the following basic observation. Consider the cross-correlation function of the recorded signals $u(t, x_1)$ and $u(t, x_2)$, with a time lag τ:

$$C_T(\tau, x_1, x_2) = \frac{1}{T}\int_0^T u(t, x_1)u(t + \tau, x_2)dt. \tag{3.1}$$

It turns out that if the noisy sources form a space-time stationary random field, then the cross-correlation encodes the Green's function between the points x_1 and x_2:

$$\frac{\partial C_T(\tau, x_1, x_2)}{\partial \tau} = G(\tau, x_1, x_2) - G(-\tau, x_1, x_2) + o(1), \quad \text{as } T \to +\infty. \tag{3.2}$$

Naturally, it is very rare that the sources are distributed randomly in all of space, and we have sensors in their midst. A more common situation is that the sources are distributed randomly in a bounded set, and the sensors are located away from them. Our goal here is to explain relation (3.2), as well as its generalization to other spatial configurations of random sources, and indicate some implications to the inverse problems. We will see that one important factor is that there should be some energy flux between the sensors. For example, if the line connecting the sensors points toward the noisy sources, the cross-correlation will carry more information about the Green's function than if it is orthogonal to the direction toward the sources. This leads to the idea that "directional diversity is good" – a medium in which waves "propagate in all directions" is better for us than a non-scattering medium. Examples when this is the case are an ergodic cavity when even a small set of random sources will create directional diversity due to reverberations, and a random medium with a large number of (weak) random scatterers. However, a random medium has to have a "just right" transport mean free path: on one hand, it has to be sufficiently small so that multiple scattering would create directional diversity. On the other, the coherent part of the signal traveling from x_1 to x_2 (which is exponentially attenuated in the presence of multiple scattering) should not be too weak meaning that the distance between the sensors should be smaller than the transport mean free path.

The wave equation with noisy sources

We consider the wave equation

$$\frac{1}{c^2(x)}\frac{\partial^2 u}{\partial t^2} - \Delta u = n(t, x). \tag{3.3}$$

Here, $c(x)$ is the deterministic background sound speed, and $n(t,x)$ is a random distribution of sources. We assume that it is stationary in time, with the correlation function

$$\langle n(t_1,x_1)n(t_2,x_2)\rangle = \Gamma(x_1,x_2)F(t_2-t_1). \tag{3.4}$$

We also assume that the coherence time of the sources is much smaller than the travel time between the sensors – this is a very important physical assumption. If we denote the ratio of these time scales by $\varepsilon \ll 1$, the time correlation function $F(t)$ takes the form

$$F_\varepsilon(t) = F\big(\frac{t}{\varepsilon}\big), \tag{3.5}$$

and its Fourier transform is

$$\hat{F}_\varepsilon(\omega) = \varepsilon\hat{F}(\varepsilon\omega). \tag{3.6}$$

For simplicity, we will assume that the spatial correlation function is a modulated delta-function:

$$\Gamma(x_1,x_2) = \theta(x_1)\delta(x_1-x_2). \tag{3.7}$$

The function $\theta(x)$ characterizes the support and strength of the sources. This assumption can be weakened but the spatial decorrelation length of the random sources should be much smaller than the distance between the sensors.

The stationary in time solution of the wave equation has the form

$$u(t,x) = \int_{-\infty}^{t}\int G(t-s,x,y)n_\varepsilon(s,y)dyds = \int_0^\infty\int G(s,x,y)n_\varepsilon(t-s,y)dyds. \tag{3.8}$$

Here, $G(t,x,y)$ is the solution of

$$\frac{1}{c^2(x)}\frac{\partial^2 G}{\partial t^2} - \Delta G = \delta(t)\delta(x-y), \quad t \ge 0, \tag{3.9}$$

with $G(0,x,y) = G_t(0,x,y) = 0$. We may extend $G(t,x,y) = 0$ for $t \le 0$ and write

$$u(t,x) = \int_{-\infty}^{\infty}\int G(s,x,y)n_\varepsilon(t-s,y)dyds. \tag{3.10}$$

The stationarity of $n_\varepsilon(t)$ implies that the wave-fields $u(t,x)$ are themselves stationary, hence the mean of C_T does not depend on T and is given by

$$C_1(\tau,x_1,x_2) := \langle C_T(\tau,x_1,x_2)\rangle = \langle u(0,x_1)u(\tau,x_2)\rangle. \tag{3.11}$$

We may now compute the mean correlation:

$$\begin{aligned} C_1(\tau,x_1,x_2) &= \int G(s_1,x_1,y_1)G(s_2,x_2,y_2)\langle n_\varepsilon(-s_1,y_1)n_\varepsilon(\tau-s_2,y_2)\rangle dy_1dy_2ds_1ds_2 \\ &= \int G(s_1,x_1,y)G(\tau+s_1+s_2,x_2,y)\theta(y)F_\varepsilon(s_2)dyds_1ds_2. \end{aligned} \tag{3.12}$$

This expression may be re-written in the Fourier domain:

$$
\begin{aligned}
C_1(\tau, x_1, x_2) &= \int e^{-i\omega_1 s_1 - i\omega_2(\tau+s_1+s_2)} \hat{G}(\omega_1, x_1, y)\hat{G}(\omega_2, x_2, y)\theta(y)F_\varepsilon(s_2) dy ds_1 ds_2 \frac{d\omega_1 d\omega_2}{(2\pi)^2} \\
&= \int e^{-i\omega_2(\tau+s_2)} \hat{G}(-\omega_2, x_1, y)\hat{G}(\omega_2, x_2, y)\theta(y)F_\varepsilon(s_2) dy ds_2 \frac{d\omega_2}{2\pi} \\
&= \int \overline{\hat{G}(\omega, x_1, y)}\hat{G}(\omega, x_2, y)\hat{F}_\varepsilon(\omega)e^{-i\omega\tau}\theta(y)\frac{dyd\omega}{2\pi}. \qquad (3.13)
\end{aligned}
$$

Here, we use the convention as in [3]:

$$
\hat{f}(\omega) = \int e^{i\omega t} f(t) dt, \quad f(t) = \int e^{-i\omega t} \hat{f}(\omega)\frac{d\omega}{2\pi}.
$$

An important observation is that the correlation is self-averaging in the limit $T \to +\infty$, in other words,

$$
C_T(\tau, x_1, x_2) \to C_1(\tau, x_1, x_2), \quad \text{as } T \to +\infty, \qquad (3.14)
$$

in probability. This is extremely important for potential applications in inverse problems: it follows that the cross-correlation is not random in the large T limit, and is thus an appropriate quantity to be used as an input into the inverse problems. We will not prove it here: the proof is by a direct computation of the variance of $C_T(\tau, x_1, x_2)$ and showing that it tends to zero as $T \to +\infty$, at the rate $O(1/T)$ – see Appendix A in [3] for details.

The Green's function from the correlations in a homogeneous medium

We now show how the Green's function emerges from the correlations in the simplest case: the medium is homogeneous, $c(x) = c_0$, and the sources are uniformly (statistically) distributed in space: $\theta(x) \equiv 1$. In this exact situation, the wave field will diverge so we need to introduce some absorption:

$$
\frac{1}{c_0^2}\left(\frac{1}{T_a} + \frac{\partial}{\partial t}\right)^2 u - \Delta u = n_\varepsilon(t, x). \qquad (3.15)
$$

Proposition 3.1 *Assume that the dimension $n = 3$ and the sources are distributed statistically homogeneously in space: $\theta(x) \equiv 1$, then*

$$
\frac{\partial C_1(\tau, x_1, x_2)}{\partial \tau} = -\frac{c_0^2 T_a}{4} e^{-|x_1 - x_2|/(c_0 T_a)} \left[F_\varepsilon \star G(\tau, x_1, x_2) - F_\varepsilon \star G(-\tau, x_1, x_2)\right]. \qquad (3.16)
$$

Here, $\star$ denotes the convolution in τ, and

$$
G(t, x_1, x_2) = \frac{1}{4\pi|x_1 - x_2|}\delta\Big(t - \frac{|x_1 - x_2|}{c_0}\Big)
$$

is the Green's function for the wave equation in a homogeneous medium without the dissipation.

Now, if the decoherence time ε is much smaller than the travel time between the sensors, that is, if $\varepsilon \ll 1$, we can approximate $F_\varepsilon(t)$ by the delta-function, and (3.16) turns into

$$\frac{\partial C_1(\tau, x_1, x_2)}{\partial \tau} \approx -\frac{c_0^2 T_a}{4} e^{-|x_1 - x_2|/(c_0 T_a)} \left[G(\tau, x_1, x_2) - G(-\tau, x_1, x_2) \right]. \tag{3.17}$$

Therefore, we may estimate the travel time between x_1 and x_2, up to the decorrelation time of the random sources.

In order to prove Proposition 3.1, recall that the Green's function of a homogeneous medium with dissipation is

$$G_a(t, x_1, x_2) = G(t, x_1, x_2) e^{-t/T_a}.$$

The cross-correlation function is given then by (recall that $\theta \equiv 1$)

$$\begin{aligned} C_1(\tau, x_1, x_2) &= \int G_a(s, x_1, y) G_a(\tau + s + s', x_2, y) F_\varepsilon(s') dy ds ds' \\ &= \frac{1}{16\pi^2} \int \frac{e^{-s/T_a}}{|x_1 - y|} \frac{e^{-(\tau+s+s')/T_a}}{|x_2 - y|} \delta\Big(s - \frac{|x_1 - y|}{c_0}\Big) \delta\Big(\tau + s + s' - \frac{|x_2 - y|}{c_0}\Big) F_\varepsilon(s') dy ds ds' \\ &= \frac{1}{16\pi^2} \int \frac{e^{-|x_1 - y|/(c_0 T_a)}}{|x_1 - y|} \frac{e^{-|x_2 - y|/(c_0 T_a)}}{|x_2 - y|} F_\varepsilon(\tau - \frac{|x_2 - y| - |x_1 - y|}{c_0}) dy. \end{aligned}$$

Let us use the coordinate axes such that $x_1 = (h, 0, 0)$ and $x_2 = (-h, 0, 0)$, and use the change of variables for $y = (y_1, y_2, y_3)$:

$$y_1 = h \sin\theta \cosh r, \quad y_2 = h \cos\theta \sinh r \cos\psi, \quad y_3 = h \cos\theta \sinh r \sin\psi,$$

with $r \in (0, +\infty)$, $\theta \in (-\pi/2, \pi/2)$ and $\psi \in (0, 2\pi)$. The Jacobian is

$$J = h^3 \cos\theta \sinh r (\cosh^2 r - \sin^2\theta),$$

while

$$|x_1 - y| = h(\cosh r - \sin\theta), \quad |x_2 - y| = h(\cosh r + \sin\theta).$$

Using these expressions in the integral above we obtain

$$\begin{aligned} C_1(\tau, x_1, x_2) &= \frac{h^3}{16\pi^2} \int_0^\infty dr \int_{-\pi/2}^{\pi/2} d\theta \int_0^{2\pi} d\psi \frac{\cos\theta \sinh r (\cosh^2 r - \sin^2\theta) e^{-2h\cosh r/(c_0 T_a)}}{h^2 (\cosh r - \sin\theta)(\cosh r + \sin\theta)} \\ &\times F_\varepsilon(\tau - \frac{2h \sin\theta}{c_0}) = \frac{h}{8\pi} \int_0^\infty dr \int_{-\pi/2}^{\pi/2} d\theta \cos\theta \sinh r e^{-2h\cosh r/(c_0 T_a)} F_\varepsilon(\tau - \frac{2h\sin\theta}{c_0}). \end{aligned}$$

After another change of variables, $w = h\cosh r$ and $s = (2h/c_0)\sin\theta$, this becomes

$$C_1(\tau, x_1, x_2) = \frac{c_0}{16\pi h} \int_h^\infty dw \int_{-2h/c_0}^{2h/c_0} ds e^{-2w/(c_0 T_a)} F_\varepsilon(\tau - s) = \frac{c_0^2 T_a e^{-2h/(c_0 T_a)}}{32\pi h} \int_{-2h/c_0}^{2h/c_0} ds F_\varepsilon(\tau - s).$$

Differentiating in τ leads to

$$\frac{\partial C_1(\tau, x_1, x_2)}{\partial \tau} = \frac{c_0^2 T_a e^{-2h/(c_0 T_a)}}{32\pi h} [F_\varepsilon(\tau + 2h/c_0) - F_\varepsilon(\tau - 2h/c_0)]. \tag{3.18}$$

Now, as $|x_1 - x_2| = 2h$, we get (3.16).

Travel time estimation with spatially localized noisy sources

We now consider the cross-correlation of signals when the noisy sources are localized, so that the function $\theta(x) \not\equiv 1$. The medium has a smooth sound speed profile $c_0(x)$, which is homogeneous outside of a large sphere that encloses both the sensors and the sources. The outgoing time-harmonic Green's function is the solution of

$$\Delta_x \hat{G}_0(\omega, x, y) + \frac{\omega^2}{c_0^2(x)} \hat{G}_0(\omega, x, y) = -\delta(x - y), \tag{3.19}$$

together with the radiation condition at infinity. When the medium is uniform, the Green's function is

$$\hat{G}_0(\omega, x, y) = \frac{e^{ik|x-y|}}{4\pi|x - y|}, \quad k = \omega/c_0,$$

in three dimensions. When the medium is slowly varying, the high frequency (WKB, for Wentzell-Kramers-Brillouin) asymptotics for Green's function is

$$\hat{G}_0\Big(\frac{\omega}{\varepsilon}, x, y\Big) \sim a(x, y) e^{i\omega\tau(x,y)/\varepsilon}. \tag{3.20}$$

Here, the functions $a(x, y)$ and $\tau(x, y)$ are smooth except at $x = y$. The travel time $\tau(x, y)$ can be obtained from Fermat's principle:

$$\tau(x, y) = \inf_{\gamma} T_\gamma, \tag{3.21}$$

where the infimum is taken over all C^1-curves $\gamma : [0, T_\gamma] \to \mathbb{R}^3$ such that $X(0) = x$, $X(T_\gamma) = y$ and

$$\left|\frac{dX_t}{dt}\right| = c_0(X_t).$$

The minimizing curve in (3.21) is the ray, and we assume that the profile $c_0(x)$ is such that there is a unique ray joining any two points x and y in the region of interest. Recall that the rays satisfy the Hamiltonian system

$$\begin{aligned}
\frac{dX}{dt} &= c_0(X)\hat{K}, \quad X(0) = x, \\
\frac{dK}{dt} &= -\nabla c_0(X)|K(t)|, \quad K(0) = k.
\end{aligned}$$

Our assumption on the uniqueness of the ray means that for every x and y there exists one k with $|k| = 1$, such that $X(0) = x$, $K(0) = k$, and $X(t) = y$ for some t, and then this time t is the travel time from x to y.

Lemma 3.2 *If $\nabla_y \tau(x_1, y) = \nabla_y \tau(x_2, y)$ then x_1 and x_2 lie on the same ray issuing from y and*

$$|\tau(x_1, y) - \tau(x_2, y)| = \tau(x_1, x_2).$$

On the other hand, if $\nabla_y \tau(x_1, y) = -\nabla_y \tau(x_2, y)$ then x_1 and x_2 lie on the opposite sides of the same ray issuing from y and

$$\tau(x_1, y) + \tau(x_2, y) = \tau(x_1, x_2).$$

Proof. Let us look at the ray connecting x_1 and y. We can look at it as "starting at y in the direction k_0" or, equivalently, as "starting at x_1 in the direction k_1". Then we have

$$X(t; y, k_0) = X(\tau - t; x_1, k_1), \quad K(t; y, k_0) = -K(\tau - t; x_1, k_1).$$

Note that we also have

$$\nabla_2 \tau(x_1, X(t; x_1, k_1)) = \frac{1}{c_0(X(t; x_1, k_1))} \hat{K}(t; x_1, k_1),$$

with the gradient taken with respect to the second variable. Let us use this identity at the time $t = \tau(x_1, y)$:

$$-k_0 = K(\tau; x_1, k_1) = \nabla_2 \tau(x_1, X(\tau; x_1, k_1)) = \nabla_y \tau(x_1, y).$$

Therefore, if $\nabla_y \tau(x_1, y) = \nabla_y \tau(x_2, y)$ then the ray connecting x_1 and y, and x_2 and y has to start at the same angle, whence x_1 and x_2 lie on the same ray going through y, and on the same side of y. On the other hand, if $\nabla_y \tau(x_1, y) = -\nabla_y \tau(x_2, y)$ then, for the same reason, they have to lie on the same ray passing through y but on two different sides from y.

We are now ready to prove the following proposition.

Proposition 3.3 *As $\varepsilon \to 0$, the cross-correlation $C_1(\tau, x_1, x_2)$ has singular components if and only if the ray going through x_1 and x_2 reaches into the source region, that is, into the support of the function θ. In this case, there are either one or two singular components at $\tau = \pm\tau(x_1, x_2)$. More precisely, any ray going from the source region to x_2 and then to x_1 gives rise to a singular component at $\tau = -\tau(x_1, x_2)$, while rays going frist from the source region to x_1 and then to x_2 give rise to the singular component at $\tau = \tau(x_1, x_2)$.*

This proposition explains why travel time estimation is bad when the ray joining the two sensors is nearly orthogonal to the direction toward the noisy sources.

In order to prove Proposition 3.3 we use expression (3.13), and recall that $\hat{F}_\varepsilon(\omega) = \varepsilon \hat{F}(\varepsilon\omega)$:

$$\begin{aligned} C_1(\tau, x_1, x_2) &= \int \overline{\hat{G}(\omega, x_1, y)} \hat{G}(\omega, x_2, y) \hat{F}_\varepsilon(\omega) e^{-i\omega\tau} \theta(y) \frac{dy d\omega}{2\pi} \\ &= \varepsilon \int \overline{\hat{G}(\omega, x_1, y)} \hat{G}(\omega, x_2, y) \hat{F}(\varepsilon\omega) e^{-i\omega\tau} \theta(y) \frac{dy d\omega}{2\pi} \\ &= \int \overline{\hat{G}(\frac{\omega}{\varepsilon}, x_1, y)} \hat{G}(\frac{\omega}{\varepsilon}, x_2, y) \hat{F}(\omega) e^{-i\omega\tau/\varepsilon} \theta(y) \frac{dy d\omega}{2\pi}. \end{aligned} \tag{3.22}$$

Using the WKB-approximation of Green's function gives

$$C_1(\tau, x_1, x_2) = \int \bar{a}(x_1, y) a(x_2, y) \hat{F}(\omega) e^{i\omega T(y)/\varepsilon} \frac{dy d\omega}{2\pi},$$

with the phase

$$\omega T(y) = \omega[\tau(x_2, y) - \tau(x_1, y) - \tau].$$

The stationary phase method implies that the main contribution to the integral comes from the critical points of the phase, where

$$\frac{\partial}{\partial\omega}(\omega T(y)) = 0, \quad \nabla_y(\omega T(\omega, y)) = 0.$$

It follows that

$$\tau(x_2, y) - \tau(x_1, y) = \tau, \quad \nabla_y \tau(x_2, y) = \nabla_y \tau(x_1, y). \tag{3.23}$$

Now, Lemma 3.2 and the second condition in (3.23) imply that x_1 and x_2 lie on the same side of a ray issuing from y. If the points are aligned so that $y \to x_1 \to x_2$ then the first condition in (3.23) implies that $\tau = \tau(x_1, x_2)$. On the other hand, if they are aligned so that $y \to x_2 \to x_1$ then the first condition in (3.23) implies that $\tau = -\tau(x_1, x_2)$. Finally, in order for a stationary point y to contribute to the integral, we should have $\theta(y) \neq 0$, which means that y has to lie in the source region.

Here, we have only touched upon the possibilities of imaging using passive sensors and random noise sources. This method extends to many other imaging problems, such as in the presence of reflectors, and in heterogeneous media, and we refer to [3] for various extensions, as well as to [4] for more recent results.

References

[1] G. Bal, T. Komorowski and L. Ryzhik, Self-averaging of the Wigner transform in random media, Comm. Math. Phys., 2003, **242**, 81–135.

[2] 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.

[3] J. Garnier and G. Papanicolaou, Passive sensor imaging using cross correlations of noisy signals in a scattering medium, SIAM J. Imaging Sciences, **2**, 2009, 396–437.

[4] J. Garnier, G. Papanicolaou, A. Semin, and C. Tsogka, Signal-to-noise ratio estimation in passive correlation-based imaging, SIAM J. Imaging Sciences, **6**, 2013, 1092–1110.

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

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

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

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

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

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

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

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