Kinetic and Related Models
©American Institute of Mathematical Sciences
Volume **3**, Number **4**, December **2010**

doi:10.3934/krm.2010.3.xx

pp. **xx–XX**

KINETIC LIMITS FOR WAVES IN A RANDOM MEDIUM

Guillaume Bal

Department of Applied Physics & Applied Mathematics
Columbia University, New York, NY 10027, USA

Tomasz Komorowski

Institute of Mathematics, UMCS, pl. Marii Curie-Skłodowskiej 1, 20-031, Lublin
and
IMPAN, ul. Śniadeckich 8, 00-956 Warsaw, Poland

Lenya Ryzhik

Department of Mathematics, Stanford University
Stanford, CA 94305, USA

(Communicated by Pierre Degond)

Contents

2000 *Mathematics Subject Classification.* Primary: 60H25; Secondary: 35Q40.

Key words and phrases. Waves in random media, Particles in random flow, Wigner transform, Wave-wave correlation, Kinetic equation, Radiative transport equation, Fokker-Planck equation, Self-averaging, Scintillation function, Time reversal.

Guillaume Bal and Lenya Ryzhik were partially supported by NSF and AFOSR. Tomasz Komorowski has been partially supported by the grant grant NN201419139 from the Polish Ministry of Science and Higher Education.

1. **Introduction.** This paper reviews several mathematical techniques that have been developed to analyze the asymptotic behavior of waves and particles propagating in a heterogeneous medium. The heterogeneous medium is typically not known precisely and is thus modeled as a realization of a collection of (random) media with known statistics. In many applications, the scale at which the medium varies, the correlation length l_c, is much shorter than the scale at which observations are made, the size of the domain L. We are interested in the asymptotic description of physical observables as $l_c/L \ll 1$. In many problems, the medium also varies rapidly in time with the correlation time t_c. If v is a characteristic speed of propagation, then we also assume that $vt_c/L \ll 1$ in the asymptotic description.

We should also underline that the problems discussed here involve weakly random media – this means that the overall effect of the weak fluctuations on the macroscopic quantities becomes significant only after a sufficiently long time – and these long times are the regime we are interested in, with the typical goal of capturing the long time effect of the randomness by means of an effective macroscopic equation. Hence, the notion of "weak" fluctuations does not mean that on the time scales we consider the effect of the random medium is small. Nevertheless, these models behave usually in a qualitatively different way from "strongly" random models that we do not consider here.

Particles in random flows. Waves in heterogeneous media form our main object of study. We begin our review by the simpler problem of particle propagation in a random flow. The main object of interest is then the dynamics of the spatial density of particles. As particles propagate in a temporally and spatially varying, centered, random flow, their speeds undergo rapid changes, whose macroscopic effects tend to cancel out by an application of the law of large numbers (LLN). As time progresses, a correction to the LLN emerges as an application of the central limit theorem (CLT). Since this argument is central to many techniques reviewed here, we describe the CLT method in detail with complete proofs in Section 2. Under appropriate assumptions on the random flow, we show that the ensemble average of the spatial particle density asymptotically solves a *diffusion* equation.

This first result provides a macroscopic description of particle densities in a statistical average sense. It fails to provide a description of particle propagation in the physically given realization of the heterogeneous medium. What we also would like to know is whether the *deterministic* limiting diffusion equation is a good approximation for the whole random distribution of particle densities, and not only its ensemble average. When the whole distribution of particle densities converges *in probability* to its deterministic limit, then we say that it is *statistically stable*. This is an important property in many applications that we shall describe in more detail below.

The methodology we follow to prove statistical stability consists of considering correlations, which are ensemble averages of quadratic quantities in the particle density, and deriving their limiting equations. If X_ε is a sequence of random variables for which we can show that $\mathbb{E}\left\{X_\varepsilon\right\}-X$ and $\mathbb{E}\left\{X_\varepsilon^2\right\}-(\mathbb{E}\left\{X_\varepsilon\right\})^2$ go to 0 with ε, then we are guaranteed that X_ε converges to X in mean square, and hence in probability by an application of the Chebyshev inequality. Although this does not give access to *almost sure* convergence, which would be the ideal tool to show that a quantity is independent of the realization of the random medium, its generalization to correlations provides a basic tool to show statistical stability as it was described above. Moreover, we observe that correlations, after an appropriate change of coordinates, are the solutions of (phase space) linear kinetic equations. Our first encounter with kinetic models is described at the end of Section 2.

Waves in random media and observables. Particle propagation in random flows displays several of the main difficulties that we face in the analysis of wave propagation in random media. The latter is, however, significantly more challenging mathematically as it involves the propagation of a whole (wave) field, that is both delocalized and parameterized by an infinite number of degrees of freedom, rather than that of a particle, which solves a finite system of ordinary differential equations. Moreover, interesting applications of wave propagation involve typical wavelengths $\lambda \ll L$ that are typically much smaller than the overall distance of propagation.

Waves in this paper will primarily mean quantum waves solution of an evolution Schrödinger equation. We will also consider classical waves, such as solutions of a system of acoustic wave equations but much less than the Schrödinger equation both because fewer results are available for the wave equation, and because the technicalities that are not small even in the Schrödinger case grow significantly for the wave equations. In all settings, the random medium (random potential for Schrödinger or random sound speed in acoustics) will depend on space and also possibly on time. Depending on the relationship between λ, l_c, and vl_t, different macroscopic regimes will emerge. What all regimes have in common is that the observables for which limiting models are available are not the wave fields themselves but rather field-field correlations, which are quadratic observables in the wave fields.

That quadratic observables play an essential role is not surprising and similar to the approximation of quantum mechanics by classical mechanics: whereas quantum waves are spatially varying fields, classical particles require a phase space (position and momentum) description. Except for one result describing the limiting behavior of a properly rescaled wave field in Section 5.2, all limiting theorems presented in this review are for quadratic observables.

The Wigner transforms. The natural tool to derive kinetic limits is the Wigner transform of two wave fields. It is defined as the Fourier transform in the $y \to k$ variable of the field-field correlation written as $\psi_1(t, x-y/2)\psi_2^*(t, x+y/2)$

and satisfies a kinetic evolution equation in the phase space variables (x, k). It is known to be very useful in many microlocal analysis problems, and it is indispensable in wave propagation in random media as well for the very simple reason that multiple scattering by the medium heterogeneities creates waves propagating in many directions at each point in the physical space making the microlocal analysis tools necessary. We recall some basic facts about the Wigner transforms in Section 3 as well as its application to the derivation of classical mechanics, in the form of a Liouville equation, from wave fields propagating in slowly varying media. Both quantum and classical wave propagation models are considered.

Kinetic models for wave equations. Closed-form equations for the Wigner transform can be obtained independent of the regime of wave propagation. That such equations can be obtained for a quadratic quantity in the wave field should not come as a surprise since the number of independent variables is twice as large. Once such equations are obtained, however, their limit crucially depends on the scaling properties of the random medium.

The most difficult cases concern wave propagation in time-independent random media. The reason is that waves can re-visit the same spatial locations multiple times and thus build statistical correlations (in a sense that has nothing to do with the field-field correlations mentioned earlier) whose understanding generates serious mathematical difficulties. Very few rigorous results exist in this setting. When the correlation length and the wavelength $\lambda \sim l_c$ are comparable, namely in the *weak-coupling* regime, rigorous results of derivation of radiative transfer models were obtained for quantum waves [37, 45, 74] and for discrete classical waves (thus displaying useful dispersive effects) [62]. Such results, which are based on Duhamel expansions of the wave equation, are obtained for random media modeled as Gaussian random field. We briefly mention them in Section 5.1.

When the correlation length is much larger than the wavelength, the radiative transfer model is replaced by another kinetic model, the Fokker-Planck model. The reason is the following. Since the random medium oscillates at a larger scale than the wavelength, we may first replace wave propagation by its classical mechanics approximation, which is a Liouville equation with a random potential. As particles propagate through the random potential, their velocities approximate Brownian motion as an application of a CLT. The resulting evolution equation for the law of the limiting process is thus a Fokker-Planck equation involving a second-order operator in the momentum variable. The derivation of the Fokker-Planck equation is described, albeit not in full mathematical detail, in Section 4. There, the *diffusion approximation* for the Fokker-Planck model obtained for longer times of wave propagation (or equivalently, for more highly disordered random media generating small transport mean free paths) is also presented. The proof of the Fokker-Planck limit involves the Kesten-Papanicolaou cut-off method, that we explain, without going into too many technicalities in Section 4.

Time-dependent random media. The derivation of kinetic models is significantly simplified when the random medium is allowed to vary in the time variable as well. The reason is quite simple: time is visited only once so that when the temporal behavior of the random medium decorrelates sufficiently rapidly, the wave fields and related (quadratic) observables are functionals of the random field for which LLN- and CLT- type results may be applied.

Heuristically, we expect that the radiative transfer models mentioned above will be unchanged when fluctuations in time are slower than fluctuations in space, that is, when $vt_c \ll l_c$. The simplest mathematical descriptions, however, arise when the opposite holds, namely $vt_c \gg l_c$. There, mixing in time is preponderant, which explains why the mathematical arguments simplify. In this highly simplified regime, quantum wave propagation is modeled by an Itô-Schrödinger equation. The statistical moments of its solution satisfy closed-form equations by an application of the Itô calculus. We are therefore in a unique situation where the *ensemble average* of the Wigner transform of wave fields satisfies an exact kinetic equation. This model is presented in Section 5.4.

In the practically often more interesting regimes where (i) $vt_c \sim l_c$ or (ii) $vt_c \gg l_c$, wave propagation is modeled by a Schrödinger equation with a time dependent potential. The two regimes (i) and (ii) are treated in Sections 5.5 and 5.7, respectively. The theory of Section 5.5 is presented in full mathematical detail. To simplify the presentation, we assume that the random medium is Markov in time. The random medium and the Wigner transform are then jointly Markov as well. We analyze the limiting properties of their infinitesimal generator using the perturbed test function technique that is very useful in such problems. The same methodology can then be extended to other regimes of wave propagation. This is done in Section 5.7. In the simplified setting of time-dependent random media, we obtain all standard kinetic models for wave propagation in random media: radiative transfer equations when $\lambda \sim l_c$ and $vt_c \ll l_c$, Fokker Planck equations when $\lambda \ll l_c$ and $vt_c \ll l_c$, and their respective *diffusion approximations* when evolution occurs over longer times or equivalently when the transport mean free path is small.

The results presented in this paper are drawn from works by the authors and collaborators [10, 11]. Similar results under slightly different assumptions on the randomness of the coefficients have also been obtained in, e.g., [39, 70].

Statistical stability and correctors. So far, we have presented several kinetic models without being specific on the sense in which the random Wigner transform converges to its deterministic kinetic limit. The minimum we expect is that the ensemble average of phase space moments of the random Wigner transform W (a quantity of the form $\mathbb{E}\left\{(W(t,x,k),\phi(x,k))_{\mathcal{D}'(\mathbb{R}^{2d}),\mathcal{D}(\mathbb{R}^{2d})}\right\}$) converges to the corresponding moment of the kinetic solution. This is already an involved exercise in the weak-coupling regime [37, 45, 62, 74]. It turns out that in many regimes, we can prove that the whole Wigner transform $W(t,x,k)$ (at least weakly in the phase-space variables) converges in probability to its limit. This is what we referred to earlier as statistical stability.

In many applications we shall discuss in more detail briefly, it is important to understand in which sense convergence occurs and possibly to obtain convergence rates or better yet characterization of correctors. Consider the *scintillation function*

$$J(t,x,k,y,p) = \mathbb{E}\left\{(W(t,x,k) - \mathbb{E}\left\{W(t,x,k)\right\})(W(t,y,p) - \mathbb{E}\left\{W(t,y,p)\right\})\right\}.$$

Its name is drawn from the scintillation of stars in the sky, whose position is somewhat statistically unstable as the (random) atmosphere fluctuates in time. We would like to show that J converges to zero but also possibly obtain a rate of convergence and exhibit its limit after proper rescaling.

It is in the mathematically simpler regime of Itô-Schrödinger propagation that the most complete results are available. There, the scintillation function solves an explicit kinetic equation, again by application of standard Itô calculus. We present error estimates and convergence results for the scintillation function in Section 5.4.2. These results show that scintillation is affected by a much wider array of parameters, including how the wave initial conditions concentrate in phase space and how fast the random media decorrelate in the spatial variables.

Statistical stability results can also be obtained for the derivation of the Fokker Planck regime when $\lambda \ll l_c$ and for time-dependent Schrödinger models when $vt_c \sim l_c$. Such results are presented in sections 4.3 and 5.5.3, respectively. The proof of convergence in the time dependent Schrödinger setting is presented in detail.

Statistical stability is a result on the scintillation function. In cases where the scintillation function can be estimated and a limit obtained, we may expect that the random fluctuations $W - \mathbb{E}\{W\}$ are approximately Gaussian, with the rescaled scintillation function giving the correlation function that characterizes the Gaussian field. That the random corrector is approximately Gaussian is understood in few cases that are presented in Section 5.6.

Kinetic models for field-field correlations. As we mentioned earlier, the Wigner transform is a quantity that is quadratic in the wave field. The results that are presented in Sections 4 and 5 apply to the auto-correlation of the wave field. The Wigner transform may then be identified with the phase-space energy density of the wave fields. The average of the Wigner transform over wavenumbers is precisely the spatial energy density of the propagating fields.

We wish to stress that more general correlations may be considered, and in particular the cross-correlation of two different wave fields with possibly different initial conditions and propagating in possibly different random media. Applications include the analysis of time reversal of waves that we will discuss shortly. How the kinetic models should be generalized to account for such correlations in described in Section 6. There, we revisit the derivation of kinetic models in the Itô-Schrödinger, time-dependent Schrödinger, and Fokker-Planck regimes of wave propagation.

Application to time reversal. Wave equations, whether quantum or classical, admit for solution operators continuous (semi-)groups that are unitary operators of the form e^{itA}. The inverse of such operators, which is thus also their adjoint, is given by e^{-itA} and is simply obtained by reversing the role of time: $t \to -t$. As spatially localized waves spread though a medium, heterogeneous or not, and are measured at a given time $T > 0$, the reconstruction of the initial condition may thus be obtained by re-compression of the measurements at time $T > 0$ by application of e^{-itA}. Applying the operator requires that the field be known everywhere. A striking behavior of time reversal is that in the presence of spatially limited measurements, re-compression is *better* when propagation occurs in a heterogeneous medium than when it occurs in a homogeneous medium.

Mathematically, time reversal is nothing but the composition of two Green's operators (the Schwartz kernels of the unitary operators) plus some filtering processes describing measurements and time reversion. The back-propagated field is therefore a quadratic quantity in wave field, and more precisely a cross-correlation function of two wave fields. This was observed first in [25].

The models presented in the preceding sections are therefore perfectly adapted to the description of the refocusing properties of time reversed waves propagating in highly heterogeneous media. We consider the detailed analysis of time reversal using the kinetic models in section 7.

Applications of wave propagation, imaging, and inverse problems. The derivation of kinetic models to describe wave propagation in random media has a long history; see, e.g., the references [31, 49, 72, 73]. It models certain quantum waves in semi-conductors to light in the atmosphere to some seismic waves in the

Earth crust. Except in one dimension of space, where Anderson localization occurs [30, 41], radiative transfer and other kinetic models are well-accepted to describe wave propagation in such environments. The numerical and experimental validity of radiative transfer equations was also addressed in [6, 13, 17, 61].

Another interesting application of kinetic models pertains to the field of imaging and inverse problems. Consider the problem of the reconstruction of buried inclusions in highly heterogeneous media. The details of the medium can often not be reconstructed from limited, noisy, available measurements. Nor is it always necessary to perform such a detailed reconstruction when kinetic models can be derived. The inclusion may be modeled as a constitutive parameter of the kinetic equation. The inverse wave problem for the reconstruction of the heterogeneous medium and the inclusion is replaced by an inverse kinetic problem for the reconstruction of the statistical properties of the random medium, i.e., essentially the transport mean free path, and of the buried inclusion. Several reconstruction techniques are then available depending on the given measurements [5].

Statistical stability then becomes a necessity. Only stable observables may be used for the solution of the inverse problem as typically measurements for only one realization of the random medium are accessible. What are stable observables and how stable they are was precisely the objective of the results in sections 4 to 6. Moreover, the accuracy in the reconstruction is strongly correlated to the "noise" in the measurements, i.e., to the error between the random observable W_ε and its deterministic limit W solution of the kinetic equation. The scintillation function is therefore an accurate description of the noise correlation function, which governs the statistical instability in the reconstructions [77]. We refer the reader to the recent review [4] on the use of kinetic models to image buried inclusions in random media.

This paper focuses on field-field correlations at two spatial locations for given time. More general correlations in space-time are also useful and can also be modeled by using kinetic models [3]. Such correlations are also extremely useful in imaging in random media when the scattering of the waves off the random scatterers is not modeled but rather treated as noise. Which correlations should be back-propagated, because they are statistically stable, and which should not, because they are not, is the corner stone of the very powerful Coherent Interferometry (CINT) methodology developed, e.g., in [26, 27, 28]. Which correlations to use in inverse kinetic problems has also been investigated in, e.g., [14, 16].

2. **Diffusive limit for a particle in a random flow.**

2.1. **Diffusion of a particle in a time-dependent random flow.**

The simplest non-dissipative problem with weakly random coefficients for which one may establish a long time diffusive behavior is the first order equation

$$\frac{\partial \phi}{\partial t} + \varepsilon v(t,x) \cdot \nabla \phi = 0. \tag{2.1}$$

Here $\varepsilon \ll 1$ is a small parameter, and $v(t,x)$ is a flow that varies randomly both in time and space - a random (vector) field using probabilistic terminology. We will specify the precise assumptions on this random field a little later. Though (2.1) by no means captures the precise behavior of the solutions of, say, the wave equation with a random sound speed, it is sufficiently rich and complex to serve as a good toy model, both to develop appropriate mathematical methods, and exhibit some of the basic phenomena.

2.1.1. *The central limit theorem, purely time-dependent flows, and diffusion.* Since the flow amplitude ε is small, one expects the flow to have a non-trivial effect on long time scales, as "nothing happens" to solutions of (2.1) for $t \sim O(1)$. In order to understand how long one should wait until one may observe some non-trivial behavior, let us recall the central limit theorem. It says that if V_j are independent, identically distributed random variables with mean zero and variance one, that is, $\mathbb{E}(V_j) = 0$, $E(V_jV_k) = \delta_{jk}$, then the variable

$$X_N = \frac{V_1 + \cdots + V_N}{\sqrt{N}}$$

converges as $N \to \infty$ to the standard Gaussian random variable X such that $E(X) = 0$ and $E(X^2) = 1$. One may reformulate this result thinking of V_j as the velocity of a particle on the time-interval $j \le t < j+1$ and as $\varepsilon = 1/\sqrt{N}$ as the velocity amplitude – then the central limit theorem says that the time it takes the particle to behave in a non-trivial way is of the order $t \sim O(\varepsilon^{-2})$. A slight generalization of the usual central limit theorem says, accordingly, that if $V(t)$ is a stationary random process with $E(V(t)) = 0$, and sufficiently rapidly decorrelating in time, then the process

$$X_\varepsilon(t) = \varepsilon \int_0^{t/\varepsilon^2} V(s)ds \tag{2.2}$$

converges, as $\varepsilon \to 0$, to a Brownian motion $B(t)$ with the variance

$$\begin{aligned}
\mathbb{E}(B^2(t)) &= \lim_{\varepsilon\to 0} \mathbb{E}(X_\varepsilon^2(t)) \\
&= \lim_{\varepsilon\to 0} \varepsilon^2 \int_0^{t/\varepsilon^2}\int_0^{t/\varepsilon^2} \mathbb{E}[V(s)V(s')]dsds' \\
&= 2\lim_{\varepsilon\to 0} \varepsilon^2 \int_0^{t/\varepsilon^2}\int_s^{t/\varepsilon^2} R(s'-s)dsds' \\
&= 2\lim_{\varepsilon\to 0} \varepsilon^2 \int_0^{t/\varepsilon^2}\int_0^{t/\varepsilon^2-s} R(s')dsds' \\
&= 2\lim_{\varepsilon\to 0} \int_0^{t/\varepsilon^2} (t-\varepsilon^2 s')R(s')ds = t\int_{-\infty}^{\infty} R(s)ds.
\end{aligned}$$

Here $R(s) = \mathbb{E}(V(t)V(t+s))$ is the covariance function of the process $V(t)$. A precisely formulation of a result of this type can be found in [48], see Theorem 18.7.1.

One may rephrase this result in terms of a PDE as follows. Let $\phi(t)$ be the solution of the initial value problem

$$\frac{\partial\phi}{\partial t} + \varepsilon V(t)\frac{\partial\phi}{\partial x} = 0,\ \phi(0,x) = \phi_0(x), \tag{2.3}$$

and set $\phi_\varepsilon(t) = \phi(t/\varepsilon^2, x)$. Then $\mathbb{E}(\phi_\varepsilon(t,x)) \to \bar\phi(t,x)$, with the function $\bar\phi(t,x)$ that solves the diffusion equation

$$\frac{\partial\bar\phi}{\partial t} = D\frac{\partial^2\bar\phi}{\partial x^2},\ \bar\phi(0,x) = \phi_0(x), \tag{2.4}$$

with the diffusion constant

$$D = \frac{1}{2}\int_{-\infty}^{\infty} R(s)ds. \tag{2.5}$$

2.1.2. *Using formal asymptotic expansions.* One may wonder whether the limit equation (2.4) could be obtained without resorting to the central limit theorem. Let us show a simple formal derivation, taking spatial dimension $d = 1$ for simplicity, as we have done so far above. The function $\phi_\varepsilon(t,x)$ satisfies

$$\frac{\partial \phi}{\partial t} + \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}\right)\frac{\partial \phi}{\partial x} = 0,\ \ \phi_\varepsilon(0,x) = \phi_0(x). \tag{2.6}$$

Consider a formal asymptotic multiple scales expansion

$$\phi_\varepsilon(t,x) = \psi(t,x) + \varepsilon\psi_1(t,t/\varepsilon^2,x) + \varepsilon^2\psi_2(t,t/\varepsilon^2,x) + \dots$$

We assume here that the leading order term $\psi(t,x)$ is deterministic and does not depend on the fast time scale $\tau = t/\varepsilon^2$. Inserting this ansatz for ϕ_ε into (2.6) gives, and collecting terms of the order $O(\varepsilon^{-1})$

$$\frac{\partial \psi_1}{\partial \tau} = -V(\tau)\frac{\partial \psi(t,x)}{\partial x},$$

or

$$\psi_1(t,\tau,x) = -\frac{\partial \psi(t,x)}{\partial x}\int_{-\infty}^{\tau} V(s)ds. \tag{2.7}$$

As the integral in (2.7) has no reason to converge, we introduce a regularization parameter $\theta \ll 1$, that we will later send to zero:

$$\psi_1(t,\tau,x) = -\frac{\partial \psi(t,x)}{\partial x}\int_{-\infty}^{\tau} e^{\theta s}V(s)ds. \tag{2.8}$$

Terms of the order $O(1)$ in (2.6) combine to give

$$\frac{\partial \psi}{\partial t} + \frac{\partial \psi_2}{\partial \tau} + V(\tau)\frac{\partial \psi_1}{\partial x} = 0. \tag{2.9}$$

Assuming that ψ_2 is stationary in τ leads to $\mathbb{E}(\partial\psi_2/\partial\tau) = 0$. Therefore, averaging (2.9) leads to (recall that $\psi(t,x)$ is assumed to be deterministic!)

$$\frac{\partial \psi}{\partial t} = -\mathbb{E}\left[V(\tau)\frac{\partial \psi_1}{\partial x}\right]. \tag{2.10}$$

Let us now compute the term in the right side above, using (2.8):

$$\begin{aligned}
-\mathbb{E}\left[V(\tau)\frac{\partial \psi_1}{\partial x}\right] &= \int_{-\infty}^{\tau} e^{\theta s}\mathbb{E}[V(\tau)V(s)]\frac{\partial^2 \psi(t,x)}{\partial x^2}ds \\
&= \int_{-\infty}^{\tau} e^{\theta s}R(s-\tau)\frac{\partial^2 \psi(t,x)}{\partial x^2}ds \\
&\to \int_{-\infty}^{0} R(s)ds\frac{\partial^2 \psi(t,x)}{\partial x^2} = D\frac{\partial^2 \psi(t,x)}{\partial x^2},
\end{aligned}$$

as $\theta \to 0$, with the diffusion constant as in (2.5). We recover the diffusion equation (2.4) that we have obtained previously using the central limit theorem:

$$\frac{\partial \psi}{\partial t} = D\frac{\partial^2 \psi}{\partial x^2},\ \ \psi(0,x) = \psi_0(x). \tag{2.11}$$

Unfortunately, as far as we know, this formal derivation can not be made rigorous in any straightforward way, and an explanation of "the reason" it works would sidetrack us for too long if we attempt it right now. However, as we explain later, this provides a very effective tool to find the (usually) correct limit equation in a large class of weakly random problems, where rigorous proofs require quite sophisticated probabilistic techniques.

2.1.3. *Random flows with spatial-temporal dependence.* Let us now turn to a more complex situation when the random flow in (2.1) depends both on time and space:

$$\frac{\partial \phi}{\partial t} + \varepsilon V(t,x)\cdot\nabla\phi = 0,\ \phi(0,x) = \phi_0(x), \tag{2.12}$$

and now we will consider the general situation $d \geq 1$. Motivated by the previous discussion, we should expect non-trivial behavior on a time scale $t \sim O(\varepsilon^{-2})$, and we rescale the time accordingly $t \to \varepsilon^2 t$:

$$\frac{\partial \phi}{\partial t} + \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, x\right)\cdot\nabla\phi = 0,\ \phi(0,x) = \phi_0(x), \tag{2.13}$$

The probabilistic approach to this problem relies on understanding the behavior of characteristics $X_\varepsilon(t) = (X_{1,\varepsilon}(t),\ldots,X_{d,\varepsilon}(t))$ when $\varepsilon \ll 1$. Equivalently one may wish to describe the asymptotics of solutions to ordinary differential equation

$$\dot X_\varepsilon(t) = \frac{1}{\varepsilon} V\left(\frac{t}{\varepsilon^2}, X_\varepsilon(t)\right),\ X_\varepsilon(0) = x, \tag{2.14}$$

as $\varepsilon \to 0$. This question goes back to the papers by Khasminskii [54] from the 60's with subsequent contributions by various authors: without any attempt at completeness we mention the work of Borodin [29], Papanicolaou and Kohler [67], and Kesten and Papanicolaou [52]. We present below a version of the limit theorem due to Komorowski [55].

When does one expect the trajectories of (2.14) to behave diffusively? If the flow V is spatially uniform then, as we have seen above, one needs V to have mean zero and decorrelate rapidly in time: the covariance function $R(s)$ should be "sufficiently" rapidly decaying (though we have not made this requirement overly precise above, it is clear that at least one needs the diffusion coefficient D to be finite, requiring integrability of $R(s)$). Hence, first of all, $V(t,x)$ has to have mean zero so that the mean displacement would not be clearly biased, and we will need some decay assumptions on the covariance matrix

$$R_{mn}(t,x) = \mathbb{E}[V_m(s,y)V_n(s+t,y+x)],\ 1 \leq m,n \leq d. \tag{2.15}$$

This is formalized by the mixing assumption below that eliminates the memory effect. Second, V should "mix things around" which means that the flow should be incompressible – this will prevent existence of spatial traps. Finally, statistically there should be no distinguished times and spatial positions – this requires stationarity of V in time and space (in particular this implies that the covariance matrix in (2.15) does not depend on (s,y)). These assumptions should, in principle, ensure a diffusive limit – after all, the Lagrangian velocity field "observed" by a particle moving along the characteristics (2.14) will likely be statistically indistinguishable from a stationary time-dependent but spatially uniform mixing field observed by a particle moving in characteristics of (2.6). Physically, there is little reason to expect a major difference. However, mathematically the problem is much more complicated.

A formal asymptotic limit. Let us first show that from the point of view of formal asymptotics there is a little difference between the spatially uniform and spatially random fields (provided that they are "sufficiently random" in time!). Both the derivation and the answer are essentially the same. We consider an asymptotic expansion

$$\phi(t,x) = \psi(t,x) + \varepsilon\psi_1(t,\frac{t}{\varepsilon^2},x) + \varepsilon^2\psi_2(t,\frac{t}{\varepsilon^2},x) + \ldots \tag{2.16}$$

for the solution of (2.13). Inserting it into (2.13) gives in the order $O(\varepsilon^{-1})$:

$$\frac{\partial \psi_1}{\partial \tau} = -V(\tau, x) \cdot \nabla \psi(t, x),$$

and, as before, we introduce a regularization parameter $\theta \ll 1$, that we will later send to zero:

$$\psi_1(t, \tau, x) = -\int_{-\infty}^{\tau} e^{\theta s} V(s, x) \cdot \nabla \psi(t, x) ds. \tag{2.17}$$

Terms of the order $O(1)$ in (2.13) combine to give

$$\frac{\partial \psi}{\partial t} + \frac{\partial \psi_2}{\partial \tau} + V(\tau, x) \cdot \nabla \psi_1 = 0. \tag{2.18}$$

Also as before, assuming that ψ_2 is stationary in τ, so that, $\mathbb{E}(\partial \psi_2/\partial \tau) = 0$, and averaging (2.18) leads to

$$\frac{\partial \psi}{\partial t} = -\mathbb{E}\left[V(\tau, x) \cdot \nabla \psi_1\right]. \tag{2.19}$$

The right side above,is, once again:

$$\begin{aligned}
-\mathbb{E}\left[V(\tau, x) \cdot \nabla \psi_1\right] &= \sum_{m,n=1}^{d} \int_{-\infty}^{\tau} e^{\theta s} \mathbb{E}[V_m(\tau, x) V_n(s, x)] \frac{\partial^2 \psi(t, x)}{\partial x_m \partial x_n} ds \\
&= \sum_{m,n=1}^{d} \int_{-\infty}^{\tau} e^{\theta s} R_{mn}(s - \tau, 0) \frac{\partial^2 \psi(t, x)}{\partial x_m \partial x_n} ds \\
&\to \sum_{m,n=1}^{d} {}_{mn} \frac{\partial^2 \psi(t, x)}{\partial x_m \partial x_n},
\end{aligned}$$

as $\theta \to 0$, with the diffusion matrix

$$D_{mn} = \frac{1}{2} \int_{-\infty}^{\infty} R_{mn}(s, 0) ds. \tag{2.20}$$

We recover the diffusion equation

$$\frac{\partial \psi}{\partial t} = \sum_{m,n=1}^{d} D_{mn} \frac{\partial^2 \psi}{\partial x_m \partial x_n}, \;\; \psi(0, x) = \psi_0(x). \tag{2.21}$$

It is quite remarkable that this very simple formal derivation of the diffusive limit is quite difficult to justify by purely PDE methods, and, to the best of our knowledge, there are no PDE methods available when randomness is time-independent.

Assumptions on the random field. Generally, we will try to limit lengthy rigorous proofs in this review article. However, since the turbulent diffusion problem is the simplest setting when such proof is non-trivial we present it now. We begin with assumptions on the random field, with some explanations as to why they are needed. The need for such hypotheses is quite common in many problems where averaging of small scales is possible.

Stationarity. The random field $V(t, x)$ is strictly stationary in time and space. This means that for any $t_1, t_2, \ldots, t_m \in \mathbb{R}$, $x_1, \ldots, x_m \in \mathbb{R}^d$, and each $h \in \mathbb{R}$ and $y \in \mathbb{R}^d$ the joint distribution of the random vectors $V(t_1 + h, x_1 + y), V(t_2 + h, x_2 + y), \ldots, V(t_m + h, x_m + y)$ is the same as that of $V(t_1, x_1), V(t_2, x_2), \ldots, V(t_m, x_m)$. We will denote by $R_{nm}(t, x)$ the covariance matrix of the field $V(t, x)$ defined in (2.15). One may think of the stationarity assumption as the analog of periodicity in standard homogenization in periodic media – some assumption of this sort is needed to make the effect of the small scale media fluctuations "identical in the statistical sense" in various regions of space and at different time.

The spatial stationarity of the field is not really necessary here but it allows to simplify a few expressions in what follows. This can be seen already from the formal computation done above. We adopt it here simply for convenience. On the other hand, stationarity in time is essential for the limit theorem. In other problems, in particular, there randomness is time-independent, spatial statistical homogeneity is essential.

Mixing. Given $C,\rho>0$ let us denote by $\mathcal{V}_a^b(C,\rho)$ the σ-algebra generated by the events of the form $\{\omega:\ V(t,x,\omega)\in A\}$, where $a\le t\le b$, $|x|\le C(1+t^\rho)$ and A is a Borel set in $\mathbb{R}^d$. When $C=\infty$ in the above definition (ρ is of no importance anymore) we obtain that the respective σ-algebra $\mathcal{V}_a^b(\infty)$ is generated by the sets of the form $\{\omega:\ V(t,x,\omega)\in A\}$ where $a\le t\le b$, $x\in\mathbb{R}^n$ (there is no restriction on x now) and A is a Borel set in $\mathbb{R}^n$.

Define the (uniform) mixing coefficient of the field by

$$\phi_{C,\rho}(h)=\sup_t\sup_{A\in\mathcal{V}_{t+h}^\infty(C,\rho),B\in\mathcal{V}_0^t(C,\rho)}\frac{|P(A\cap B)-P(A)P(B)|}{P(B)} \tag{2.22}$$

Our principal assumption is that there exists $C_V>0$ and $1/2<\rho_V$ such that

$$C_{mix}:=\sup_{h>0}h^{100}\phi_{C_V,\rho_V}(h)<+\infty. \tag{2.23}$$

The mixing coefficient corresponding to $C_V=\infty$ shall be denoted by $\phi_\infty(h)$. It shall be customary in our notation of the mixing coefficient and the respective σ-algebras to omit parameters C_V,ρ_V if their value is obvious from the context.

One may wonder why the definition of the mixing coefficient involves parameters ρ,C. After all we could have formulated the mixing assumption, somewhat more customarily, using only $\phi_\infty(h)$. However this hypothesis 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 a bit more complicated definition of the mixing condition as in (2.22).

While the mixing assumption may be taken in various forms, weakened or strengthened, depending on a particular problem, from the physical point of view a mixing assumption in any form ensures (in a non-trivial way) that the particle (or a wave in a wave problem) experiences different and "nearly independent" randomness in various regions of space, allowing to obtain some form of a central limit theorem type result (which may be well hidden and obscured by the technique of a particular proof) that eventually leads to a diffusive or kinetic limit.

Boundedness. The random field $V(t,x)$ has two spatial derivatives and

$$\text{esssup}\left[\|V\|_\infty+\sum_{j=1}^d\left\|\frac{\partial V}{\partial x_j}\right\|_\infty+\sum_{i,j=1}^d\left\|\frac{\partial^2 V}{\partial x_i\partial x_j}\right\|_\infty\right]<+\infty$$

for all $1\le i,j,l\le n$. Here $\|\cdot\|_\infty$ is the supremum norm over $\mathbb{R}^d$ and essup corresponds to the essential supremum over the random variable. This assumption is purely technical and can often be weakened.

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

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

The limit theorem. Let us define the diffusion matrix

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

and its (unique) symmetric, non-negative definite, square-root matrix σ, i.e. $\sigma \geq 0$, $\sigma^T = \sigma$ and $\sigma^2 = a$.

Suppose that $\{Y_\varepsilon(t),\, t \geq 0\}$ is a family of continuous trajectory stochastic processes. We say that they converge weakly, as $\varepsilon \to 0$, to a process $\{Y(t),\, t \geq 0\}$ if for any bounded and continuous functional $F : C[0,+\infty) \to \mathbb{R}$ we have $\lim_{\varepsilon\to 0} \mathbb{E}F(Y_\varepsilon(\cdot)) = \mathbb{E}F(Y(\cdot))$. The following theorem holds.

Theorem 2.1. *Suppose that the random field $V(t,x)$ satisfies the assumptions made above. Then the processes $X_\varepsilon(t)$, solution of (2.14), converge weakly, as $\varepsilon \to 0$, to the limit process $\bar{X}(t) = x + \sigma W_t$. Here W_t is the standard Brownian motion.*

The main result of [55] 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. [55] also accounts for the possible small scale variations of the random field looking at equations of the form

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

with $0 \leq \alpha < 1$. We will not describe this generalization 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 2.1 under a simplifying assumption that the matrix σ is invertible. While this does not subtract any of the essential aspects of the proof, it does shorten many expressions and calculations which are sufficiently long even without them. The proof proceeds in several (typical for such limit theorems) 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 $X_\varepsilon(t)$. In the last step, we identify the limit of the processes $X_\varepsilon(t)$ as a Brownian motion multiplied by the matrix σ by means of the martingale characterization of the Brownian motion.

2.2. **The proof of Theorem 2.1.**

The mixing lemmas. We start with the proof of the tightness of the family of probability measures generated by the process $X_\varepsilon(t)$. 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. In our case it is Lemma 2.2 below. Ultimately this allows us to split expectations into product of expectations and either "justify", or explain away the closure assumptions that are often made formally. In our particular problem it explains why the formal assumption that the leading order term in the asymptotic expansion (2.16) is deterministic produced the correct answer.

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

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

Incompressibility of $V(t,x)$ and its spatial stationarity imply that $\mathbb{E}\{G_1(s_1,s_2,x)\}=0$.

Lemma 2.2. *Suppose that mixing condition* (2.23) *holds. Then for any $T\geq 0$ there exists a constant $C>0$ such that for any $0\leq u\leq s\leq s_2\leq s_1\leq T$, $\varepsilon\in(0,1]$ and Y, that is a $\mathcal{V}_0^{s/\varepsilon^2}$-measurable random vector, we have*

$$\left|\mathbb{E}\left\{V\left(\frac{s_1}{\varepsilon^2},X_\varepsilon(u)\right)\cdot Y\right\}\right|\leq C\phi\left(\frac{s_1-s}{\varepsilon^2}\right)\mathbb{E}\,|Y|, \tag{2.24}$$

$$\left|\mathbb{E}\left\{\partial_{x_k}V\left(\frac{s_1}{\varepsilon^2},X_\varepsilon(u)\right)\cdot Y\right\}\right|\leq C\phi\left(\frac{s_1-s}{\varepsilon^2}\right)\mathbb{E}\,|Y| \tag{2.25}$$

and

$$\left|\mathbb{E}\left\{G_1\left(\frac{s_1}{\varepsilon^2},\frac{s_2}{\varepsilon^2},X_\varepsilon(u)\right)\cdot Y\right\}\right|\leq C\phi^{1/2}\left(\frac{s_1-s_2}{\varepsilon^2}\right)\phi^{1/2}\left(\frac{s_2-s}{\varepsilon^2}\right)\mathbb{E}\,|Y| \tag{2.26}$$

$$\left|\mathbb{E}\left\{\partial_{x_k}G_1\left(\frac{s_1}{\varepsilon^2},\frac{s_2}{\varepsilon^2},X_\varepsilon(u)\right)\cdot Y\right\}\right|\leq C\phi^{1/2}\left(\frac{s_1-s_2}{\varepsilon^2}\right)\phi^{1/2}\left(\frac{s_2-s}{\varepsilon^2}\right)\mathbb{E}\,|Y| \tag{2.27}$$

for all $1\leq k\leq n$. Here ϕ is the mixing coefficient defined in (2.22).

Proof. First of all, we note that for $C_*:=1+\sup|V(t,x)|$ one can find $\varepsilon_0>0$, depending only on C_V, C_* and T, such that for any $\varepsilon\in(0,\varepsilon_0]$ the process $\{X_\varepsilon(t),\ t\in[0,u]\}$ does not leave the ball of the radius $C_V[1+(u/\varepsilon^2)^{\rho_V}]$, centered at the origin, hence it is $\mathcal{V}_0^{u/\varepsilon^2}$-measurable. Indeed,

$$|X_\varepsilon(t)|\leq\frac{1}{\varepsilon}\int_0^t\left|V\left(\frac{s}{\varepsilon^2},X_\varepsilon(s)\right)\right|ds\leq\frac{C_*u}{\varepsilon}\leq C_V\left(1+\frac{u^{\rho_V}}{\varepsilon^{2\rho_V}}\right),\ t\in[0,u],$$

provided $\varepsilon\in(0,\varepsilon_0]$ and $\varepsilon_0:=(C_V/C_*T^{1-\rho_V})^{1/(2\rho_V-1)}$.

The conclusion of the lemma is obviously true for $\varepsilon\in(\varepsilon_0,1]$ upon a suitable choice of constant C so we only consider the case when $\varepsilon\in(0,\varepsilon_0]$.

We will first prove (2.25), and the reader may check that the proof of (2.24) is identical. The recurring idea in such proofs is to replace the random variable $X_\varepsilon(u)$ by a deterministic value and use the mixing properties of the field $V(t,x)$ in time. Let $M\in\mathbb{N}$ be a fixed positive integer and $l\in\mathbb{Z}^n$. Define the event

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

The event $A(l)$ is $\mathcal{V}_0^{s/\varepsilon^2}$ measurable since $u \le s$. For almost every realization ω there exists exactly one $l \in \mathbb{Z}^n$ so that $\omega \in A(l)$. Then we may decompose the expectation in (2.25) using the fact that the random variable $X_\varepsilon(u)$ is close to the non-random value l/M on the event $A(l)$ as follows:

$$\begin{aligned}\left|\mathbb{E}\left\{\partial_{x_k} V\left(\tfrac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right)\cdot Y\right\}\right| &= \left|\sum_l \mathbb{E}\left\{\partial_{x_k} V\left(\tfrac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right)\cdot Y,\, A(l)\right\}\right| \\ &\le \left|\sum_l \mathbb{E}\left\{\left[\partial_{x_k} V\left(\tfrac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right) - \partial_{x_k} V\left(\tfrac{s_1}{\varepsilon^2}, \tfrac{l}{M}\right)\right] Y, A(l)\right\}\right| \\ &\quad + \left|\sum_l \mathbb{E}\left\{\partial_{x_k} V\left(\tfrac{s_1}{\varepsilon^2}, \tfrac{l}{M}\right)\cdot Y, A(l)\right\}\right| \\ &= I + II.\end{aligned}$$

The second term above may be now estimated using (2.23) and the fact that $\mathbb{E}\left\{\partial_{x_k} V\left(s_1/\varepsilon^2, l/M\right)\right\} = 0$ by

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

uniformly in M.

Since we have assumed that two spatial derivatives of the field $V(t,x)$ are bounded by a deterministic constant, $\partial V/\partial x_k$ is uniformly continuous in space. Therefore, using the Lebesgue dominated convergence theorem we conclude that $I \to 0$ as $M \to +\infty$ and (2.25) follows. An identical proof shows that in addition we have the same bound for the second derivatives of the random field V:

$$\left|\mathbb{E}\left\{\partial^2_{x_k,x_m} V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right)\cdot Y\right\}\right| \le C\phi\left(\frac{s_1 - s}{\varepsilon^2}\right)\mathbb{E}\left|Y\right|. \tag{2.28}$$

We now prove (2.27) – the proof of (2.26) is identical. Let us first write out the expression for G_1:

$$\begin{aligned}&\left|\mathbb{E}\left\{\partial_{x_k} G_1\left(\tfrac{s_1}{\varepsilon^2}, \tfrac{s_2}{\varepsilon^2}, X_\varepsilon(u)\right)\cdot Y\right\}\right| \\ &\le \sum_{p=1}^n \left|\mathbb{E}\left\{\partial_{x_k} V_p\left(\frac{s_2}{\varepsilon^2}, X_\varepsilon(u)\right)\partial_{x_p} V\left(\frac{s_1}{\varepsilon^2}, X_\varepsilon(u)\right)\cdot Y\right\}\right|.\end{aligned}$$

Now we may apply (2.25), (2.28) in two different ways using different parts of the inequality $s \le s_2 \le s_1$. First, we may use (2.25), (2.28) with the gap between s_1 and s_2, that is, we group into "Y" in (2.25), (2.28) all terms that involve s and s_2. Using in addition the uniform bounds on V and its derivatives this leads to

$$\left|\mathbb{E}\left\{\partial_{x_k} G_1\left(\frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u)\right)\cdot Y\right\}\right| \le C\phi\left(\frac{s_1 - s_2}{\varepsilon^2}\right)\mathbb{E}\left|Y\right|.$$

Second, note that (2.25) may be slightly generalized to apply with $\partial V/\partial x_k$ replaced by a sufficiently smooth in space $\mathcal{V}^\infty_{s_2/\varepsilon^2}$ random variable with an expectation equal to zero. Since

$$\mathbb{E}\{G_1\left(s_1/\varepsilon^2, s_2/\varepsilon^2, x\right)\} = 0$$

we can use use this modified version of (2.25) with the gap between s_2 and s, taking "Y" in (2.25) to be simply Y:

$$\left|\mathbb{E}\left\{\partial_{x_k} G_1\left(\frac{s_1}{\varepsilon^2}, \frac{s_2}{\varepsilon^2}, X_\varepsilon(u)\right)\cdot Y\right\}\right| \le C\phi\left(\frac{s_2 - s}{\varepsilon^2}\right)\mathbb{E}\left|Y\right|.$$

Multiplying these two inequalities and taking the square root we conclude that (2.27) holds. This finishes the proof of Lemma 2.2. □

The proof of tightness. Now, we are ready to prove tightness of the family of the laws of the processes $\{X_\varepsilon(t),\, t\ge 0\}$. For that purpose we shall establish that for any $T>0$ there exists constants $C,\nu>0$ such that

$$\mathbb{E}\left\{|X_\varepsilon(t)-X_\varepsilon(s)|^2|X_\varepsilon(s)-X_\varepsilon(u)|^2\right\}\le C(t-u)^{1+\nu},\ 0\le u\le s\le t\le T. \quad (2.29)$$

This implies tightness in the space $C[0,+\infty)$ for the family of laws of continuous trajectory processes, see Chapter 3 of [23]. It can be seen as follows. Thanks to estimate (15.22) p. 129 of ibid., (2.29) implies that for any $T,\eta>0$

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\mathbb{P}\left[w''_T(X_\varepsilon;\delta)>\eta\right]=0.$$

Here for any function $f:[0,T]\to\mathbb{R}$ and $\delta>0$ we define

$$w''_T(f;\delta):=\sup\min\{|f(t_1)-f(t)|,|f(t_2)-f(t)|:0<t_2-t_1<\delta, 0\le t_1\le t\le t_2\le T\}.$$

It can be easily seen that for a continuous function f we have $1/2w_T(f;\delta)\le w''_T(f;\delta)\le w_T(f;2\delta)$, where $w_T(f;\delta):=\sup\{|f(t_2)-f(t_1)|:0<t_2-t_1<\delta,0\le t_1\le t_2\le T\}$. This in turn yields that

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\mathbb{P}\left[w_T(X_\varepsilon;\delta)>\eta\right]=0.$$

Since we have assumed that $X_\varepsilon(0)=0$, tightness in $C[0,+\infty)$ is a consequence of Theorem 8.2, p. 55 of [23].

The main step in the proof is to find $C>0$ and $\gamma\in(1,2)$ such that for all times $t,s\in[0,T]$ such that $t-s>10\varepsilon^\gamma$ we have an estimate for the conditional expectation

$$\mathbb{E}\left\{|X_\varepsilon(t)-X_\varepsilon(s)|^2\,\Big|\,\mathcal{V}_0^{s/\varepsilon^2}\right\}\le C(t-s)\ \text{for}\ t-s>10\varepsilon^\gamma. \quad (2.30)$$

The gap between the times t and s is needed to make use of the mixing lemma.

Step 1. Nearby times. Estimate (2.30) itself is sufficient to establish tightness in $D[0,+\infty)$ for the family $X_\varepsilon(t)$ if it were to hold for all $t>s$, as it clearly implies (2.29). Indeed, we would get trivially

$$\mathbb{E}|X_\varepsilon(s)-X_\varepsilon(u)|^2\le C(s-u) \quad (2.31)$$

and the left hand side of (2.29) could be estimated by

$$\begin{aligned}&\mathbb{E}\left\{\mathbb{E}\left\{|X_\varepsilon(t)-X_\varepsilon(s)|^2\mid\mathcal{V}_0^{s/\varepsilon^2}\right\}|X_\varepsilon(s)-X_\varepsilon(u)|^2\right\}\\ &\le C(t-s)\mathbb{E}|X_\varepsilon(s)-X_\varepsilon(u)|^2\\ &\le C^2(t-s)(s-u)\\ &\le C^2(t-u)^2.\end{aligned}$$

Thus, (2.29) would follow with $\nu=1$.

Since (2.30) will be shown only for pairs of time with a gap: $t-s>10\varepsilon^\gamma$, we may at the moment conclude only that

$$\mathbb{E}\left\{|X_\varepsilon(t)-X_\varepsilon(s)|^2|X_\varepsilon(s)-X_\varepsilon(u)|^2\right\}\le C(t-u)^2\ \text{for}\ t-s>10\varepsilon^\gamma\ \text{and}\ s-u>10\varepsilon^\gamma.$$

Adjusting $\gamma \in (1,2)$ we prove that (2.29) still holds for some ν, if either $t-s \le 10\varepsilon^\gamma$ or $s-u \le 10\varepsilon^\gamma$ Suppose that both $t-s \le 10\varepsilon^\gamma$ and $s-u \le 10\varepsilon^\gamma$. Then directly from (2.14) we have:

$$\begin{aligned}
&\mathbb{E}\left\{|X_\varepsilon(t)-X_\varepsilon(s)|^2|X_\varepsilon(s)-X_\varepsilon(u)|^2\right\}\\
&\le C\varepsilon^{-4}(t-s)^2(s-u)^2\\
&\le C\varepsilon^{11\gamma/4-4}(t-u)^{5/4}\\
&\le C(t-u)^{5/4}
\end{aligned}$$

provided that $\gamma > 16/11$. On the other hand, if, say, $t-s \le 10\varepsilon^\gamma$ but $s-u > 10\varepsilon^\gamma$, then (2.31) holds. On the other hand (2.14) implies that with probability one

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

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

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

provided that $\gamma > 8/7$. We see that, indeed, (2.30) together with (2.14) are sufficient to prove the tightness criterion (2.29). The rest of the proof of tightness of the processes $X_\varepsilon(t)$ is concerned with verifying (2.30).

Step 2. Taking a time-step backward. Suppose we are given 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

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

where $[x]$ is the integer part of x. Then the time step Δt is such that $\varepsilon^\gamma/2 \le \Delta t \le 2\varepsilon^\gamma$ and the partition $s = t_0 < t_1 < \cdots < t_{M+1} = t$ is taken with a time step Δt. Here $M = [\varepsilon^{-\gamma}(t-s)]$. 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 (2.30) is “to use two term Taylor expansion for $X_\varepsilon(t)-X_\varepsilon(s)$ ” for a time step of $O(\Delta t)$ size with explicitly computable terms. The corresponding error terms, which are nominally large, are shown to be negligible using mixing Lemma 2.2.

Dropping the subscript ε of $X_\varepsilon(t)$ we write:

$$X(t)-X(s) = \frac{1}{\varepsilon}\int_s^t V\left(\frac{u}{\varepsilon^2},X(u)\right)du = \frac{1}{\varepsilon}\sum_{i=0}^{M}\int_{t_i}^{t_{i+1}} V\left(\frac{u}{\varepsilon^2},X(u)\right)du. \tag{2.32}$$

Therefore our task is to estimate the integral inside the summation in the right side of (2.32). In the preparation for the application of the mixing lemma the integrand on the interval $t_i \le u \le t_{i+1}$ can be rewritten as

$$\begin{aligned} V\left(\tfrac{u}{\varepsilon^2}, X(u)\right) &= V\left(\tfrac{u}{\varepsilon^2}, X(t_{i-1})\right) + \int_{t_{i-1}}^{u} \tfrac{d}{du_1} V\left(\tfrac{u}{\varepsilon^2}, X(u_1)\right) du \\ &= V\left(\tfrac{u}{\varepsilon^2}, X(t_{i-1})\right) + \int_{t_{i-1}}^{u} \sum_{p=1}^{n} \partial_{x_p} V\left(\tfrac{u}{\varepsilon^2}, X(u_1)\right)\left(\tfrac{1}{\varepsilon} V_p\left(\tfrac{u_1}{\varepsilon^2}, X(u_1)\right)\right) du_1 \\ &= V\left(\tfrac{u}{\varepsilon^2}, X(t_{i-1})\right) + \tfrac{1}{\varepsilon}\int_{t_{i-1}}^{u} G_1\left(\tfrac{u}{\varepsilon^2}, \tfrac{u_1}{\varepsilon^2}, X(u_1)\right) du_1. \end{aligned}$$

The next step is to expand G_1 as well, also around the "one-step-backward" time t_{i-1}:

$$G_1\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(u_1)\right) = G_1\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, X(t_{i-1})\right) + \frac{1}{\varepsilon}\int_{t_{i-1}}^{u_1} G_2\left(\frac{u}{\varepsilon^2}, \frac{u_1}{\varepsilon^2}, \frac{u_2}{\varepsilon^2}, X(u_2)\right) du_2$$

with

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

Putting together the above calculations we see that

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

The triple integral in the last line is bounded by $C\varepsilon^{2\gamma-3}(t-s)$ for some deterministic constant $C > 0$, so it becomes small for $\gamma > 3/2$. Indeed, the time interval in each integration is smaller than ε^γ and the total number of terms is at most $M + 1 \approx (t-s)/\varepsilon^\gamma$. Since G_2 is deterministically bounded we obtain an estimate of the term in question by $C\varepsilon^{-3}(M+1)(\Delta t)^3 \approx C\varepsilon^{2\gamma-3}(t-s)$.

In fact 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 $|E(s,t)| \le C\varepsilon^p(t-s)$ with some deterministic constants $p, C > 0$. This finishes the first preliminary step in the proof of tightness.

Step 3. Application of the mixing lemma. Now we are ready to prove (2.30). 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 $0 \le s \le t \le T$ such that $t \ge s + 10\varepsilon^\gamma$:

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

Our estimates in Step 2 show that it is actually enough to verify that

$$\mathbb{E}\left\{L_m^2(s,t)Y\right\} \le C(t-s)\mathbb{E}Y, \ m = 1, 2. \tag{2.33}$$

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

$$\begin{aligned}
&\mathbb{E}\left\{L_1^2(s,t)Y\right\} \\
&= \frac{2}{\varepsilon^2}\sum_{i<j}\sum_{p=1}^{n}\int_{t_i}^{t_{i+1}}\int_{t_j}^{t_{j+1}}\mathbb{E}\left\{V_p\left(\tfrac{u}{\varepsilon^2}, X(t_{i-1})\right)V_p\left(\tfrac{u'}{\varepsilon^2}, X(t_{j-1})\right)Y\right\}dudu' \\
&\quad + \frac{1}{\varepsilon^2}\sum_{j}\sum_{p=1}^{n}\int_{t_j}^{t_{j+1}}\int_{t_j}^{t_{j+1}}\mathbb{E}\left\{V_p\left(\tfrac{u}{\varepsilon^2}, X(t_{j-1})\right)V_p\left(\tfrac{u'}{\varepsilon^2}, X(t_{j-1})\right)Y\right\}dudu' \\
&= \mathcal{I}_1 + \mathcal{I}_2.
\end{aligned}$$

The terms $\mathcal{I}_1$ and $\mathcal{I}_2$ correspond to the summation ranges $i \le j-1$ and $i = j$. We estimate them separately as they end up being of a different order. The idea is to use the separation between t_i and t_j and apply Lemma 2.2. In case of $\mathcal{I}_1$ it implies that, as the time gap between the times u' and u is in general larger than the correlation time ε^2, we have

$$|\mathcal{I}_1| \le o(1)(t-s)\mathbb{E}Y, \quad \text{as } \varepsilon \to 0. \tag{2.34}$$

Indeed, from Lemma 2.2 we get

$$|\mathcal{I}_1| \le \frac{C}{\varepsilon^2}\sum_{j=0}^{M}\sum_{i\le j-1}\int_{t_i}^{t_{i+1}}\int_{t_j}^{t_{j+1}}\phi\left(\frac{u'-u}{\varepsilon^2}\right)\mathbb{E}Y\,dudu'. \tag{2.35}$$

Since $u' - u \ge \varepsilon^{\gamma-2}$ the sum over indices $i \le j-2$ can be estimated further by

$$\frac{C}{\varepsilon^2}(M\Delta t)^2\phi\left(\varepsilon^{\gamma-2}\right)\mathbb{E}Y \le \frac{C}{\varepsilon^2}\phi\left(\varepsilon^{\gamma-2}\right)(t-s)^2\mathbb{E}Y.$$

We select γ from $(16/11, 2)$ in such a way that (2.23) implies that the expression above can be estimated by $C\varepsilon^p(t-s)\mathbb{E}Y$. The remaining part of the sum appearing in (2.35), corresponding to $i = j-1$, upon performing the change of variables $u' := (u' - t_{i+1})\varepsilon^{-2}$ and $u := (t_{i+1} - u)\varepsilon^{-\gamma}$ transforms to

$$C\mathbb{E}Y\varepsilon^\gamma M\int_0^{\varepsilon^\gamma\Delta t}\left\{\int_0^{\varepsilon^{-2}\Delta t}\phi\left(u' + \frac{u}{\varepsilon^{2-\gamma}}\right)du'\right\}du \le o(1)(t-s)\mathbb{E}Y$$

and (2.34) follows.

On the other hand term $\mathcal{I}_2$, corresponding to $i=j$, can be estimated using again Lemma 2.2. From the fact that t_{j-1} is smaller than both u and u' it follows that:

$$
\begin{aligned}
|\mathcal{I}_2| &\le \frac{C}{\varepsilon^2}\sum_{j=0}^{M}\int_{t_j}^{t_{j+1}}\int_{t_j}^{t_{j+1}} \mathbb{E}\left\{V_p\left(\tfrac{u}{\varepsilon^2}, X(t_{j-1})\right)V_p\left(\tfrac{u'}{\varepsilon^2}, X(t_{j-1})\right)Y\right\}du'du \\
&\le \tfrac{2C}{\varepsilon^2}\sum_{j=0}^{M}\int_{t_j}^{t_{j+1}}\int_{u'}^{t_{j+1}} \phi\left(\tfrac{u-u'}{\varepsilon^2}\right)dudu'\mathbb{E}Y \\
&\le C(t-s)\mathbb{E}Y\int_0^\infty \phi(u)du.
\end{aligned}
$$

Thus (2.33) corresponding to $m=1$ follows.

A better estimate for L_1. Let us now go one step further and actually identify the limit of $E\{L_{1,j}(s,t)L_{1,m}(s,t)Y\}$ with $1\le j,m\le n$. The previous calculations already show that the term corresponding to $\mathcal{I}_1$ vanishes so we are interested only in the limit of $\mathcal{I}_2$. Using more carefully Lemma 2.2 we actually get:

$$
\begin{aligned}
\mathcal{I}_2 &= \tfrac{1}{\varepsilon^2}\sum_{j=0}^{M}\int_{t_j}^{t_{j+1}}\int_{t_j}^{t_{j+1}} \mathbb{E}\left\{V_p\left(\tfrac{u}{\varepsilon^2}, X(t_{j-1})\right)V_m\left(\tfrac{u'}{\varepsilon^2}, X(t_{j-1})\right)Y\right\}du'du \\
&= \tfrac{1}{\varepsilon^2}\sum_{j=0}^{M}\int_{t_j}^{t_{j+1}}\int_{t_j}^{t_{j+1}} R_{pm}\left(\tfrac{u-u'}{\varepsilon^2},0\right)dudu'\mathbb{E}Y + o(1)(t-s)\mathbb{E}Y \\
&= (a_{pm}+o(1))(t-s)\mathbb{E}Y,
\end{aligned}
$$

where a_{pm} are given by (2.24).

An estimate for L_2. Following a similar computation one can also obtain estimate (2.33) for $m=2$. In fact

$$
\mathbb{E}\left\{L_2^2(s,t)Y\right\} \le \varepsilon^p(t-s)^2\mathbb{E}Y \tag{2.36}
$$

for a suitable $p>0$. To see it we rewrite $\mathbb{E}\left\{L_2^2(s,t)Y\right\}$ in the form

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

Once again, we split the sum above into terms $\mathcal{I}_1'$ and $\mathcal{I}_2'$ that correspond to the summation over index ranges $i\le j-1$ and $i=j$.

The important difference with L_1 is that the term corresponding to $i=j$ is small:

$$
\begin{aligned}
&\mathcal{I}_2' \\
&= \tfrac{1}{\varepsilon^4}\sum_{i}\int_{t_i}^{t_{i+1}}du\int_{t_i}^{t_{i+1}}du'\int_{t_{i-1}}^{u}du_1\int_{t_{i-1}}^{u'}du_1'\mathbb{E}\left\{G_1\left(\tfrac{u}{\varepsilon^2},\tfrac{u_1}{\varepsilon^2},X(t_{i-1})\right)G_1\left(\tfrac{u'}{\varepsilon^2},\tfrac{u_1'}{\varepsilon^2},X(t_{i-1})\right)Y\right\} \\
&\le CM(\Delta t)^4\varepsilon^{-4}\mathbb{E}Y \le C\varepsilon^{3\gamma-4}(t-s)\mathbb{E}Y.
\end{aligned}
$$

Recall that $M\approx(t-s)/\varepsilon^\gamma$ and $\Delta t\approx\varepsilon^\gamma$. This means that if we take $\gamma>4/3$ this term is bounded by the right side of (2.36).

As for $\mathcal{I}_1'$ we estimate first the sum corresponding to summation range $i \le j-2$. Using Lemma 2.2 it can be bounded by

$$CM(\Delta t)^2\varepsilon^{-4}\sum_{j=1}^{M}\int\limits_{t_j}^{t_{j+1}} du' \int\limits_{t_{j-1}}^{u'} \phi^{1/2}\left(\tfrac{u'-u_1'}{\varepsilon^2}\right)\phi^{1/2}\left(\tfrac{u_1'-t_{j-1}}{\varepsilon^2}\right) du_1'\mathbb{E}Y$$
$$\le CM^2(\Delta t)^4\varepsilon^{-4}\phi^{1/2}\left(\varepsilon^{\gamma-2}\right)\mathbb{E}Y.$$

The last estimate follows from the fact that at either $u'-u_1'$ or $u_1'-t_{j-1}$ has to be greater than or equal to $\varepsilon^\gamma/2$. The right hand side of the above expression is bounded by

$$C\varepsilon^{2\gamma-4}\phi^{1/2}\left(\varepsilon^{\gamma-2}\right)(t-s)^2\mathbb{E}Y \le C\varepsilon^p(t-s)^2\mathbb{E}Y$$

for some $p>0$ by virtue of (2.23). The contribution of the terms with $i=j-1$ is estimated identically as in (2.37) and we end up with the estimate $|\mathcal{I}_1'| \le C\varepsilon^p(t-s)^2\mathbb{E}Y$ for some $p>0$, provided that γ is chosen as in the estimate for L_1.

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

$$\mathbb{E}\left\{(X_{p,\varepsilon}(t)-X_{p,\varepsilon}(s))(X_{m,\varepsilon}(t)-X_{m,\varepsilon}(s))Y\right\} = \left(a_{pm}+o(1)\right)(t-s)\mathbb{E}Y \qquad (2.37)$$

for all t,s with $t-s\ge 10\varepsilon^\gamma$ and $\varepsilon\to 0$. This, of course, implies (2.30) and hence the tightness of the laws of $\{X_\varepsilon(t),\, t\ge 0\}$ follows.

Identification of the limit. Suppose that $f(\cdot)$ belongs to $C_0^\infty(\mathbb{R}^d)$, the space of all compactly supported function C^∞ functions on $\mathbb{R}^d$, and $\Phi(\cdot)$ is a nonnegative, continuous and bounded function on $\mathbb{R}^{dN}$ for some integer N. A slight modification of our previous calculation can be carried out in order to compute

$$\mathbb{E}\left\{[f(X_\varepsilon(t))-f(X_\varepsilon(s))]\Psi_\varepsilon\right\},$$

where $\Psi_\varepsilon := \Psi(X_\varepsilon(s_1),\dots,X_\varepsilon(s_N))$ for some $0\le s_1\le\dots\le s_N\le s$. In this case we can write

$$f(X_\varepsilon(t))-f(X_\varepsilon(s))$$
$$=\sum_{i=0}^{M}\left\{\nabla f(X_\varepsilon(t_{i-1}))\cdot\left[\frac{1}{\varepsilon}\int\limits_{t_i}^{t_{i+1}} V\left(\tfrac{u}{\varepsilon^2},X_\varepsilon(t_{i-1})\right)du+\frac{1}{\varepsilon^2}\int\limits_{t_i}^{t_{i+1}}\int\limits_{t_{i-1}}^{u} G_1\left(\tfrac{u}{\varepsilon^2},X_\varepsilon(t_{i-1})\right)du\right]\right.$$
$$+\frac{1}{\varepsilon^2}\partial^2_{x_jx_m}f(X_\varepsilon(t_{i-1})\sum_{j,m=1}^{n}\int_{t_i}^{t_{i+1}}\int_{t_{i-1}}^{u} V_j\left(\tfrac{u}{\varepsilon^2},X_\varepsilon(t_{i-1})\right)V_m\left(\tfrac{u_1}{\varepsilon^2},X_\varepsilon(t_{i-1})\right)du_1du$$
$$+\frac{1}{\varepsilon^3}\sum_{i=0}^{M}\int_{t_i}^{t_{i+1}}\left[\int_{t_{i-1}}^{u}\left[\int_{t_{i-1}}^{u_1} r\left(\tfrac{u}{\varepsilon^2},\tfrac{u_1}{\varepsilon^2},\tfrac{u_2}{\varepsilon^2},X(u_2)\right)du_2\right]du_1\right]du,$$

where $r(\cdot)$ is a certain function that is deterministically bounded. Thus, the corresponding term in the formula above can be estimated by $C\varepsilon^p$ for some $p>0$, provided that γ is suitably chosen. In fact one can choose an arbitrary $\gamma\in(3/2,2)$. Denote the remaining terms on the right hand side of (2.38) by $J_{1,\varepsilon}$ and $J_{2,\varepsilon}$ respectively. The application of the mixing lemma, in the same way as we have done it before, leads to the conclusion that $\lim_{\varepsilon\to 0}\mathbb{E}\left\{J_{1,\varepsilon}\Psi_\varepsilon\right\}=0$, and, by the same token,

$$\mathbb{E}\left\{J_{2,\varepsilon}\Psi_\varepsilon\right\}$$
$$=\frac{1}{\varepsilon^2}\sum_{j,m=1}^{n}\mathbb{E}\left\{\partial^2_{x_jx_m}f(X_\varepsilon(t_{i-1}))\Psi_\varepsilon\int_{t_i}^{t_{i+1}}\int_{t_{i-1}}^{u} R_{j,m}\left(\tfrac{u-u_1}{\varepsilon^2},0\right)dudu_1\right\}+o(1)$$

as $\varepsilon \to 0$. After straightforward computations we obtain that

$$\mathbb{E}\left\{[f(X_\varepsilon(t)) - f(X_\varepsilon(s))]\Psi_\varepsilon\right\} = \frac{1}{2}\sum_{j,m=1}^{n} \mathbb{E}\left[a_{jm}\int_s^t \partial^2_{x_j x_m} f(X_\varepsilon(u))du\Psi_\varepsilon\right] + o(1)$$

for all $f(\cdot) \in C_0^\infty(\mathbb{R}^n)$ and bounded non-negative functions Ψ_ε as described above. Since we already know that the family of laws of $\{X_\varepsilon(t),\, t \geq 0\}$ is tight in $C[0, +\infty)$ we conclude that any limiting measure for the family must be a solution of the martingale problem corresponding to the operator $Lf := (1/2)\sum_{j,m=1}^{n} a_{jm}\partial^2_{x_j x_m} f$, i.e. the process $f(X(t)) - \int_0^t Lf(X(s))ds$ is a martingale under this measure. There is only one solution to this problem and it is a Wiener measure that is the law of the Brownian motion appearing in the statement of the theorem.□

The proof of Theorem 2.1 is, in a sense, very generic in limit theorems based on trajectory considerations in the weak coupling regime: one needs to establish a mixing lemma of some sort, then verify tightness by using arguments with stepping back in time, and using the Taylor expansion until some order where the correction becomes deterministically small. Then the mixing lemma is applied to eliminate some explicit "apparently large but small due to mixing" terms, and the rest is computed explicitly.

2.3. **One and two particles in a random flow with a strong drift.** An important random flow that is close in spirit to a one-way wave equation is a flow of the form

$$V(x) = \bar{u} + \varepsilon v(x). \tag{2.38}$$

Here $\bar{u} \neq 0$ is the mean flow and $v(x)$ is a (time-independent) random perturbation. We assume that $v(x)$ is divergence-free: $\nabla \cdot v = 0$. The corresponding trajectory is

$$\frac{dX_\varepsilon}{dt} = \bar{u} + \varepsilon v(X_\varepsilon),\ X_\varepsilon(0) = x. \tag{2.39}$$

Note that the flow $V(x)$ is time-independent. This, however, is not a problem since the large mean flow $\bar{u}$ "always takes the particle to new places" – hence, from the point of view of the trajectory $X_\varepsilon(t)$ it always sees a new medium, and the increments of $X_\varepsilon(t)$ are nearly independent if $v(x)$ is sufficiently strongly mixing in space. We will assume here that $v(x)$ satisfies the assumptions of Theorem 2.1 (apart, obviously, from the mixing requirement in time, but we do assume mixing in space). This problem was studied by Kesten and Papanicolaou in [52].

We are interested here in two questions: first, how does the deviation from the straight line $\bar{X}(t) = \bar{u}t$ develop in time, and how do two particles starting at initially nearby positions, diverge? Both questions can be answered with the help of Theorem 2.1, and, as we will see, the particle separation satisfies an equation of a kinetic type. As before, in order for the effect of a weak random fluctuation to have order one, we have to consider the times of the order $t \sim O(\varepsilon^{-2})$. Accordingly, we introduce the deviation from the straight line

$$Z_\varepsilon(t) = X_\varepsilon\left(\frac{t}{\varepsilon^2}\right) - \frac{\bar{u}t}{\varepsilon^2} - x.$$

It satisfies

$$\frac{dZ_\varepsilon}{dt} = \frac{1}{\varepsilon}v\left(x + \frac{\bar{u}t}{\varepsilon^2} + Z_\varepsilon\right),\ Z_\varepsilon(0) = 0. \tag{2.40}$$

A straightforward application of Theorem 2.1 shows that Z_ε converges in law, as $\varepsilon \to 0$ to the Brownian motion with the diffusion matrix given by (2.24):

$$a_{ij}^Z = \int_{-\infty}^{\infty} R_{ij}(\bar{u}t)dt, \tag{2.41}$$

where

$$R_{ij}(x) = \mathbb{E}[v_i(y)v_j(x+y)]$$

is the covariance tensor of the field $v(x)$.

Now, consider two solutions of (2.39) that start initially at the points x and $x-y$. Their separation

$$Y_\varepsilon(t) = X_\varepsilon\left(\frac{t}{\varepsilon^2}, x\right) - X_\varepsilon\left(\frac{t}{\varepsilon^2}, x-y\right)$$

satisfies

$$\frac{dY_\varepsilon}{dt} = \frac{1}{\varepsilon}\left[v\left(x + \frac{\bar{u}t}{\varepsilon^2} + Z_\varepsilon\right) - v\left(x + \frac{\bar{u}t}{\varepsilon^2} + Z_\varepsilon - Y_\varepsilon\right)\right],\ Y_\varepsilon(0) = y. \tag{2.42}$$

A slight modification of Theorem 2.1 needed to account for the explicit dependence of the right side in (2.42) on Z_ε implies that the processes $(Z_\varepsilon, Y_\varepsilon)$ converge weakly to the correlated diffusion processes Z and Y whose joint generator is

$$Lf(z,y) = \frac{1}{2}\sum_{i,j=1}^{n}\left(a_{ij}^Z \frac{\partial^2 f}{\partial z_i \partial z_j} + b_{ij}^{Z,Y}\frac{\partial^2 f}{\partial y_i \partial y_j} + c_{ij}^Y \frac{\partial^2 f}{\partial z_i \partial y_j}\right), \tag{2.43}$$

where a_{ij}^Z are given by (2.41),

$$b_{ij}^{Z,Y}(y) := \int_{-\infty}^{\infty}\left(2R_{ij}(\bar{u}t) - R_{ij}^s(\bar{u}t + y) - R_{ij}^s(\bar{u}t - y)\right)dt$$

and

$$c_{ij}^Y(y) := 2\int_{-\infty}^{\infty}\left(R_{ij}(\bar{u}t) - R_{ij}^s(\bar{u}t - y)\right)dt.$$

Here $R_{ij}^s(y) := (1/2)[R_{ij}(y) + R_{ji}(y)]$. The individual generators for Z and Y are:

$$L_Z f(z) = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}^Z \frac{\partial^2 f}{\partial z_i \partial z_j}, \tag{2.44}$$

and

$$L_Y f(y) = \frac{1}{2}\sum_{i,j=1}^{n} c_{ij}^Y(y)\frac{\partial^2 f}{\partial y_i \partial y_j}, \tag{2.45}$$

The diffusion coefficient in (2.45) vanishes for $y = c\bar{u}$, so that if two particles start at two nearby positions on the same straight line in the direction of $\bar{u}$ then their separation is not changed by the flow in the limit. The generator L_Y is asymptotically close to L_Z for all y that have large component in the direction perpendicular to the mean flow $\bar{u}$. The reason for this is that when the two starting points are separated by a large distance in the direction normal to $\bar{u}$, their trajectories are almost independent, and the rescaled difference trajectory behaves like the fluctuations of each individual trajectory.

Interpretation in terms of the first order PDEs. Let us now relate the foregoing discussion to solutions of the first order linear PDEs. Let ϕ be the solution of

$$\frac{\partial \phi}{\partial t} + (\bar{u} + \sqrt{\varepsilon} v(x)) \cdot \nabla \phi = 0, \tag{2.46}$$

with an incompressible random flow $v(x)$ as above. Rescaling time and space as $t \to t/\varepsilon$, $x \to x/\varepsilon$, leads to the following initial value problem $\phi^\varepsilon(t,x) = \varepsilon^{-n/2}\phi(t/\varepsilon, x/\varepsilon)$ (the rescaling factor $\varepsilon^{-n/2}$ keeps the L^2-norm of the initial data fixed)

$$\frac{\partial \phi^\varepsilon}{\partial t} + \left(\bar{u} + \sqrt{\varepsilon} v\left(\frac{x}{\varepsilon}\right)\right) \cdot \nabla \phi^\varepsilon = 0,\ \phi^\varepsilon(0,x) = \frac{1}{\varepsilon^{n/2}}\phi_0\left(\frac{x}{\varepsilon}\right). \tag{2.47}$$

More generally, we may consider (2.47) with the initial data $\phi_0^\varepsilon(x)$ forming an ε-oscillatory family [43], such as the WKB data

$$\phi^\varepsilon(x) = A(x) e^{iS(x)/\varepsilon} \tag{2.48}$$

where $A(x)$ and $S(x)$ are smooth functions, and $S(x)$ is real valued. The latter family describes the distribution of tracers which have the form of high frequency waves propagating in the direction $\nabla S(x)$ with amplitude $A(x)$.

We are interested in the (non-symmetrized) Wigner transform

$$W^\varepsilon(t,x,k) = \int e^{ik\cdot y} \phi^\varepsilon(t, x - \varepsilon y)\phi^{\varepsilon *}(t,x)\frac{dy}{(2\pi)^n}, \tag{2.49}$$

and the two-point correlation function

$$\hat{W}^\varepsilon(t,x,y) = \phi^\varepsilon(t, x-\varepsilon y)\phi^{\varepsilon *}(t,x), \tag{2.50}$$

so that its expectation is the correlation function of the filed $\phi^\varepsilon(x)$ at two nearby points separated by εy. The Wigner transform has a weak limit $W(t,x,y)$ in the space of Schwartz distributions that is a non-negative measure [43]. The diffusive limit for the process $Y^\varepsilon(t)$ described above implies the following: let $\hat{W}(t,x,y)$ be the weak limit of $\mathbb{E}[\hat{W}^\varepsilon(t,x,y)]$. Then $\hat{W}(t,x,y)$ satisfies the diffusion equation with a drift:

$$\frac{\partial \hat{W}}{\partial t} + \bar{u} \cdot \nabla_x \hat{W} = L_Y \hat{W}, \tag{2.51}$$

where the operator L_Y is given by (2.45).

Applying the inverse Fourier transform to (2.51) we conclude that the $\mathbb{E}\left\{W^\varepsilon(t,x,k)\right\}$ converges to $W(t,x,k)$ weakly in $\mathcal{S}'$, where $W(t,x,k)$ satisfies a kinetic equation

$$\begin{aligned}&\frac{\partial W}{\partial t} + \bar{u} \cdot \nabla_x W \\ &= \int \frac{dk'}{(2\pi)^{n-1}} k_i k_j \hat{R}^s_{ij}(k'-k)\delta((k'-k)\cdot \bar{u})(W(t,x,k') - W(t,x,k)).\end{aligned} \tag{2.52}$$

Equation (2.52) has the form of a radiative transport equation with the dispersion law $\omega(k) = \bar{u} \cdot k$.

3. **The Wigner transform.**

3.1. **The basic properties of the Wigner transform.**

3.1.1. *The unscaled Wigner transform.* The Schrödinger equation

$$i\phi_t + \frac{1}{2}\Delta\phi - V(t,x)\phi = 0. \tag{3.1}$$

with a real potential $V(t,x)$ preserves the total energy of the solution (or the total number of particles depending on the point of view or physical application):

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

This 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 of 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, for instance, 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 as 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 [78] (where he gives credit to Szilard for this idea). Wigner introduced the following object:

$$W(t,x,k) = \int \phi\left(t, x - \frac{y}{2}\right)\bar{\phi}\left(t, x + \frac{y}{2}\right) e^{ik\cdot y}\frac{dy}{(2\pi)^d}. \tag{3.2}$$

It is immediate to check that

$$\int W(t,x,k)dk = |\phi(t,x)|^2 = E(t,x), \tag{3.3}$$

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

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

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

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

reinforcing 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.

In order to get an equation for W in the special case $V = 0$, we differentiate it with respect to t, use the Schrödinger equation for ϕ and integrate by parts:

$$\begin{aligned}\frac{\partial W}{\partial t} &= -\tfrac{1}{2i}\int e^{ik\cdot y}\left[\bar{\phi}\left(t,x+\tfrac{y}{2}\right)\Delta\phi\left(t,x-\tfrac{y}{2}\right)-\phi\left(t,x-\tfrac{y}{2}\right)\Delta\bar{\phi}\left(t,x+\tfrac{y}{2}\right)\right]\tfrac{dy}{(2\pi)^d}\\ &= \tfrac{1}{i}\int e^{ik\cdot y}\nabla_x\cdot\nabla_y\left[\phi\left(t,x-\tfrac{y}{2}\right)\bar{\phi}\left(t,x+\tfrac{y}{2}\right)\right]\tfrac{dy}{(2\pi)^d}\\ &= -k\cdot\nabla_x W.\end{aligned}$$

We obtain a (closed) kinetic equation for $W(t, x, k)$:

$$W_t + k\cdot\nabla_x W = 0. \tag{3.4}$$

Therefore, one may describe energy density evolution for the Schrödinger equation with zero potential as follows: compute the initial data $W(0, x, k)$, solve the kinetic equation (3.4) and find $|\phi(t, x)|^2$ using (3.3).

However, there is one drawback in the interpretation of $W(t, x, k)$ as electron energy density resolved over positions and momenta – there is no reason for $W(t, x, k)$ to be non-negative! Moreover, the same analysis for the Schrödinger equation (3.1) with a non-zero potential V leads to the following evolution equation for $W(t, x, k)$:

$$\frac{\partial W}{\partial t} + k\cdot\nabla_x W = \frac{1}{i}\int e^{ip\cdot x}\hat{V}(p)\left[W\left(k-\frac{p}{2}\right)-W\left(k-\frac{p}{2}\right)\right]\frac{dp}{(2\pi)^n}. \tag{3.5}$$

While the uniform kinetic equation (3.4) possesses 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 straight lines $\dot{X} = K$, $\dot{K} = 0$,– the Wigner equation (3.5) has very few attractive features. 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) until the work of P. Gérard and L. Tartar in the late eighties. They realized that the Wigner transforms became 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).

3.1.2. *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 (3.2) is

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

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 to $\delta(k - k_0)$. This motivates the following definition of the (rescaled) Wigner transform of a family of functions $\phi_\varepsilon(x)$:

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

that may be more conveniently re-written as

Definition 3.1. The Wigner transform (or the Wigner distribution) of a family of functions $\phi_\varepsilon(x)$ is a distribution $W_\varepsilon(x,k) \in \mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ defined by

$$W_\varepsilon(t,x,k) = \int \phi_\varepsilon\left(x - \frac{\varepsilon y}{2}\right) \bar{\phi}_\varepsilon\left(x + \frac{\varepsilon y}{2}\right) e^{ik\cdot y} \frac{dy}{(2\pi)^d}. \tag{3.6}$$

Expression (3.6) shows that $W_\varepsilon(x,k)$ is well suited to study functions oscillating 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 $O(1)$.

We will be mostly using the Wigner transform for families of solutions of non-dissipative evolution equations that conserve the L^2-norm (or a weighted L^2-norm). The scaling in (3.6) is particularly well suited for families of functions $\phi_\varepsilon(x)$ that are uniformly (in $\varepsilon \in (0,1)$) bounded in $L^2(\mathbb{R}^d)$. To see that we multiply W_ε by a test function $\lambda(x,k)$ and integrate:

$$\begin{aligned}\langle W_\varepsilon, \lambda\rangle &= \int W_\varepsilon(x,k)\bar{\lambda}(x,k)dxdk \\ &= \int \phi_\varepsilon\left(x - \tfrac{\varepsilon y}{2}\right) \bar{\phi}_\varepsilon\left(x + \tfrac{\varepsilon y}{2}\right) \bar{\lambda}(x,k) \tfrac{e^{ik\cdot y}dydxdk}{(2\pi)^n} \\ &= \int \phi_\varepsilon\left(x - \tfrac{\varepsilon y}{2}\right) \bar{\phi}_\varepsilon\left(x + \tfrac{\varepsilon y}{2}\right) \overline{\tilde{\lambda}}(x,y) \tfrac{dydx}{(2\pi)^n}.\end{aligned}$$

Here $\tilde{\lambda}(x,y)$ is the Fourier transform of λ in the variable k only:

$$\tilde{\lambda}(x,y) = \int e^{-ik\cdot y}\lambda(x,k)dk. \tag{3.7}$$

Using the Cauchy-Schwartz inequality we arrive at the following estimate:

$$|\langle W_\varepsilon, \lambda\rangle| \le \|\phi_\varepsilon\|^2_{L^2(\mathbb{R}^d)} \int \sup_x \left[\left|\tilde{\lambda}(x,y)\right|\right] \frac{dy}{(2\pi)^d}$$

Let us define the space of test functions

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

with the norm

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

We have just shown that if the family of functions $\phi_\varepsilon(x)$ be uniformly bounded in $L^2(\mathbb{R}^d)$ then the corresponding family of Wigner transforms $W_\varepsilon(x,k)$ is uniformly bounded in $\mathcal{A}'(\mathbb{R}^d \times \mathbb{R}^d)$. Thus, the family $W_\varepsilon(x,k)$ has a weak-$\star$ converging subsequence in the space $\mathcal{A}'(\mathbb{R}^d \times \mathbb{R}^d)$.

Remarkably, any limit point of $W_{\varepsilon_k}(x,k)$ in $\mathcal{A}'$ is a non-negative measure of a bounded total mass. A convenient way to see this is to use the Husimi function, which is a convolution of the Wigner function with an approximation of a delta function on the intermediate scale $O(\sqrt{\varepsilon})$

$$\tilde{W}_\varepsilon(x,k) = W_\varepsilon \star G_\varepsilon,\ \ G_\varepsilon(x,k) = \frac{1}{(\pi\varepsilon)^n} e^{-[|x|^2+|k|^2]/\varepsilon}.$$

A straightforward computation shows that $\tilde{W}_\varepsilon$ is non-negative. It is also straightforward to verify that for any function $\phi \in \mathcal{A}$ the sequence $\phi \star G_{\varepsilon_k}$ converges strongly to ϕ in $\mathcal{A}$ as $\varepsilon_k \to 0$. Hence, any limit point of the family W_ε has to be non-negative, and, in addition the families W_ε and $\tilde{W}_\varepsilon$ have the same limit points.

In addition, as

$$\int \tilde{W}_\varepsilon(x,k)dxdk = \int |\phi_\varepsilon(x)|^2 dx$$

and the family ϕ_ε is uniformly bounded in $L^2(\mathbb{R}^d)$, it follows that the measure $W(dxdk)$ has a bounded total mass.

Positivity of the Husimi function provides a quantitative way to measure the potential non-positivity of the Wigner transform: its local averages over regions of size $\sqrt{\varepsilon}$ are non-negative.

We summarize the above into the following theorem.

Theorem 3.2. *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 $L^2(\mathbb{R}^d \times \mathbb{R}^d)$ as the L^2-norm of W_ε is unbounded unless $\phi_\varepsilon(x)$ converges strongly to zero:

$$\int |W_\varepsilon(x,k)|^2 dxdk = \int \left|\phi_\varepsilon\left(x - \frac{\varepsilon y}{2}\right)\phi_\varepsilon\left(x + \frac{\varepsilon y}{2}\right)\right|^2 \frac{dydx}{(2\pi)^d} = \frac{1}{(2\pi\varepsilon)^d}\|\phi_\varepsilon\|^4_{L^2(\mathbb{R}^d)}.$$

Therefore, it is impossible to expect even weak convergence of W_ε in $L^2(\mathbb{R}^d \times \mathbb{R}^d)$ unless the family ϕ_ε converges strongly to zero. In that case, however, $W_\varepsilon = 0$, which is a case of somewhat limited interest.

3.1.3. *The semiclassical operators.* In order to better understand the limits of the Wigner transforms we note that for any test function $a(x,k) \in \mathcal{S}(\mathbb{R}^d \times \mathbb{R}^d)$ we have

$$\langle a, W_\varepsilon\rangle = \langle a^w(x,\varepsilon D)\phi_\varepsilon, \phi_\varepsilon\rangle. \tag{3.8}$$

Here we have associated to the function $a(x,k)$ the Weyl operator $a^w(x,\varepsilon D)$ defined by

$$[a^w(x,\varepsilon D)f](x) = \int a\left(\frac{x+y}{2}, \varepsilon k\right) f(y) e^{ik\cdot(x-y)}\frac{dydk}{(2\pi)^n}. \tag{3.9}$$

The Weyl operators provide one way to associate an operator to a symbol $a(x,k)$, another example of such semiclassical operator is

$$[a(x,\varepsilon D)f](x) = \int a\left(x, \varepsilon k\right) f(y) e^{ik\cdot(x-y)}\frac{dydk}{(2\pi)^n} \tag{3.10}$$

The operator $a(x,\varepsilon D)$ is called the standard quantization of the symbol $a(x,k)$.

Semiclassical operators, both in the Weyl quantization (3.9) and in the standard quantization (3.10) provide a useful and effective tool in understanding the limits of the Wigner transforms. We will now recall some of their properties, without proofs. The next proposition provides flexibility in the choice of quantization.

Proposition 3.3. *Let $a \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$ be a test function, then the Weyl operators $a^w(x,\varepsilon D)$ and the semiclassical operators $a(x,\varepsilon D)$ are asymptotically equivalent in the sense that*

$$\|a(x,\varepsilon D) - a^w(x,\varepsilon D)\|_{L^2\to L^2} \to 0 \tag{3.11}$$

as $\varepsilon \to 0$.

The next proposition provides the uniform L^2-bounds on $a(x,\varepsilon D)$ and the H^s-estimates.

Proposition 3.4. *Let $a(x,k) \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$ then for $s \geq 0$ there exist positive constants $C_s > 0$ so that*

$$\|a(x,\varepsilon D)\|_{L^2\to L^2} + \|a^w(x,\varepsilon D)\|_{L^2\to L^2} \leq C_0(a) \tag{3.12}$$

$$\varepsilon^s \|a(x,\varepsilon D)\|_{H^{-s}\to L^2} + \varepsilon^s \|a^w(x,\varepsilon D)\|_{H^{-s}\to L^2} \leq C_s(a) \tag{3.13}$$

$$\varepsilon^s ||a(x,\varepsilon D)||_{L^2\to H^s} + \varepsilon^s ||a^w(x,\varepsilon D)||_{L^2\to H^s} \leq C_s(a). \tag{3.14}$$

But the key property that makes the semiclassical operators so useful for us is that the product of two semiclassical operators corresponds to the operator which is nearly the product of their symbols.

Proposition 3.5. *The product of two operators $a(x,\varepsilon D)$, $b(x,\varepsilon D)$ is*

$$b(x,\varepsilon D)a(x,\varepsilon D) = (ba)(x,\varepsilon D) + \frac{\varepsilon}{i}(\nabla_k b \cdot \nabla_x a)(x,\varepsilon D) + \varepsilon^2 R_\varepsilon, \tag{3.15}$$

where the operators R_ε are uniformly bounded on $L^2(\mathbb{R}^n)$. The product of the Weyl quantized operators is

$$\begin{aligned} &b^w(x,\varepsilon D)a^w(x,\varepsilon D)\\ &= (ba)^w(x,\varepsilon D) + \tfrac{\varepsilon}{2i}\left[(\nabla_k b \cdot \nabla_x a)^w(x,\varepsilon D) - (\nabla_x b \cdot \nabla_k a)^w(x,\varepsilon D)\right] + \varepsilon^2 Q_\varepsilon, \end{aligned} \tag{3.16}$$

where the operators Q_ε are uniformly bounded on L^2. One of the symbols a and b above may grow polynomially in k.

Finally, the adjoint operators are as follows.

Proposition 3.6. *The operators $a^w(x,\varepsilon D)$ and $a(x,\varepsilon D)$ have the following adjoints:*

$$[a^w(x,\varepsilon D)]^* = [\bar{a}]^w(x,\varepsilon D) \tag{3.17}$$

and

$$a(x,\varepsilon D)^* = \bar{a}(x,\varepsilon D) + \varepsilon R_\varepsilon \tag{3.18}$$

with the operators R_ε uniformly bounded on L^2.

3.1.4. *Examples of the Wigner measures.* We now present some examples of the Wigner measures – they are easy to compute with the help of the semiclassical operators.

A strongly converging sequence. Let $\phi_\varepsilon(x)$ converge strongly in $L^2(\mathbb{R}^d)$ to a limit $\phi(x)$. Then the limit Wigner measure is $W(x,k) = |\phi(x)|^2\delta(k)$. To see this, we take a test function $a(x,k)$ and write

$$(a(x,\varepsilon D)\phi_\varepsilon, \phi_\varepsilon) = (a(x,\varepsilon D)[\phi_\varepsilon - \phi], \phi_\varepsilon) + (a(x,\varepsilon D)\phi, [\phi_\varepsilon - \phi]) + (a(x,\varepsilon D)\phi, \phi).$$

The first two terms above tend to zero as $\varepsilon \to 0$ as $\|\phi_\varepsilon - \phi\|_{L^2} \to 0$. Moreover, we also have

$$a(x,\varepsilon D)\phi \to a(x,0)\phi(x) \text{ in } L^2(\mathbb{R}^d)$$

as $\varepsilon \to 0$. It follows that

$$\langle a, W_\varepsilon\rangle \to \int a(x,0)|\phi(x)|^2 dx,$$

and thus the limit Wigner measure is indeed $W(x,k) = |\phi(x)|^2\delta(k)$. This means that unless the family ϕ_ε oscillates on a small scale, 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 given by $W(x,k) = (2\pi)^{-d}|\hat{\phi}(k)|^2\delta(x)$. This is verified as follows:

$$\begin{aligned}\langle a, W_\varepsilon\rangle &= \int a(x,k)\phi\left(\frac{x}{\varepsilon}-\frac{y}{2}\right)\bar{\phi}\left(\frac{x}{\varepsilon}+\frac{y}{2}\right)e^{ik\cdot y}\frac{dydxdk}{(2\pi\varepsilon)^d}\\ &= \int a(\varepsilon x,k)\phi\left(x-\frac{y}{2}\right)\bar{\phi}\left(x+\frac{y}{2}\right)e^{ik\cdot y}\frac{dydxdk}{(2\pi)^d}\\ &\to \int a(0,k)\phi(x)\bar{\phi}(z)e^{ik\cdot(z-x)}\frac{dzdxdk}{(2\pi)^d} = \int\int a(0,k)|\hat{\phi}(k)|^2\frac{dk}{(2\pi)^d}.\end{aligned}$$

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

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

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

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

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

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

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

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

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.

3.1.5. *Basic properties of the Wigner measures.* We have shown in Theorem 3.2 that if $\phi_\varepsilon(x)$ is a family of uniformly bounded functions in $L^2(\mathbb{R}^n)$ then the limit Wigner distribution is a non-negative measure of finite mass. Positivity of the limit has been proved using the Husimi transform. Another way to see the positivity of the limit is using the semiclassical operators. The distribution W is non-negative if for any non-negative test function $a(x,k) \geq 0$ the pairing $\langle a, W\rangle$ is non-negative. It suffices to check this property for functions of the form $a(x,k) = |b(x,k)|^2$ as these are dense among non-negative functions. However, for such functions we have

$$\langle a, W\rangle = \lim_{\varepsilon\to 0}\langle |b|^2, W_\varepsilon\rangle = \lim_{\varepsilon\to 0}\langle [|b|^2]^w(x,\varepsilon D)\phi_\varepsilon, \phi_\varepsilon\rangle = \lim_{\varepsilon\to 0}\langle b^w(x,\varepsilon D)\phi_\varepsilon, b^w(x,\varepsilon D)\phi_\varepsilon\rangle \geq 0.$$

We used (3.16) and (3.17) in the third equality above. Yet another way to show the positivity (non-negativity) of the limit Wigner distribution is by using the sharp Gärding inequality

$$\langle a^w(x,\varepsilon D)u, u\rangle \geq -Ch\|u\|_{L^2}^2,\ 0 \leq h \leq h_0, \tag{3.19}$$

which holds for all non-negative symbols $a(x,k) \geq 0$. The proof of this inequality may be found in [38], which provides an excellent introduction to the semiclassical analysis.

Another important fact is that the Wigner measure is a local notion in space. 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 3.7 (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 neighborhood of a point x_0. Then the the limit Wigner measures $\mu[\phi]$ and $\mu[\psi]$ coincide in this neighborhood.*

The localization property is quite useful because it allows the consideration of Wigner measures for families of functions ϕ_ε that are uniformly bounded in $L^2_{loc}(\mathbb{R}^d)$ and not in $L^2(\mathbb{R}^d)$.

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 3.8 (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" can be seen on the family

$$\phi_\varepsilon(x) = \theta\left(x - \frac{1}{\varepsilon}\right) \tag{3.20}$$

with $\theta(x) \in C_c^\infty(\mathbb{R}^d)$. 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 [42]. The ε-oscillatory property guarantees that the functions ϕ_ε oscillate on a scale which is not smaller than $O(\varepsilon)$, and is conveniently characterized by the following definition.

Definition 3.9. 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 $\theta(x)$

$$\limsup_{\varepsilon\to 0} \int_{|\xi|\geq R/\varepsilon} |\widehat{\theta\phi_\varepsilon}(\xi)|^2 d\xi \to 0 \text{ as } R \to +\infty. \tag{3.21}$$

A simple and intuitive sufficient condition for (3.21) is that there exist a positive integer j and a constant C independent of ε such that

$$\varepsilon^j \left\| \frac{\partial^j f_\varepsilon}{\partial x^j} \right\|_{L^2_{loc}} \leq C. \tag{3.22}$$

Condition (3.22) 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 happens with the family (3.20)) we introduce the following definition.

Definition 3.10. A bounded family $\phi_\varepsilon(x) \in L^2(\mathbb{R}^n)$ is said to be compact at infinity if

$$\limsup_{\varepsilon\to 0} \int_{|x|\geq R} |\phi_\varepsilon(x)|^2 dx \to 0 \text{ as } R \to +\infty. \tag{3.23}$$

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.

Proposition 3.11. *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*

$$\int_{\mathbb{R}^d} \mu(\cdot, dk) \leq \nu \tag{3.24}$$

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

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \mu(dx,dk) \leq \limsup_{\varepsilon\to 0} \int_{\mathbb{R}^d} |\phi_\varepsilon(x)|^2 dx \tag{3.25}$$

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 *on the right side of (3.25).*

With all the above properties we can interpret the limiting Wigner measure as the limit phase space energy density of the family ϕ_ε, that is, energy density resolved over directions and wavenumbers.

3.2. **The evolution of the Wigner transform.**

3.2.1. *The Liouville equation and geometric optics.* 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

$$i\varepsilon\frac{\partial\phi_\varepsilon}{\partial t}+\frac{\varepsilon^2}{2}\Delta\phi_\varepsilon-V(x)\phi_\varepsilon(x)=0 \tag{3.26}$$

with a smooth potential $V(x)$. The initial data $\phi_\varepsilon(0,x)=\phi_\varepsilon^0(x)$ forms an ε-oscillatory and compact at infinity family of functions uniformly bounded in $L^2(\mathbb{R}^d)$. As (3.26) preserves the L^2-norm of solutions, the family $\phi_\varepsilon(t,x)$ is bounded in $L^2(\mathbb{R}^d t)$ for each $t\geq 0$ and it makes sense to define the Wigner transform

$$W_\varepsilon(t,x,k)=\int\psi_\varepsilon\left(t,x-\frac{\varepsilon y}{2}\right)\bar{\psi}_\varepsilon\left(t,x+\frac{\varepsilon y}{2}\right)e^{ik\cdot y}\frac{dy}{(2\pi)^n}. \tag{3.27}$$

We first obtain the equation for the limit Wigner transform directly "by hand". Differentiating (3.27) with respect to time, using (3.26) and a calculation similar to (3.4), we arrive at the following equation for the Wigner transform

$$W_t^\varepsilon+k\cdot\nabla_x W^\varepsilon-\mathcal{L}^\varepsilon W^\varepsilon=0. \tag{3.28}$$

Here the operator $\mathcal{L}^\varepsilon$ is defined by

$$\mathcal{L}^\varepsilon Z(x,k)=\frac{i}{\varepsilon}\int_{\mathbb{R}^d}e^{ip\cdot x}\hat{V}(p)\left[Z(x,k-\frac{\varepsilon p}{2})-Z(x,k+\frac{\varepsilon p}{2})\right]\frac{dp}{(2\pi)^d} \tag{3.29}$$

acting on a smooth function $Z(x,k)$ in phase space.

It is easy to verify that for any smooth and decaying function $Z(x,k)$ we have

$$\mathcal{L}^\varepsilon Z\to\mathcal{L}Z=\nabla_x V\cdot\nabla_k Z \text{ in } \mathcal{A}. \tag{3.30}$$

We conclude that, first, for any test function $Z\in\mathcal{S}(\mathbb{R}^n\times\mathbb{R}^n)$ we have

$$\left|\frac{\partial}{\partial t}\langle W_\varepsilon(t),Z\rangle\right|\leq C(Z)$$

so that $W_\varepsilon(t)$ is uniformly continuous in $\mathcal{S}'$. Further, we can pass to the limit in

$$\langle\frac{\partial W_\varepsilon}{\partial t},Z\rangle+\langle k\cdot\nabla_x W_\varepsilon,Z\rangle=\langle\mathcal{L}_\varepsilon W_\varepsilon,Z\rangle$$

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

$$W_t+k\cdot\nabla_x W-\nabla V\cdot\nabla_k W=0 \tag{3.31}$$

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

Proposition 3.12. *Let the family $\phi_\varepsilon^0(x)$ 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 (3.31) with the initial data $W(0,x,k)=W_0(x,k)$.*

The formalism of semiclassical pseudodifferential operators provides a fast and elegant way to derive the Liouville equation (3.31). For any test function $a(x,k)\in$

$\mathcal{S}(\mathbb{R}^d \times \mathbb{R}^d)$ we have, using (3.26) and the product formula (3.15) with $\omega(x,k) = \dfrac{k^2}{2} + V(x)$:

$$\begin{aligned}
\langle a, \frac{\partial W}{\partial t}\rangle &= \lim_{\varepsilon\to 0}\left[(a(x,\varepsilon D)\phi_\varepsilon, \frac{\partial \phi^\varepsilon}{\partial t}) + (a(x,\varepsilon D)\frac{\partial \phi_\varepsilon}{\partial t}, \phi^\varepsilon)\right] \\
&= \lim_{\varepsilon\to 0}\left\{\left(a(x,\varepsilon D)\phi_\varepsilon, \left[-\tfrac{\varepsilon}{2i}\Delta + \tfrac{1}{i\varepsilon}V\right]\phi_\varepsilon\right) + \left(a(x,\varepsilon D)\left[-\tfrac{\varepsilon}{2i}\Delta + \tfrac{1}{i\varepsilon}V\right]\phi_\varepsilon, \phi_\varepsilon\right)\right\} \\
&= \lim_{\varepsilon\to 0}\tfrac{1}{i\varepsilon}\left(\left[\tfrac{\varepsilon^2}{2}\Delta - V(x), a(x,\varepsilon D)\right]\phi_\varepsilon, \phi_\varepsilon\right) \\
&= -\lim_{\varepsilon\to 0}\tfrac{1}{i\varepsilon}\left(\left[\omega(x,\varepsilon D), a(x,\varepsilon D)\right]\phi_\varepsilon, \phi_\varepsilon\right) \\
&= -\lim_{\varepsilon\to 0}\tfrac{1}{i\varepsilon}\tfrac{\varepsilon}{i}\left(\left[(\nabla_k\omega\cdot\nabla_x a)(x,\varepsilon D) - (\nabla_x\omega\cdot\nabla_k a)(x,\varepsilon D)\right]\phi_\varepsilon, \phi_\varepsilon\right) \\
&= \langle\left[\nabla_k\omega\cdot\nabla_x a - (\nabla_x\omega\cdot\nabla_k a)\right], W\rangle \\
&= \langle k\cdot\nabla_x a - \nabla V\cdot\nabla_k a, W\rangle \\
&= \langle a, -k\cdot\nabla_x W + \nabla_x V\cdot\nabla_k W\rangle,
\end{aligned}$$

which is the weak form of (3.31).

Let us now compare the information one may obtain from the Liouville equation (3.31) to the standard geometric optics. First, we derive the eikonal and transport equations for the semiclassical Schrödinger equation (3.26). We consider initial data of the form

$$\phi^\varepsilon(0,x) = e^{iS_0(x)/\varepsilon}A_0(x) \tag{3.32}$$

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 (3.26) in the same form as the initial data (3.32), with an evolved phase and amplitude

$$\phi^\varepsilon(t,x) = e^{iS(t,x)/\varepsilon}(A(t,x) + \varepsilon A_1(t,x) + \dots). \tag{3.33}$$

Inserting this form into (3.26) and equating the powers of ε we get evolution equations for the phase and amplitude

$$S_t + \frac{1}{2}|\nabla S|^2 + V(x) = 0,\ \ S(0,x) = S_0(x) \tag{3.34}$$

and

$$(|A|^2)_t + \nabla\cdot(|A|^2\nabla S) = 0,\ \ |A(0,x)|^2 = |A_0(x)|^2. \tag{3.35}$$

The phase equation (3.34) is called the eikonal and the amplitude equation (3.35) 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 (3.31)? As we have computed before, for the WKB initial data (3.32) the initial Wigner distribution has the form

$$W_0(x,k) = |A_0(x)|^2\delta(k - \nabla S_0(x)). \tag{3.36}$$

As long as the geometric optics approximation (3.33) remains valid we expect the solution of the Liouville equation (3.31) to have the same form:

$$W(t,x,k) = |A(t,x)|^2\delta(k - \nabla S(t,x)). \tag{3.37}$$

We insert this ansatz into (3.31) :

$$\left(\frac{\partial}{\partial t}+k\cdot\nabla_x-\nabla V\cdot\nabla_k\right)\left(|A(t,x)|^2\delta(k-\nabla S(t,x))\right)=0. \tag{3.38}$$

or, equivalently,

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

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 (3.39) we obtain the transport equation (3.35) from the term in the first line, while the coefficient at D_m gives the eikonal equation (3.34) differentiated with respect to x_m. Expression (3.37) 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 at a caustic. Consider the Schrödinger equation (3.26) with $V=0$ – the corresponding Liouville equation is

$$W_t+k\cdot\nabla_x W=0,\ W(0,x,k)=W_0(x,k). \tag{3.39}$$

Its solution is $W(t,x,k)=W_0(x-kt,k)$ and clearly exists for all time. Consider 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 (3.39) 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 (3.34) 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

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

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 (3.36).

3.2.2. *Wigner transforms of mixtures of states.* We have noted that the $L^2(\mathbb{R}^d\times\mathbb{R}^d)$-norm of the Wigner transform blows up in the limit $\varepsilon\to 0$ unless the underlying family of functions ϕ_ε converges strongly to zero in $L^2(\mathbb{R}^d)$. On the other hand, the L^2-norm of the Wigner transforms for each $\varepsilon>0$ is preserved – it just so happens that it blows up in the limit. The L^2-norm is often much more convenient to use than the norm in $\mathcal{A}'$ and its conservation is typically an easy consequence of the evolution equation for the Wigner transform. For example, if ϕ_ε satisfy the Schrödinger equation

$$i\varepsilon\frac{\partial\phi_\varepsilon}{\partial t}+\frac{\varepsilon^2}{2}\Delta\phi_\varepsilon-V(x)\phi_\varepsilon=0, \tag{3.40}$$

then the Wigner transform W_ε satisfies

$$\frac{\partial W_\varepsilon}{\partial t}+k\cdot\nabla_x W_\varepsilon=\frac{1}{i\varepsilon}\int e^{ip\cdot x}\hat{V}(p)\left[W_\varepsilon\left(k-\frac{\varepsilon p}{2}\right)-W_\varepsilon\left(k+\frac{\varepsilon p}{2}\right)\right]\frac{dp}{(2\pi)^n}. \tag{3.41}$$

It is immediate to verify that (3.41) preserves the L^2-norm:

$$\frac{d}{dt}\int|W_\varepsilon(t,x,k)|^2dxdk=0.$$

It is much more difficult to verify that the $\mathcal{A}'$-norm of solutions does not grow. Therefore, it would be convenient to have a tool of working with the L^2-norm of the Wigner transform. This is what mixtures of state do. They arise, either naturally or artificially when families of solutions are considered rather than one solution. That is, we consider a measure $P(d\omega)$ on a state space Ω (which can be a probability space but needs not be) and introduce a family of initial data $\psi_0^\varepsilon(x,\omega)$ for the Schrödinger equation parametrized by $\omega\in\Omega$. Accordingly we may define a mixture of states (the terminology comes from the quantum mechanics)

$$\bar{W}_\varepsilon(t,x,k)=\int_\Omega W_\varepsilon(t,x,k,\omega)P(d\omega)$$

with

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

The point is that while the L^2-norm of $W_\varepsilon(t,x,k,\omega)$ blows up for each fixed state $\omega\in\Omega$, the L^2-norm of the average Wigner transform $\bar{W}_\varepsilon(t,x,k)$ may remain bounded. In particular, in the case of the Schrödinger equation, as $\bar{W}_\varepsilon$ satisfies (3.41), its L^2-norm is bounded as long as the L^2-norm of the initial data $\bar{W}_\varepsilon(0,x,k)$ is uniformly bounded. Let us give a couple of examples when this might happen. The first one arises when the initial data is random, and the second comes from the analysis of the time-reversal experiments that we will study in some detail later.

Statistical averaging: take the initial data for the Schrödinger equation of the form $\phi_0^\varepsilon(x;\omega)=\psi(x)V(x/\varepsilon;\omega)$, where $V(y;\omega)$ is a mean zero, scalar spatially homogeneous random process with a rapidly decaying two-point correlation function $R(z)$:

$$E\left\{V(y)V(y+z)\right\}=\int V(y;\zeta)V(y+z;\zeta)dP(\omega)=R(z)\in\mathcal{S}(\mathbb{R}^d),$$

and $\psi(x)\in C_c^\infty(\mathbb{R}^d)$. The "average" Wigner transform is then

$$\begin{aligned}\bar{W}_\varepsilon(x,k)&=\int_\Omega\left(\int e^{ik\cdot y}\phi_\varepsilon\left(x-\tfrac{\varepsilon y}{2},\omega\right)\bar{\phi}_\varepsilon\left(x-\tfrac{\varepsilon y}{2},\omega\right)\tfrac{dy}{(2\pi)^d}\right)dP(\omega)\\&=\int_\Omega\left(\int e^{ik\cdot y}\psi\left(x-\tfrac{\varepsilon y}{2}\right)\bar{\psi}\left(x-\tfrac{\varepsilon y}{2}\right)V\left(\tfrac{x}{\varepsilon}-\tfrac{y}{2},\omega\right)\right.\\&\quad\left.\times V\left(\tfrac{x}{\varepsilon}+\tfrac{y}{2},\omega\right)\tfrac{dy}{(2\pi)^d}\right)dP(\omega)\\&=\int e^{ik\cdot y}R(y)\psi\left(x-\tfrac{\varepsilon y}{2}\right)\bar{\psi}\left(x-\tfrac{\varepsilon y}{2}\right)\tfrac{dy}{(2\pi)^n}\to|\psi(x)|^2\hat{R}(k).\end{aligned}$$

Hence the limit Wigner distribution is given by $W(x,k) = |\psi(x)|^2 \hat{R}(k)$, where $\hat{R}(k)$ is the inverse Fourier transform of $R(y)$. In addition, convergence is strong in $L^2(\mathbb{R}^d \times \mathbb{R}^d)$:

$$\begin{aligned} \|\bar{W}_\varepsilon - W\|_{L^2}^2 &= \int |R(y)|^2 \left(\psi\left(x - \tfrac{\varepsilon y}{2}\right)\bar{\psi}\left(x - \tfrac{\varepsilon y}{2}\right) - |\psi(x)|^2\right)^2 \tfrac{dxdy}{(2\pi)^n} \\ &= \int I_\varepsilon(y)|R(y)|^2 \tfrac{dy}{(2\pi)^n} \end{aligned}$$

with

$$I_\varepsilon(y) = \int \left(\psi\left(x - \frac{\varepsilon y}{2}\right)\bar{\psi}\left(x - \frac{\varepsilon y}{2}\right) - |\psi(x)|^2\right)^2 dx.$$

However, we have $|I_\varepsilon(y)| \le 4\|\psi\|_{L^4}^4$ and

$$I_\varepsilon(y) = \int \left(\psi(x - \frac{\varepsilon y}{2})\psi^*(x + \frac{\varepsilon y}{2}) - |\psi(x)|^2\right)^2 dx \to 0$$

as $\varepsilon \to 0$ since $\psi \in C_c(\mathbb{R}^d)$, pointwise in y. Therefore $\|\bar{W}_\varepsilon - W\|_2 \to 0$ by the Lebesgue dominated convergence theorem.

Smoothing of oscillations: the initial data is of the form $\phi_0^\varepsilon(x;\zeta) = \psi(x)e^{i\zeta\cdot x/\varepsilon}$, where $\psi(x) \in C_c(\mathbb{R}^d)$. The state space $S = \mathbb{R}^d$, and the measure P is $P(d\omega) = g(\omega)d\omega$, $\omega \in \mathbb{R}^d$, and $g \in \mathcal{S}(\mathbb{R}^d)$. Then the limit Wigner distribution is $W(x,k) = |\psi(x)|^2 g(k)$ and convergence of $\bar{W}_\varepsilon(x,k)$ to the limit is strong in $L^2(\mathbb{R}^d \times \mathbb{R}^d)$. This is verified exactly in the same way as in the previous example.

3.3. **The high frequency limit for symmetric hyperbolic systems.**

3.3.1. *Matrix-valued Wigner transform.* The definition of the Wigner transform may be generalized in a straightforward manner for families of vector-valued functions $u_\varepsilon(x) \in L^2(\mathbb{R}^n; \mathbb{C}^m)$. The Wigner transform is then an $m \times m$ matrix

$$W_\varepsilon(x,k) = \int e^{ik\cdot y} u_\varepsilon\left(x - \frac{\varepsilon y}{2}\right) u_\varepsilon^*\left(x + \frac{\varepsilon y}{2}\right)\frac{dy}{(2\pi)^d}. \tag{3.42}$$

Here we denote by u^* the conjugate-transpose of the vector u. The basic properties of the scalar Wigner transform can be immediately generalized to the matrix case. In particular, $W_\varepsilon(x,k)$ is a self-adjoint matrix, and we have the following:

Theorem 3.13. *Let the family of vector-valued functions $u_\varepsilon(x)$ be uniformly bounded in $L^2(\mathbb{R}^n;\mathbb{C}^m)$. Then the matrix-valued Wigner transform W_ε converges weakly along a subsequence $\varepsilon_k \to 0$ to a matrix-valued distribution $W(x,k) \in \mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{C}^m \times \mathbb{C}^m)$. Any such limit point $W(x,k)$ is a non-negative matrix for each (x,k).*

The proof is identical to that of Theorem 3.2 and uses the fact that under the assumptions of this theorem the family of Wigner transforms $W_\varepsilon(x,k)$ is uniformly bounded in $(\mathcal{A}^{m\times m})'$, the dual space of

$$\mathcal{A}^{m\times m} = \{\lambda(x,k) : \int \sup_x \|\tilde{\lambda}(x,y)\| dy\}$$

where $\lambda(x,k)$ is a vector-valued function, and $\|X\| = \mathrm{Tr}(XX^*) = \left(\sum_{j,k=1}^m |x_{jk}|^2\right)^{1/2}$. We may also define the semiclassical operators $a^w(x,\varepsilon D)$ and $a(x,\varepsilon D)$ in the Weyl and standard quantizations, respectively, by expressions (3.9) and (3.10) with a matrix valued function $a(x,k)$ and a vector valued function $f(x)$. Then Propositions 3.3, 3.4, 3.5, 3.6 all apply verbatim and we do not restate them here. Note, however, that in the product Lemma 3.5 the operators should be applied in the correct order as they are matrix valued and their symbols do not commute. The localization, orthogonality and "energy capturing" Propositions 3.7, 3.8 and 3.11 also hold.

In the same spirit we may define a "cross"-Wigner transform for a pair of (vector-valued) functions $u_\varepsilon(x)$ and $v_\varepsilon(x)$ as

$$W_\varepsilon[u_\varepsilon, v_\varepsilon] = \int e^{ik\cdot y} u_\varepsilon\left(x - \frac{\varepsilon y}{2}\right) v_\varepsilon^*\left(x + \frac{\varepsilon y}{2}\right) \frac{dy}{(2\pi)^d}. \tag{3.43}$$

We will denote the corresponding limit by $W[u^\varepsilon, v^\varepsilon]$. It has the property that

$$\langle W[u^\varepsilon, v^\varepsilon], a\rangle = \sum_{i,j} \langle W[u_i^\varepsilon, v_j^\varepsilon], a\rangle = \lim_{\varepsilon\to 0} \sum_{i,j} \langle \bar{v}_j^\varepsilon, a^w(x, \varepsilon D) u_i^\varepsilon\rangle.$$

A useful fact that follows immediately from the aforementioned properties of the Wigner transforms is that for a symbol $P(x,k)$ we have for the limit Wigner matrix

$$\begin{aligned} W[P(x,\varepsilon D)u_\varepsilon, v_\varepsilon](x,k) &= P(x,k) W[u_\varepsilon, v_\varepsilon](x,k), \\ W[u_\varepsilon, P(x,\varepsilon D)u_\varepsilon](x,k) &= W[u_\varepsilon, v_\varepsilon](x,k) P^*(x,k). \end{aligned} \tag{3.44}$$

3.3.2. *The evolution of the Wigner transform: constant coefficients.* We now address consider the evolution of the Wigner transform for general equations other than the linear Schrödinger equation. We begin with systems of equations with constant coefficients of the form

$$\begin{aligned} &\varepsilon \frac{\partial u^\varepsilon}{\partial t} + P(\varepsilon D) u^\varepsilon = 0 \\ &u^\varepsilon(t=0) = u_0^\varepsilon \end{aligned} \tag{3.45}$$

with u^ε being a $\mathbb{C}^m$-valued vector function. A typical example we have in mind is a symmetric hyperbolic system

$$\frac{\partial u}{\partial t} + D^j \frac{\partial u}{\partial x_j} = 0$$

with symmetric matrices D^j, $j = 1, \dots, n$ – in that case $P(k) = ik_j D^j$. In general, the operator $P(\varepsilon D)$ is associated with a multiplier $P(k)$. We assume that $P \in C^\infty(\mathbb{R}^d \backslash \{0\})$ and $P^*(k) = -P(k)$. It follows that the total energy is conserved:

$$N(t) = \int n^\varepsilon(t,x) dx = \int n_0^\varepsilon(x) dx = N(0).$$

Here $n_\varepsilon(t,x) = |u^\varepsilon(t,x)|^2$ is the energy density and $n_0^\varepsilon(x)$ its initial value. Therefore, it makes sense to consider the Wigner transform of solutions and their weak limits.

We impose the following conditions on the symbol: all eigenvalues $\omega_\alpha(k)$ of the self-adjoint matrix $iP(k)$ may be ordered as

$$\omega_1(k) < \cdots < \omega_p(k)$$

with the multiplicities r_α independent of k, for $k \neq 0$. We denote by $\Pi_\alpha(k)$ the orthogonal projection onto the eigenspace corresponding to $\omega_\alpha(k)$ and assume that $\omega_\alpha(k)$ and $\Pi_\alpha(k)$ are smooth functions of k away from $k = 0$. Under these assumptions the evolution of the Wigner matrix is described by the following theorem.

Theorem 3.14. *Let the initial data $u_0^\varepsilon(x)$ for (3.45) be a pure family, uniformly bounded in $L^2(\mathbb{R}^n)$, ε-oscillatory and compact at infinity with the unique limit Wigner matrix measure $W_0(x,k)$. Assume that $\hat{u}_0^\varepsilon(k)$ vanishes for $|k| \le r$ for some*

$r > 0$. *Then the Wigner transform* $W_\varepsilon(t,x,k)$ *converges weakly in* $\mathcal{S}'(\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d)$ *to*

$$W(t,x,k) = \sum_{\alpha=1}^{p} W_\alpha(t,x,k).$$

The matrices $W_\alpha(t,x,k)$ *satisfy the Liouville equations*

$$\frac{\partial W_\alpha}{\partial t} + \nabla_k \omega_\alpha(k) \cdot \nabla_x W_\alpha = 0,\ W_\alpha(0,x,k) = \Pi_\alpha(k) W_0(x,k) \Pi_\alpha(k). \tag{3.46}$$

We can make the above statement somewhat more explicit if we introduce the orthonormal basis of eigenvectors b_α^m that correspond the eigenvalue ω_α with $m = 1,\dots,r_\alpha$ – here r_α is the multiplicity of ω_α. Then the matrix Π_α has the form

$$\Pi_\alpha = \sum_{m,l=1}^{r_\alpha} b_\alpha^m(k) b_\alpha^{l*}(k).$$

The matrix W_α may, in turn, be written as

$$W_\alpha(t,x,k) = \sum_{m,l} w_{ml}^\alpha(t,x,k) b_\alpha^m(k) b_\alpha^{l*}(k).$$

The entries w_{ml}^α are given by

$$w_{ml}^\alpha = \mathrm{Tr}\left[W_\alpha b_\alpha^l b_\alpha^{m*}\right] = \mathrm{Tr}\left[\Pi_\alpha W \Pi_\alpha b_\alpha^l b_\alpha^{m*}\right] = \mathrm{Tr}\left[W b_\alpha^l b_\alpha^{m*}\right].$$

Let us organize the elements w_{mn}^α into an $r_\alpha \times r_\alpha$ matrix w^α. When $r_\alpha > 1$ the mode α is polarized and w^α is called the coherence matrix of this mode. Each coherence matrix satisfies the Liouville equation

$$\frac{\partial w^\alpha}{\partial t} + \nabla_k \omega_\alpha \cdot \nabla_x w^\alpha = 0,\ w_{mn}^\alpha(0,x,k) = \mathrm{Tr}\left[W_0(x,k)) b_\alpha^n(k) b_\alpha^{m*}(k)\right]. \tag{3.47}$$

This form of the Liouille equation is more convenient in the derivation of the radiative transport equations in random media that we will consider later.

Energy propagation for solutions of (3.45) is described by the following theorem.

Theorem 3.15. *Under the same assumptions the energy density* $n^\varepsilon(t,x)$ *converges weakly (for each time* $t \geq 0$*) to the measure* $n^0(t,x)$ *given by*

$$n^0(t,x) = \sum_{\alpha=1}^{p} \int w_\alpha^0(x - t\nabla\omega_\alpha(k), dk). \tag{3.48}$$

Here $w_\alpha^0(x,k) = Tr(\Pi_\alpha W \Pi_\alpha)(x,k)$. *Moreover, convergence is uniform on finite time intervals.*

The reason why we do not have uniform in time convergence of the matrix Wigner transform but do have it for the energy density lies in the cross-mode terms $\Pi_\alpha W_\varepsilon \Pi_\beta$ with $\alpha \neq \beta$ – they have fast temporal oscillations but do not necessarily go to zero uniformly in time. For example, consider a special solution of (3.45) which is a sum of two plane waves with the same wave vector:

$$u_\varepsilon(x) = A_\alpha b_\alpha(k_0) e^{ik_0\cdot x/\varepsilon - i\omega_\alpha(k_0)t/\varepsilon} + A_\beta b_\beta(k_0) e^{ik_0\cdot x/\varepsilon - i\omega_\beta(k_0)t/\varepsilon}$$

with $\omega_\alpha(k_0) \neq \omega_\beta(k_0)$. Then the matrix Wigner transform is

$$\begin{aligned} W_\varepsilon(t,x,k) = \Big[& |A_\alpha|^2 a b_\alpha(k_0) a b_\alpha^*(k_0) + |A_\beta|^2 a b_\beta(k_0) a b_\beta^*(k_0) \\ & + A_\alpha \bar{A}_\beta b_\beta(k_0) a b_\alpha^*(k_0) e^{i(\omega_\beta(k_0) - \omega_\alpha(k_0))t/\varepsilon} \\ & + \bar{A}_\alpha A_\beta b_\alpha(k_0) a b_\beta^*(k_0) e^{i(\omega_\alpha(k_0) - \omega_\beta(k_0))t/\varepsilon} \Big] \delta(k - k_0). \end{aligned}$$

The cross-terms are oscillating rapidly in time – hence they vanish as $\varepsilon \to 0$ but only in the weak sense. On the other hand, these terms have zero energy – their trace vanishes. Therefore, the energy does not have these temporally oscillating terms – this simple example captures the basic phenomenon that the cross-mode terms are oscillatory in time but carry no energy.

3.3.3. *The evolution of the Wigner transform: slowly varying coefficients.* We now consider the Wigner transforms of solutions of symmetric hyperbolic systems of the form

$$\frac{\partial u_\varepsilon}{\partial t} + B(x)D^j \frac{\partial}{\partial x_j}\left(B(x)u_\varepsilon\right) = 0. \tag{3.49}$$

The matrix $B(x)$ is positive-definite and the constant matrices D^j are symmetric and independent of t and x. The total energy

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

is conserved:

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

with the energy density $E(t,x) = |u(t,x)|^2$ and the flux $F_j(t,x) = (D^j Bu, Bu)$. We will assume in this section, as usually, that away from $k = 0$ the dispersion matrix $L(x,k) = B(x)k_j D^j B(x)$ has eigenvalues $\omega_\alpha(x,k)$ with constant multiplicity r_α independent of x and $k \neq 0$, and both ω_α and the corresponding eigenvectors b^i_α, $i = 1,\dots,r_\alpha$ are smooth functions of $x \in \mathbb{R}^d$ and $k \in \mathbb{R}^d\backslash\{0\}$.

Energy conservation allows us to talk about the matrix Wigner transforms of the solutions and study their limits. We take a matrix-valued test function $a(t,x,k) \in \mathcal{S}(\mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n)$ and compute

$$\varepsilon\langle a, \frac{\partial W_\varepsilon}{\partial t}\rangle + \langle a(x,\varepsilon D)u_\varepsilon, BP(\varepsilon D)[Bu_\varepsilon]\rangle + \langle a(x,\varepsilon D)\left(BP(\varepsilon D)[Bu_\varepsilon]\right), u_\varepsilon\rangle = 0. \tag{3.50}$$

Here we have denoted $P(k) = ik_j D_j$ and B is multiplication by the function $B(x)$. The operator $Mf = B(x)P(\varepsilon D)[B(x)f]$ is skew-symmetric. Hence, we may re-write (3.50) as

$$\varepsilon\langle a, \frac{\partial W_\varepsilon}{\partial t}\rangle - \langle [BP(\varepsilon D)(Ba(t,x,\varepsilon D)u_\varepsilon), u_\varepsilon\rangle + \langle a(t,x,\varepsilon D)\left(BP(\varepsilon D)[Bu_\varepsilon]\right), u_\varepsilon\rangle = 0. \tag{3.51}$$

Integrating in time and passing to the limit $\varepsilon \to 0$ using the uniform in time a priori bounds for $W_\varepsilon(t,x,k)$ in $\mathcal{A}'_{m\times m}$ we see that the limit Wigner matrix $W(t,x,k)$ satisfies

$$\mathrm{Tr}\left[B(x)P(k)B(x)a(t,x,k)W(t,x,k) - a(t,x,k)B(x)P(x,k)B(x)W(t,x,k)\right] = 0. \tag{3.52}$$

Therefore, as in the constant coefficient case, the matrix $W(x,k)$ satisfies

$$L(x,k)W(t,x,k) = W(t,x,k)L(x,k),\ L(x,k) = \frac{1}{i}B(x)P(x,k)B(x), \tag{3.53}$$

It follows that $\Pi_\alpha(x,k)W(t,x,k)\Pi_\beta(t,x,k) = 0$ for $\alpha \neq \beta$ – here $\Pi_\alpha(x,k)$ is the projection matrix on the eigenspace of the matrix $L(x,k)$ corresponding to an eigenvalue $\omega_\alpha(x,k)$. Thus, the limit Wigner matrix has a representation

$$W(t,x,k) = \sum_\alpha \Pi_\alpha(x,k)W(t,x,k)\Pi_\alpha(x,k). \tag{3.54}$$

We may also write it in a more explicit form as

$$W(t,x,k) = \sum_{\alpha} \sum_{i,j=1}^{r_\alpha} w_\alpha^{ij}(t,x,k) b_\alpha^i(x,k) b_\alpha^{j*}(x,k). \tag{3.55}$$

The vectors b_α^i form the orthonormal basis of the eigenspace corresponding to the eigenvalue ω_α. The limit energy density is simply

$$E(t,x) = \sum_{\alpha} \int \mathrm{Tr} w_\alpha(t,x,k) dk$$

for ε-oscillatory and compact at infinity families of solutions – we will see that this property is preserved by evolution.

Somewhat lengthy computations (using the Weyl operator calculus and appropriate test functions) lead to the following evolution equations for the $r_\alpha \times r_\alpha$ coherence matrices w_α:

$$\frac{\partial w_\alpha}{\partial t} + \nabla_k \omega_\alpha \cdot \nabla_x w_\alpha - \nabla_x \omega_\alpha \cdot \nabla_k w_\alpha + [\bar{N}^\alpha, w^\alpha] = 0. \tag{3.56}$$

The $r_\alpha \times r_\alpha$ matrices $\bar{N}_\alpha$ have entries

$$\bar{N}_\alpha^{ni} = \frac{1}{2}\left[((B\nabla_k P \cdot \nabla_x B) b_\alpha^i, b_\alpha^n) - (b_\alpha^i, (B\nabla_k P \cdot \nabla_x B) b_\alpha^n)\right]. \tag{3.57}$$

The matrix $\bar{N}^\alpha$ is skew-symmetric and vanishes when the eigenvalue ω_α is simple. This result can be summarized in the following theorem.

Theorem 3.16. *Let $u_\varepsilon(t,x)$ be the solution of the initial value problem*

$$\frac{\partial u_\varepsilon}{\partial t} + B(x) D^j \frac{\partial}{\partial x_j}(B(x) u_\varepsilon) = 0 \tag{3.58}$$

with an ε-oscillatory and compact at infinity pure family of initial data $u_\varepsilon(0,x) = u_\varepsilon^0(x)$. The coefficient matrices $B(x)$ are symmetric positive-definite and D^j are independent of t and x. Then the Wigner transforms $W_\varepsilon(t,x,k)$ converge weakly in $\mathcal{S}'(\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d)$ to the matrix distribution

$$W(t,x,k) = \sum_{\alpha=1}^{p} \sum_{i,j=1}^{r_\alpha} w_\alpha^{ij}(t,x,k) b_\alpha^i(x,k) b_\alpha^j(x,k).$$

The $r_\alpha \times r_\alpha$ coherence matrices w_α satisfy the matrix Liouville equations (3.56) with the initial data $w_\alpha^{mn}(0,x,k) = Tr[W_0(x,k) b_\alpha^n(x,k) b_\alpha^{m*}(x,k)]$. Here $W_0(x,k)$ is the Wigner transform of the family $u_\varepsilon^0(x)$.

As an immediate consequence, the energy of the mode α, which is $\bar{w}_\alpha = Tr(w_\alpha)$ satisfies a scalar Liouville equation

$$\frac{\partial \bar{w}_\alpha}{\partial t} + \nabla_k \omega_\alpha \cdot \nabla_x \bar{w}_\alpha - \nabla_x \omega_\alpha \cdot \nabla_k \bar{w}_\alpha = 0. \tag{3.59}$$

A few comments on the matrix Liouville equations (3.56) are in order. First of all, the coupling matrix N^α vanishes if the coefficient matrix B is independent of x – this is seen from its explicit form. Furthermore, as in the constant coefficients case equations for various modes are all decoupled. This means that slow variations (relative to the wave length) of the background material properties do not induce mode coupling in the leading order. They do, however, suffice to couple various polarizations corresponding to the same mode if the mode is polarized. Still the "coupling" commutator term in the Liouville equations may be eliminated by an appropriate choice of the basis. More precisely, if we take $w_\alpha = U\bar{w}_\alpha U^*$ with the matrix U that is a solution of the evolution equation

$$\frac{\partial U}{\partial t} + \{\omega_\alpha, U\} + \bar{N}^\alpha U = 0,\ U(0,x,k) = I,$$

then the matrix $\bar{w}_\alpha$ satisfies a Liouville equation without the commutator term

$$\frac{\partial \bar{w}_\alpha}{\partial t} + \nabla_k\omega_\alpha \cdot \nabla_x \bar{w}_\alpha - \nabla_x\omega_\alpha \cdot \nabla_k \bar{w}_\alpha = 0,\ \bar{w}_\alpha(0,x,k) = w_\alpha^0(x,k). \tag{3.60}$$

This means that the matrix $U(t,x,k)$ describes the rotation (recall that the matrix N^α is skew-symmetric) of the polarization vector along the bicharacteristics.

As in the case of constant coefficients, the non-uniform in time convergence of the matrix Wigner transform to the limit in Theorem 3.16 is not an artifact of the proof. However, the phase space energy density, that is, the trace of the Wigner matrix converges to its limit $\bar{E}(t,x) = \sum_\alpha \int \mathrm{Tr} w^\alpha(t,x,k)dk$ uniformly in time (and weakly in space). This is because (as one can see from (3.50)) the time derivative $\partial W_\varepsilon/\partial t$ is uniformly bounded in time.

The limit Liouville equations preserve the total energy $\bar{E}(t,x)$ defined above. Therefore, as long as the initial data is ε-oscillatory and compact at infinity, convergence of the trace of the Wigner matrix is tight for all $t \geq 0$. As a consequence, using Theorem 3.11 we conclude that the family of solutions of (3.49) remain ε-oscillatory and compact at infinity.

3.4. High frequency Wigner limits: examples.

3.4.1. *High Frequency Approximation for Acoustic Waves.* We will now apply the results of the previous section to acoustic waves. We will also review the usual form of the high frequency approximation and make explicit the relation between the phase space form of the high frequency approximation and the usual one.

The acoustic equations for the velocity and pressure disturbances u and p are

$$\begin{aligned} \rho\frac{\partial u}{\partial t} + \nabla p &= 0 \\ \kappa\frac{\partial p}{\partial t} + \mathrm{div} u &= 0. \end{aligned} \tag{3.61}$$

Here $\rho = \rho(x)$ is the medium density and $\kappa = \kappa(x)$ is its compressibility. Equations (3.61) can be re-written in terms of $v(t,x) = \sqrt{\rho(x)}u(t,x)$ and $q(t,x) = \sqrt{\kappa(x)}p(t,x)$ as

$$\begin{aligned} &\frac{\partial v}{\partial t} + \frac{1}{\sqrt{\rho}}\nabla\left[\frac{1}{\sqrt{\kappa}}q\right] = 0 \\ &\frac{\partial q}{\partial t} + \frac{1}{\sqrt{\kappa(x)}}\mathrm{div}\left[\frac{1}{\sqrt{\rho}}v\right] = 0. \end{aligned} \tag{3.62}$$

The energy density and flux for acoustic waves are given by

$$\mathcal{E}(t,x) = \frac{1}{2}|v(t,x)|^2 + \frac{1}{2}q^2(t,x),\ \mathcal{F}(t,x) = c(x)q(t,x)v(t,x). \tag{3.63}$$

Equations (3.62) have the form (3.49) with the matrix

$$B(x) = \mathrm{diag}\left[\frac{1}{\sqrt{\rho(x)}}, \frac{1}{\sqrt{\rho(x)}}, \frac{1}{\sqrt{\rho(x)}}, \frac{1}{\sqrt{\kappa(x)}}\right]$$

while each of the matrices $D^i = e_1 e_4^* + e_4 e_1^*$ has all zero entries except for D^i_{i4} and D^i_{4i} which are equal to one. For instance, the matrix D^1 is

$$D^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Then the dispersion matrix $L(x,k)$ has the form

$$L = v(x) \begin{pmatrix} 0 & 0 & 0 & k_1 \\ 0 & 0 & 0 & k_2 \\ 0 & 0 & 0 & k_3 \\ k_1 & k_2 & k_3 & 0 \end{pmatrix} \tag{3.64}$$

with the sound speed $c(x) = 1/\sqrt{\kappa(x)\rho(x)}$. It has one double eigenvalue $\omega_1 = \omega_2 = 0$ and two simple eigenvalues $\omega_\pm = \pm c(x)|k|$. The corresponding orthonormal basis of eigenvectors is

$$b^1 = (z^{(1)}(k), 0),\ b^2 = (z^{(2)}(k), 0),\ b^\pm = \left(\frac{\hat{k}}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right), \tag{3.65}$$

with the vectors $\hat{k}$, $z^{(1)}(k)$ and $z^{(2)}(k)$, which form an orthonormal triplet:

$$\hat{k} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}, z^{(1)} = \begin{pmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{pmatrix}, z^{(2)} = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix}. \tag{3.66}$$

The limit Wigner matrix of the family $\mathbf{v}_\varepsilon = (v_\varepsilon, q_\varepsilon)$, according to (3.55) can be represented as

$$\begin{aligned} W(t,x,k) = & \sum_{i,j=1}^{2} w_0^{ij}(t,x,k) b^i(k) b^{j*}(k) \\ & + w_+(t,x,k) b^+(k) b^{+*}(k) + w_-(t,x,k) b^-(k) b^{-*}(k). \end{aligned} \tag{3.67}$$

In order to understand better the physical meaning of these modes let us write the vector $v_\varepsilon(x)$ as a sum $v_\varepsilon(t,x) = v^\varepsilon_{in}(t,x) + v^\varepsilon_{irr}$ with an incompressible field $v^\varepsilon_{in}(t,x)$: $\nabla \cdot v^\varepsilon_{in} = 0$ and an irrotational component v^ε_{irr}: $\nabla \times v^\varepsilon_{irr} = 0$. The limit Wigner matrices W_{in} and W_{irr} of the families $v^\varepsilon_{in}(t,x)$ and $v^\varepsilon_{irr}(t,x)$ satisfy $W_{in}(t,x,k)k = 0$ and $W_{irr}(t,x,k)z = 0$ for any vector z orthogonal to k. Decomposition (3.67) tells us that $W = W_{irr} + W_{in}$ with

$$\begin{aligned} W_{in} &= \sum_{i,j=1}^{2} w_0^{ij}(t,x,k) b^i(k) b^{j*}(k), \\ W_{irr} &= w_+(t,x,k) b^+(k) b^{+*}(k) + w_-(t,x,k) b^-(k) b^{-*}(k). \end{aligned}$$

Therefore, the eigenvectors $b^1(k)$ and $b^2(k)$ correspond to transverse advection modes, orthogonal to the direction of propagation. These modes do not propagate because $\omega_{1,2} = 0$: equation (3.56) for the coherence matrix w_0 is of the form $\dfrac{\partial w_0}{\partial t} = 0$ – hence $w_0(t,x,k) = 0$ if it is zero initially. This is the case when the initial data is irrotational. The eigenvectors $b^+(k)$ and $b^-(k)$ represent acoustic waves, which are longitudinal, and which propagate with the sound speed $c(x)$: the scalar amplitudes $w_\pm(t,x,k)$ satisfy the scalar Liouville equations

$$\frac{\partial w_\pm}{\partial t} \pm c(x)\hat{k}\cdot\nabla_x w_\pm \mp |k|\nabla_x c(x)\cdot\nabla_k w_\pm = 0. \tag{3.68}$$

Next, as we did for the Schrödinger equation, we establish the connection with the usual high frequency approximation for acoustic waves. We consider acoustic equations (3.62) with initial data of the form

$$\mathbf{v}(0,x) = \mathbf{v}_0(x)e^{iS_0(x)/\varepsilon},\ \ \mathbf{v} = (v,q) \tag{3.69}$$

where S_0 is the real valued initial phase function. We look for a solution in the form

$$\mathbf{v}(t,x) = (\underline{A}_0(t,x) + \varepsilon\underline{A}_1 + \dots)e^{iS(t,x)/\varepsilon}, \tag{3.70}$$

where $\underline{A}_0 = (v_0, q_0)$. We insert (3.70) into (3.62) to get in the leading order in ε

$$\begin{pmatrix} S_t & c(x)\nabla S \\ c(x)\nabla S\cdot & S_t \end{pmatrix}\begin{pmatrix} v_0 \\ q_0 \end{pmatrix} = 0. \tag{3.71}$$

The next term in the expansion yields

$$-i\begin{pmatrix} S_t & c(x)\nabla S \\ c(x)\nabla S\cdot & S_t \end{pmatrix}\begin{pmatrix} v_1 \\ q_1 \end{pmatrix} = \begin{pmatrix} \partial_t v_0 + \dfrac{1}{\sqrt{\rho}}\nabla\left[\dfrac{1}{\sqrt{\kappa}}q_0\right] \\ \partial_t q_0 + \dfrac{1}{\sqrt{\kappa}}\nabla\left[\dfrac{1}{\sqrt{\rho}}\cdot v_0\right] \end{pmatrix}. \tag{3.72}$$

Equation (3.71) gives the eikonal equation for the phase S

$$S_t^2 - c^2(x)(\nabla S)^2 = 0. \tag{3.73}$$

Then assuming that $S_t = +c(x)|\nabla S|$ we have

$$\begin{pmatrix} v_0 \\ q_0 \end{pmatrix} = \mathcal{A}(x)\mathbf{b}^+(\nabla S(t,x)), \tag{3.74}$$

where $\mathbf{b}^+$ is given by (3.65). The amplitude $\mathcal{A}(t,x)$ is determined by the solvability condition for (3.72), which gives the transport equation

$$\frac{\partial}{\partial t}|\mathcal{A}|^2 + \nabla\cdot\left(|\mathcal{A}|^2 c(x)\frac{\nabla S}{|\nabla S|}\right) = 0. \tag{3.75}$$

The eikonal and transport equations (3.73) and (3.75) can also be derived from the Liouville equation (3.68) as we did for the Schrödinger equation. In the high frequency limit, initial conditions of the form (3.69) imply that

$$w_+(0,x,k) = |\mathcal{A}_0(x)|^2\delta(k - \nabla S_0(x)). \tag{3.76}$$

Let the functions $S(t,x)$ and $|\mathcal{A}(t,x)|^2$ be the solutions of the eikonal and transport equations (3.73) and (3.75), respectively, with the initial conditions $S(0,x) = S_0(x)$ and $|\mathcal{A}(0,x)|^2 = |\mathcal{A}_0(x)|^2$. Then the solution of equation (3.68) is

$$w_+(t,x,k) = |\mathcal{A}(t,x)|^2\delta(k - \nabla S(t,x)). \tag{3.77}$$

Conversely, given initial conditions of the form (3.76) for (3.68) and w_+ given by (3.77), then S and $\mathcal{A}$ must satisfy the eikonal and transport equations (3.73) and (3.75), respectively. This is because the eikonal equation follows by integrating (3.68) with respect to k while the transport equation follows by multiplying it by k and then integrating with respect to k.

3.4.2. *Phase space geometric optics for electromagnetic waves.* Maxwell's equations in an isotropic medium and in suitable units are

$$\begin{aligned}\frac{\partial E}{\partial t} &= \frac{1}{\epsilon}\mathrm{curl}H \\ \frac{\partial H}{\partial t} &= -\frac{1}{\mu}\mathrm{curl}E\end{aligned} \tag{3.78}$$

where the dielectric permittivity is $\epsilon(x)$ and the relative magnetic permeability is $\mu(x)$. In this section as well as in other instances when we consider electromagnetic waves ϵ denotes the dielectric permittivity while the small parameter is denoted by ε. It follows from Maxwell's equations that if at the initial time we have

$$\mathrm{div}(\epsilon E) = \mathrm{div}(\mu H) = 0 \tag{3.79}$$

then these conditions hold for all time. We will always assume that (3.79) holds.

As a symmetric hyperbolic system Maxwell's equations can be written as

$$\frac{\partial}{\partial t}\begin{pmatrix}\bar{E}\\ \bar{H}\end{pmatrix} + \begin{pmatrix}\frac{1}{\sqrt{\epsilon}} & 0 \\ 0 & \frac{1}{\sqrt{\mu}} & 0\end{pmatrix}\begin{pmatrix}0 & -\nabla\times \\ \nabla\times & 0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{\epsilon}} & 0 \\ 0 & \frac{1}{\sqrt{\mu}}\end{pmatrix}\begin{pmatrix}\bar{E}\\ \bar{H}\end{pmatrix} = 0 \tag{3.80}$$

with $\bar{E} = \sqrt{\epsilon}E$ and $\bar{H} = \sqrt{\mu}H$. The 6×6 dispersion matrix L is

$$L = -c(x)\begin{pmatrix} 0 & 0 & 0 & 0 & -k_3 & k_2 \\ 0 & 0 & 0 & k_3 & 0 & -k_1 \\ 0 & 0 & 0 & -k_2 & k_1 & 0 \\ 0 & k_3 & -k_2 & 0 & 0 & 0 \\ -k_3 & 0 & k_1 & 0 & 0 & 0 \\ k_2 & -k_1 & 0 & 0 & 0 & 0 \end{pmatrix} = c(x)\begin{pmatrix} 0 & -T \\ T & 0\end{pmatrix} \tag{3.81}$$

with the speed of light $c(x) = 1/\sqrt{\epsilon(x)\mu(x)}$ and the matrix $T(k)$ defined by $T(k)p = k\times p$ or

$$T(k) = \begin{pmatrix} 0 & -k_3 & k_2 \\ k_3 & 0 & -k_1 \\ -k_2 & k_1 & 0\end{pmatrix}. \tag{3.82}$$

The dispersion matrix L has three eigenvalues, each with multiplicity two. They are $\omega_0 = 0$, $\omega_+ = c|k|$, $\omega_- = -c|k|$. The basis formed by the corresponding eigenvectors is

$$\begin{aligned} & b^{(01)} = (\hat{k},0),\ b^{(02)} = (0,\hat{k}), \\ & b^{(+,1)} = (\frac{z^{(1)}}{\sqrt{2}},\frac{z^{(2)}}{\sqrt{2}}),\ b^{(+,2)} = (\frac{z^{2}}{\sqrt{2}},-\frac{z^{(1)}}{\sqrt{2}}), \\ & b^{(-,1)} = (\frac{z^{(1)}}{\sqrt{2}},-\frac{z^{(2)}}{\sqrt{2}}),\ b^{(-,2)} = (\frac{z^{(2)}}{\sqrt{2}},\frac{z^{(1)}}{\sqrt{2}}), \end{aligned} \tag{3.83}$$

where the vectors k, $z^{(1)}(k)$ and $z^{(2)}(k)$ form an orthonormal triplet (3.66). The coherence matrix w_0 corresponding to the mode $\omega_0 = 0$ vanishes if (3.79) holds – this is checked in the same way as the absence of the vortical modes for the acoustic waves. The other eigenvectors correspond to transverse modes propagating with the speed $c(x)$. As in the acoustic case, we need only consider the eigenspace corresponding to ω_+. The 2×2 coherence matrices $w_\pm$ satisfy the Liouville equations (3.56), for instance, the evolution equation for $w = W_+$ is

$$\frac{\partial w}{\partial t} + c(x)\hat{k}\cdot\nabla_x w - |k|\nabla_x c(x)\cdot\nabla_k w + \bar{N}w - w\bar{N} = 0. \tag{3.84}$$

The 2×2 skew symmetric coupling matrix $N_+(x,k)$ is determined by its non-zero element

$$\bar{N}_+^{12} = \frac{1}{2}\left[((B\nabla_k P\cdot\nabla_x B)b^{+,2}, b^{+,1}) - (b^{+,2}, (B\nabla_k P\cdot\nabla_x B)b^{+,1})\right]. \tag{3.85}$$

The coherence matrix $W^+(t,x,k)$ is related to the four Stokes parameters [31], which are commonly used for the description of polarized light because they are directly measurable. Let l and r be two directions orthogonal to the direction of propagation and let $I = I_l + I_r$ be the the total intensity of light, with I_l and I_r denoting the intensities in the directions l and r, respectively. Let $Q = I_l - I_r$ be the difference between the two intensities. Also let $U = 2 < E_l E_r\cos\delta >$ and $V = 2 < E_l E_r \sin\delta >$ denote the intensity coherence, with fixed phase shift δ, between the amplitude of light in the directions l and r, respectively. Light is unpolarized if $U = V = Q = 0$. If the directions l and r are chosen to be $z^{(1)}(k)$ and $z^{(2)}(k)$, given by (3.66), then the coherence matrix $w^+(t,x,k)$ is related to the Stokes parameters (I,Q,U,V) by

$$w^+(t,x,k) = \frac{1}{2}\begin{pmatrix} I+Q & U+iV \\ U-iV & I-Q\end{pmatrix}. \tag{3.86}$$

When light is unpolarized, then the coherence matrix w^+ is proportional to the 2×2 identity matrix I. We will later see that in a random medium after a long propagation time, indeed, w^+ becomes nearly proportional to identity.

4. **Kinetic limits for the Liouville equations.**

4.1. **The Fokker-Planck limit.**

4.1.1. *The Fokker-Planck equation.* As we have seen in Section 3, the Wigner measure of the oscillatory solutions of the Schrödinger equations with a slowly (relative to the scale of wave oscillations) varying potential satisfies the Liouville equation

$$W_t + k\cdot\nabla_x W - \nabla V(x)\cdot\nabla_k W = 0. \tag{4.1}$$

More generally, the energy of each mode of a high frequency wave satisfies in the limit a Liouville equation (3.59):

$$W_t + \nabla_k\omega\cdot\nabla_x W - \nabla_x\omega\cdot\nabla_k W = 0. \tag{4.2}$$

Here $\omega(x,k)$ is the dispersion law of the mode. Equation (4.1) is a particular example of (4.2) with $\omega(x,k) = k^2/2 + V(x)$. Therefore, in order to understand how high frequency waves scatter in a random medium with random fluctuations that vary on a scale much larger than the wave length, one should study solutions of the Liouville equation with a random dispersion relation. The Liouville problem in a random medium is simpler to analyze than the full wave problem because it has an interpretation in terms of the trajectories of the Hamiltonian system

$$\frac{dX}{dt} = \nabla_k\omega(X,K),\ \frac{dK}{dt} = -\nabla_x\omega(X,K),\ X(T) = x,\ K(T) = k. \tag{4.3}$$

That is, if $(X(t),K(t))$ satisfy (4.3) then solution of (4.2) is given by

$$W(t,x,k) = W_0(X(0),K(0)). \tag{4.4}$$

In particular, trajectories of (4.1) satisfy the classical laws of the Newtonian mechanics:

$$\frac{dX}{dt} = K, \ \frac{dK}{dt} = -\nabla V(X). \tag{4.5}$$

We will be interested here in the long time, large distance effect of weak random fluctuations in the dispersion law on the behavior of solutions of the Liouville equations. In order to be more concrete, we first consider the classical case (4.5) with a weakly random potential:

$$\frac{dX}{dt} = K, \ \frac{dK}{dt} = -\sqrt{\delta}\nabla V(X), \ X(0) = x, \ K(0) = k, \tag{4.6}$$

where $\delta \ll 1$ is a small parameter measuring the strength of random fluctuations (it is a common convention for the size of fluctuations to be $\sqrt{\delta}$ so that the central limit theorem time scale is δ^{-1}). The long time, large distance behavior of a massive particle in a weakly random time-independent potential field is described by the momentum diffusion: the particle momentum undergoes the Brownian motion on the energy sphere. In terms of the Liouville equation, if W is the solution of

$$W_t + k \cdot \nabla_x W - \sqrt{\delta}\nabla V(x) \cdot \nabla_k W = 0, \tag{4.7}$$

then $\bar{W}_\delta(t, x, k) = \mathbb{E}[W(t/\delta, x/\delta, k)]$ converges as $\delta \to 0$ to the solution of the Fokker-Planck equation

$$\bar{W}_t + k \cdot \nabla \bar{W} = \sum_{m,n=1}^{d} \frac{\partial}{\partial k_n}\left(D_{nm}(k)\frac{\partial \bar{W}}{\partial k_m}\right). \tag{4.8}$$

The diffusion matrix $D = [D_{mn}]$ is given by

$$D_{nm}(k) = -\frac{1}{2}\int_{-\infty}^{\infty} \frac{\partial^2 R(sk)}{\partial x_n \partial x_m} ds, \quad m, n = 1, \dots, d. \tag{4.9}$$

Here $R(x) = \mathbb{E}[V(y)V(x + y)]$ is the covariance function of the random potential. The time scale $t \sim O(\delta^{-1})$ on which we obtain a non-trivial kinetic limit comes as no surprise – it is the same as the central limit theorem time scale we considered for a particle in a random velocity field in Section 2 (recall that fluctuation strength is $\sqrt{\delta}$).

Note that the diffusion matrix has the property $Dk = 0$:

$$\begin{aligned}\sum_{m=1}^{d} D_{nm}(k)k_m &= -\sum_{m=1}^{d} \tfrac{1}{2}\int_{-\infty}^{\infty} \tfrac{\partial^2 R(sk)}{\partial x_n \partial x_m} k_m ds \\ &= -\tfrac{1}{2}\int_{-\infty}^{\infty} \tfrac{d}{ds}\left(\frac{\partial R(s)}{\partial x_n}\right) ds = 0,\end{aligned} \tag{4.10}$$

provided that $\lim_{|x|\to+\infty} |\nabla_x R(x)| = 0$. This implies that the diffusion process $(X(t), K(t))$ corresponding to (4.8) is given by a solution of the Itô stochastic differential equation:

$$dK(t) = \sigma(K(t))dB_t + b(K(t))dt, \ dX(t) = K(t)dt, \tag{4.11}$$

where $\sigma(k)$ is a $d \times d$ symmetric matrix that is a square root of $2D(k)$, $b(k) = (b_1(k), \dots, b_d(k))$, with $b_m(k) = \sum_{n=1}^{d} \partial_{k_n} D_{mn}(k)$, and $\{B_t, t \geq 0\}$ is a d -dimensional, standard Brownian motion. The K-component of the diffusion satisfies $|K(t)| = \text{const}$, that is, $K(t)$ stays on a sphere of fixed radius. This, of course, is expected: the Hamiltonian $\omega_\delta(x, k) = k^2/2 + \sqrt{\delta}V(x)$, is preserved by the dynamics, and in the limit $\delta \to 0$ this reducers to the preservation of $|K(t)|$.

The Fokker-Planck limit has been first proved in [53] in dimensions higher than two, and later extended to two dimensions in [34, 58]. Similar results can be obtained for general Hamiltonians (dispersion laws), in particular, for the classical wave Hamiltonian $\omega(x,k) = c_\delta(x)|k|$, when the sound speed is weakly random: $c_\delta(x) = c_0 + \sqrt{\delta}\tilde{c}(x)$, where $c_0 > 0$ is constant and $\tilde{c}(x)$ is a mean-zero, spatially homogeneous, random field. However, for the sake of simplicity we will first concentrate on the classical mechanics Hamiltonian as in (4.9).

4.1.2. *Formal short time asymptotics.* The easiest way to understand how the Fokker-Planck equation and momentum diffusion come about is to consider the weakly random Hamiltonian system (4.6) on time scales of the order $t \sim O(1)$ rather than $t \sim O(\delta^{-1})$ when the behavior becomes non-trivial. Consider a formal expansion

$$X(t) = X_0(t) + \sqrt{\delta}X_1(t) + \delta X_2(t) + \dots,\ K(t) = K_0(t) + \sqrt{\delta}K_1(t) + \delta K_2(t) + \dots \tag{4.12}$$

Inserting this into (4.6) gives $K_0(t) = k$, $X_0(t) = x + kt$. The next order terms are

$$K_1(t) = \int_0^t \nabla V(x+ks)ds, \quad X_1(t) = \int_0^t ds \int_0^s \nabla V(x+ks_1)ds_1$$

and $K_2(t) = (K_{2,1}(t), \dots, K_{2,d}(t))$, where

$$K_{2,m}(t) = \sum_{n=1}^d \int_0^t ds \int_0^s ds_1 \int_0^{s_1} \partial^2_{x_n x_m} V(x+ks)\partial_{x_n} V(x+ks_2)ds_2.$$

Let $\Delta K_0(t/\delta) \approx K(t/\delta) - K_0(t/\delta)$. Hence, assuming that t/δ is large but not as much to violate validity of expansion (4.12), this is the case when $t \sim \delta^\gamma$ and $\gamma \in (1/2, 1)$, we conclude that

$$\sqrt{\delta}K_1(t/\delta) \sim \sigma(k)B_t, \quad \delta K_2(t/\delta) \sim b(k)t, \tag{4.13}$$

or equivalently

$$\Delta K_0(t/\delta) \approx \sigma(K_0(t/\delta))B_t + b(K_0(t/\delta))t, \tag{4.14}$$

where $t \sim \delta^\gamma$, in agreement with the Itô equation (4.11). The asymptotics of the first term can be obtained by an argument as in Section 2.1.1 for the integral (2.2). As for the second we get, via an ergodic theorem,

$$\begin{aligned}\delta K_2(t/\delta) &\approx -\sum_{n=1}^d \delta \int_0^{t/\delta} ds \int_0^s ds_1 \int_0^{s_1} \partial^2_{x_n}\partial_{x_m} R(k(s-s_2))ds_2 \\ &= -\sum_{n=1}^d \delta \int_0^{t/\delta} ds \int_0^s (s-s_2)\partial^2_{x_n}\partial_{x_m} R(k(s-s_2))ds_2 \\ &\to b_m(k)t,\end{aligned}$$

as $\delta \to 0$. We finish the derivation with an observation that the Fokker-Planck equation corresponding to the infinitesimal version of the "stochastic" difference equation (4.14) is (4.8). The formal expansion (4.12) can not, of course, hold on macroscopic time scales longer than δ^γ when $\gamma < 1/2$. Nevertheless, this "back of the envelope" calculation gives an excellent idea of what happens.

Let us derive the diffusion operator in (4.8) yet in another way, more in the spirit of of multiscale expansion of solutions of (4.7) done in Section 2.1.2. Let W be the solution of that equation, then $W^\delta(t,x,k) = W(t/\delta, x/\delta, k)$ satisfies

$$W^\delta_t + k\cdot\nabla_x W^\delta - \frac{1}{\sqrt{\delta}}\nabla V\left(\frac{x}{\delta}\right)\cdot\nabla_k W^\delta = 0. \tag{4.15}$$

Consider an asymptotic multiple scale expansion for W^δ

$$W^\delta(t,x,k) = \bar{W}(t,x,k) + \sqrt{\delta}W_1\left(t,x,\frac{x}{\delta},k\right) + \delta W_2\left(t,x,\frac{x}{\delta},k\right) + \dots \qquad (4.16)$$

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 (4.15) and obtain in the order $O\left(\delta^{-1/2}\right)$:

$$k \cdot \nabla_z W_1 = \nabla V(z) \cdot \nabla_k \bar{W}. \qquad (4.17)$$

Let $\theta \ll 1$ be a small positive regularization parameter that will be later sent to zero, and consider a regularized version of (4.17):

$$k \cdot \nabla_z W_1 + \theta W_1 = \nabla V(z) \cdot \nabla_k \bar{W},$$

Its solution is

$$W_1(z,k) = \int_{-\infty}^0 \nabla V(z+sk) \cdot \nabla_k \bar{W}(t,x,k) e^{\theta s} ds. \qquad (4.18)$$

The next order equation becomes upon averaging (we assume, as in Section 2.1.2, that the leading order term $\bar{W}$ is deterministic)

$$\frac{\partial \bar{W}}{\partial t} + k \cdot \bar{W} = \mathbb{E}\left(\nabla V(z) \cdot \nabla_k W_1\right). \qquad (4.19)$$

The term on the right side above may be computed explicitly using expression (4.18) for W_1:

$$\begin{aligned}
\mathbb{E}\left(\nabla V(z) \cdot \nabla_z \phi_1\right) &= \int_0^\infty \mathbb{E}\left[\frac{\partial V(z)}{\partial z_n}\frac{\partial V(z+sk)}{\partial z_m}\right] \frac{\partial^2 \bar{W}(t,x,k)}{\partial k_n \partial k_m} e^{-\theta s} ds \\
&= -\frac{1}{2}\int_{-\infty}^\infty \frac{\partial^2 R(sk)}{\partial x_n \partial x_m} e^{-\theta s} ds \frac{\partial^2 \bar{W}(t,x,k)}{\partial k_n \partial k_m} \\
&\to D_{nm}(k) \frac{\partial^2 \bar{W}(t,x,k)}{\partial k_n \partial k_m},
\end{aligned}$$

as $\theta \downarrow 0$, with $D_{mn}(k)$ as in (4.9).

However, the naive asymptotic expansion (4.16) is hard to justify. The rigorous proof outlined in the next section is based on a quite different method closer in the spirit to expansion (4.12).

4.2. **Outline of the rigorous proof of the Fokker-Planck limit.** The rigorous proof of the Foker-Planck limit is quite long, and an interested reader may consult the original papers [8, 53, 57, 58] for the details. Here we will only describe the main ingredients of the proof.

The random potential. First, let us make precise our assumptions on the random potential. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, and let $\mathbb{E}$ denote the expectation with respect to $\mathbb{P}$. We assume that $\{V(x), x \in \mathbb{R}^d\}$ is a real-valued random field over the probability space that is measurable and strictly stationary. This means that for any shift $x \in \mathbb{R}^d$, and any collection of points $x_1, \dots, x_n \in \mathbb{R}^d$ the laws of $(V(x_1+x), \dots, V(x_n+x))$ and $(V(x_1), \dots, V(x_n))$ are identical. In addition, we assume that the field is centered, i.e. $\mathbb{E}V(0) = 0$, the realizations of $V(x)$ are $\mathbb{P}$–a.s. C^2-smooth in $x \in \mathbb{R}^d$ and they satisfy

$$D_j := \max_{|\alpha|=j} \operatorname*{ess\text{-}sup}_{\omega \in \Omega} \|\partial_x^\alpha V(\cdot;\omega)\|_\infty < +\infty, \quad j = 0,1,2. \qquad (4.20)$$

We suppose further that the random field is strongly mixing in the uniform sense. More precisely, for any $R>0$ we let $\mathcal{C}_R^i$ (resp. $\mathcal{C}_R^e$) be the σ–algebra generated by random variables $\{V(x), x\in B_R\}$ (resp. $\{V(x), x\in B_R^c\}$). Here $B_R := [x : |x|\le R]$. 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 sufficiently fast, i.e.

$$\phi_* := \sup_{\rho\ge 0}\rho^{100}\phi(\rho) < +\infty. \tag{4.21}$$

This condition and (4.20) imply in particular that the covariance function of $V(x)$, $R(x) := \mathbb{E}[V(x)V(0)]$, satisfies

$$R_* := \sum_{i=0}^{4}\sum_{|\alpha|=i}\sup_{x\in\mathbb{R}^d}(1+|x|^2)^{50}|\partial_x^\alpha R(x)| < +\infty. \tag{4.22}$$

We also assume that the *power spectrum of the field*, defined as the Fourier transform of $R(x)$: $\hat{R}(k) = \int R(x)\exp(-ik\cdot x)dx$, has the following regularity property

$$\hat{R}(k) \text{ does not vanish identically on any hyperplane } H_p = \{k : \; k\cdot p = 0\},\; p\in\mathbb{R}^d. \tag{4.23}$$

The main convergence to Fokker-Planck result. Let the function $\phi^\delta(t,x,k)$ satisfy the Liouville equation

$$\begin{aligned}\frac{\partial\phi^\delta}{\partial t} + k\cdot\nabla_x\phi^\delta - \sqrt{\delta}\nabla V(x)\cdot\nabla_k\phi^\delta = 0, \\ \phi^\delta(0,x,k) = \phi_0(\delta x,k).\end{aligned} \tag{4.24}$$

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

$$D_{mn}(k) = -\frac{1}{2}\int_{-\infty}^{\infty}\frac{\partial^2 R(sk)}{\partial x_n\partial x_m}ds, \qquad m,n = 1,\dots,d. \tag{4.25}$$

The following result holds.

Theorem 4.1. *Let ϕ^δ be the solution of (4.24) and let $\bar\phi$ satisfy*

$$\begin{aligned}\frac{\partial\bar\phi}{\partial t} + k\cdot\nabla_x\bar\phi = \sum_{m,n=1}^{d}\frac{\partial}{\partial k_m}\left(D_{mn}(k)\frac{\partial\bar\phi}{\partial k_n}\right) \\ \bar\phi(0,x,k) = \phi_0(x,k),\end{aligned} \tag{4.26}$$

where the diffusivity matrix $D(k) = [D_{mn}(k)]$ is given by (4.9). Suppose also that $d\ge 3$. Then, for any $M, T_0 > 0$ there exist two constants C, $\alpha_0>0$ such that for all $T\ge T_0$

$$\sup_{(t,x,k)\in[0,T]\times\mathcal{A}(M)}\left|\mathbb{E}\phi^\delta\left(\frac{t}{\delta},\frac{x}{\delta},k\right) - \bar\phi(t,x,k)\right| \le CT(1+\|\phi_0\|_{1,4})\delta^{\alpha_0}. \tag{4.27}$$

Here for any function $f(t,x,k)$ and non-negative integers k,l,m we denote by $\|f\|_{k,l,m}$ the supremum norm over the first k (resp. l,m) derivatives in t (resp. x,k) variables. In case f does not depend on some of the variable we omit writing the corresponding subscript in the notation of the norm.

We recall that the process corresponding to the generator given by the operator on the right hand side of (4.26) is a diffusion process on a sphere $\mathbb{S}_l^{d-1} := [|k| = l]$. Equivalently, the above means that after the spherical change of variables $k \mapsto (l, \hat{k})$, where $l := |k|$ and $\hat{k} := k/|k|$, equations obtained from (4.26) for different values of l are decoupled. On the other hand, assumption (4.23) implies that the matrix $D(k)$ has rank $d-1$ for each $k \in \mathbb{S}_l^{d-1}$, hence the corresponding diffusion on the sphere of momenta of a fixed modulus is non-degenerate: see [8] for details.

Sketch of the proof of Theorem 4.1. The estimate (4.27) can be translated into the statement about the rate of convergence of one dimensional marginals of $(X^{(\delta)}(t), K^{(\delta)}(t))$ – the trajectories of the following system of ordinary differential equations

$$\frac{dX^{(\delta)}}{dt} = -K^{(\delta)},\ \frac{dK^{(\delta)}}{dt} = \frac{1}{\sqrt{\delta}}\nabla V\left(\frac{X^{(\delta)}}{\delta}\right),\ X^{(\delta)}(0) = x,\ K^{(\delta)}(0) = k. \tag{4.28}$$

Estimate (4.27) is equivalent to:

$$\sup_{(t,x,k)\in[0,T]\times\mathcal{A}(M)} \left|\mathbb{E}\phi_0\left(X^{(\delta)}(t), K^{(\delta)}(t)\right) - \bar{\phi}(t,x,k)\right| \le CT(1 + \|\phi_0\|_{1,4})\delta^{\alpha_0}, \tag{4.29}$$

provided that $X^{(\delta)}(0) = x$ and $K^{(\delta)}(0) = k$.

According to (4.29), the limit of $(X^{(\delta)}(t), K^{(\delta)}(t))$ should be a Markov process. Let us try to explain how Markovianity of the limit comes about from (4.28). For any two times $t_1 < t_2$, we have

$$\Delta K^{(\delta)}(t_1,t_2) := K^{(\delta)}(t_2) - K^{(\delta)}(t_1) \approx \frac{1}{\sqrt{\delta}}\int_{t_1+\delta^\gamma}^{t_2} \nabla V\left(\frac{X^{(\delta)}(s)}{\delta}\right) ds, \tag{4.30}$$

provided that $\gamma \in (1/2, 1)$. If the trajectory $\{X^{(\delta)}(t),\ t \ge t_1 + \delta^\gamma\}$ stays away from $\Gamma_{0,t_1} := \{(X^{(\delta)}(t), K^{(\delta)}(t),\ t \in [0,t_1]\}$ farther than a distance $\sim \delta$ apart then the increment $\Delta K^{(\delta)}(t_1,t_2)$, given $(X^{(\delta)}(t_1), K^{(\delta)}(t_1))$, would become "almost independent" of the information carried by Γ_{0,t_1} and we have a good chance that the limit is indeed Markovian. The above scenario can be endangered if $X^{(\delta)}(t_2)$ comes into the vicinity of Γ_{0,t_1} (i.e. closer than distance $\sim \delta$ of Γ_{0,t_1}). Then the increment $\Delta K^{(\delta)}(t_1,t_2)$, given $(X^{(\delta)}(t_1), K^{(\delta)}(t_1))$, would be, in a stark contrast with the Markov property, far from being "nearly" independent of the information contained in Γ_{0,t_1}.

Another obstacle in getting a diffusive limit could lie in the fact that the limiting process would be indeed Markovian, but a sufficiently "irregular" behavior of the path $K^{(\delta)}(t)$ could lead to the formation of jumps for the limiting momentum component. The limit would be then a Markovian jump process rather than a diffusion.

One way to show convergence is the "bootstrapping procedure" introduced by Kesten and Papanicolaou in [53]. We consider the trajectories corresponding to the Liouville equation (4.24) and introduce a stopping time, called τ_δ, so that until this time the process $(X^{(\delta)}(t), K^{(\delta)}(t))$ is well-behaved, i.e. none of the scenarios that can endanger diffusive limit described above is realized. Among other things, until this time the process $X^{(\delta)}(t)$ does not nearly self-intersect and the velocity $K^{(\delta)}(t)$ does not violate a Hölder property. These facts ensure in particular that until the stopping time the particle is "exploring a new territory" and, thanks to the strong mixing properties of the medium, "memory effects" are wiped out so we can essentially use the scaling argument described in (4.12) - (4.15). Therefore, until τ_δ

the law of the process is approximately described by the diffusion (4.11). After the stopping time we modify the dynamics of the process by augmenting it with the "true" diffusion, i.e., the limiting process described by (4.11). The main body of the argument is devoted to the proof that the law of the augmented process is close to the law of the diffusion, with an explicit error bound. Since an analogously defined stopping time corresponding to the limiting diffusion tends to infinity in probability the same holds for the stopping time τ_δ. In fact, for any $T > 0$ the probability of the event $[\tau_\delta \le T]$ can be estimated by $C\delta^{\gamma_1}$ for some $C, \gamma_1 > 0$. 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 : 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 $\tau_\delta \to \infty$, as $\delta \to 0$, shows that with a large probability the augmented process is close to the true process.

The stopping times. We now define the stopping time τ_δ, that prevents the trajectories to have near self-intersections (recall that the intent of the stopping time is to prevent any "memory effects").

We take small positive constants ϵ_j, $j = 1, 2, 3, 4$ and set

$$N = [\delta^{-\varepsilon_1}], \quad p = [\delta^{-\varepsilon_2}], \quad q = p\,[\delta^{-\varepsilon_3}], \quad N_1 = Np\,[\delta^{-\varepsilon_4}]. \tag{4.31}$$

The precise choice of ϵ_j-s is slightly technical and we will not go into details on their definitions. We introduce the following stopping times - each of them responsible for the control of an undesirable behavior of the trajectory (as described in the previous section). Let $t_k^{(p)} := kp^{-1}$ be a mesh of times, $k \ge 0$. To simplify the notation we omit writing subscript δ by a trajectory of solution of (4.28).

The "violent turn" stopping time. Let $\hat{K}(t) := K(t)/|K(t)|$. We define

$$S_\delta := \inf\Bigg[t \ge 0 : \text{ for some } k \ge 0 \text{ we have } t \in \left[t_k^{(p)}, t_{k+1}^{(p)}\right) \text{ and} \tag{4.32}$$
$$\hat{K}(t_{k-1}^{(p)}) \cdot \hat{K}(t) \le 1 - \frac{1}{N}, \text{ or } \hat{K}\left(t_k^{(p)} - \frac{1}{N_1}\right) \cdot \hat{K}(t) \le 1 - \frac{1}{N} \Bigg],$$

The stopping time S_δ is triggered when the trajectory performs a sudden turn – this is undesirable as in the limit the momentum component of the trajectory may loose its Hölder property, thus the limit can be a jump process instead of diffusion. In addition, the trajectory may return to the region it has just visited and create in this way correlations with the past.

"Near self-instersection" stopping time. For each $t \ge 0$, we denote by $\mathfrak{X}_t := \bigcup_{0 \le s \le t} X(s)$ the trace of the spatial component of the trajectory up to time t, and by $\mathfrak{X}_t(q) := [x : \text{dist}\,(x, \mathfrak{X}_t) \le 1/q]$ a tubular region around the X component of the path. We introduce the stopping time

$$U_\delta := \inf\left[t \ge 0 : \exists\, k \ge 1 \text{ and } t \in [t_k^{(p)}, t_{k+1}^{(p)}) \text{ for which } X(t) \in \mathfrak{X}_{t_{k-1}^{(p)}}(q) \right]. \tag{4.33}$$

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

$$\tau_\delta := S_\delta \wedge U_\delta. \tag{4.34}$$

Approximate martingale property. The next step (rather technical, hence we omit it) is the construction of a concatenated process $\{(\tilde{X}(t), \tilde{K}(t)),\, t \ge 0\}$ that follows the trajectories of (4.24) until the time τ_δ but after this time they become the momentum diffusion corresponding to the limiting Fokker-Planck operator, starting from the point $(\tilde{X}(\tau_\delta), \tilde{K}(\tau_\delta))$. The augmented process is not, of course, Markov, but it is "nearly Markov" in the sense we explain below.

Let $\{\tilde{\mathcal{F}}_t,\, t \ge 0\}$ be the natural filtration corresponding to the process, $\mathcal{L}$ be the generator of the Fokker-Planck equation

$$\mathcal{L}F(x,k) = -k \cdot \nabla_x F(x,k) + \sum_{m,n=1}^{d} \frac{\partial}{\partial k_n}\left(D_{nm}(k)\frac{\partial F(x,k)}{\partial k_m}\right). \tag{4.35}$$

and

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

Then, there exist constants $\gamma_1, C > 0$ such that for any smooth function G, random variable ζ that is $\tilde{\mathcal{F}}_t$-measurable we have

$$|\mathbb{E}\{[N_v(G) - N_t(G)]\,\zeta\}| \le C\delta^{\gamma_1}(v-t)\mathbb{E}\zeta. \tag{4.36}$$

If the right side of (4.36) vanished, that would mean that the law of $(\tilde{X}(t), \tilde{K}(t))$ satisfied the so called martingale problem of Strook and Varadhan. It is well known, see Chapter 6 of [76], that under the assumptions made about the generator $\mathcal{L}$ this problem is well-posed, i.e., there exists only one process, in the sense of uniqueness of law, that could satisfy the requirement that the functional $\{N_t(G),\, t \ge 0\}$ is a martingale for all smooth test functions G. The associated process is then a diffusion, thus in particular it is Markovian. One can check, by verifying the Hörmander condition, that the generator $\mathcal{L}$ is hypoelliptic. A particular consequence of this fact is that: when $d \ge 3$ for any $T \ge 1$ one can find constants $C, \gamma > 0$ such that $\mathbb{P}[\,\tau_\delta < T] \le C\delta^\gamma T$. The details of this argument can be found in [53].

On the other hand, the fact that the law of the trajectory satisfies an approximate martingale problem implies that the following result holds.

Theorem 4.2. *Suppose that $d \ge 3$. For any $T, M > 0$ one can choose $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 > 0$, appearing in the definition of the stopping times, and constants $C, \gamma > 0$ such that*

$$\mathbb{P}[\,\tau_\delta < T] \le C\delta^\gamma T, \quad \forall\, \delta \in (0,1],\, T \ge 1,\, (\tilde{X}(0), \tilde{K}(0)) \in \mathcal{A}(M). \tag{4.37}$$

The end of the proof of Theorem 4.1. The argument leading to the proof of the theorem is as follows. Let $G(t,x,k) := \bar{\phi}(u-t,x,k)$, where $\bar{\phi}(t,x,k)$ is the solution of the Cauchy problem (4.26), and $\zeta \equiv 1$. From (4.36) (with $v = u$, $t = 0$) we get

$$\left|\mathbb{E}\left[\phi_0(\tilde{X}(u), \tilde{K}(u)) - \bar{\phi}(u, \tilde{X}(0), \tilde{K}(0)) \;\; - \int_0^u (\partial_\varrho + \mathcal{L})G(\varrho, \tilde{X}(\varrho), \tilde{K}(\varrho))\, d\varrho\right]\right|$$
$$\le C\|G\|_{1,1,3}\delta^\gamma T.$$

Since $(\partial_\varrho + \mathcal{L})G \equiv 0$ we get

$$\left|\mathbb{E}\phi_0(\tilde{X}(u), \tilde{K}(u)) - \bar{\phi}(u,x,k)\right| \le C\|G\|_{1,1,3}\delta^\gamma T.$$

This is almost (4.29) except for the fact that under the expectation we have the modified, i.e. augmented, instead of the true process. However on $[\tau_\delta \geq T]$, we have $\tau_\delta \geq u$, and applying estimate (4.37) we get

$$\left|\mathbb{E}\Big[\phi_0(\tilde{X}(u), \tilde{K}(u)) - \bar{\phi}(u, x, k),\ \tau_\delta \geq T\Big]\right| \leq C\|G\|_{1,1,3}\delta^\gamma T + C\delta^\gamma T\|\phi_0\|_{0,0},$$

Since $(\tilde{X}(u), \tilde{K}(u)) = (X(u), K(u))$ on $[\tau_\delta \geq T]$, invoking again (4.37) we conclude (4.29). We use also that $\|G\|_{1,1,3} \leq C\|\phi_0\|_{1,4}$, which holds by virtue of standard P.D.E. estimates,

4.3. **Independence of two trajectories.** In the same spirit as we have studied the correlation of two trajectories in a random velocity field starting at two different points in Section 4.1 and have shown that the particle separation becomes a diffusion process on the same time scale as when the behavior of one particle becomes non-trivial, one may ask what happens after a long time for two particles that move in a weakly random potential force [68]. Consider the two particles $(X_j(t), K_j(t))$, $j = 1, 2$ that satisfy

$$\frac{dX_j}{dt} = K_j,\ \frac{dK_j}{dt} = -\sqrt{\delta}\nabla V(X_j),\ X_j(0) = x_j,\ K_j(0) = k_j. \tag{4.38}$$

The corresponding Liouville equation is

$$\frac{\partial \phi}{\partial t} + k \cdot \nabla\phi - \sqrt{\delta}\nabla V(x) \cdot \nabla_k \phi = 0. \tag{4.39}$$

As in the study of a single particle we consider rescaled time and space: $X_j^{(\delta)}(t) = \delta X_j(t/\delta)$, $K_j^{(\delta)}(t) = K_j(t/\delta)$:

$$\begin{aligned} &\tfrac{dX_j^{(\delta)}}{dt} = K_j^{(\delta)}, \quad \tfrac{dK_j(\delta)}{dt} = -\tfrac{1}{\sqrt{\delta}}\nabla V\left(\tfrac{X_j^{(\delta)}}{\delta}\right), \\ &X_j^{(\delta)}(0) = \tilde{x}_j := \delta x_j, \quad K_j^{(\delta)}(0) = \tilde{k}_j := k_j, \end{aligned} \tag{4.40}$$

and the rescaled Liouville equation is

$$\frac{\partial \phi^\delta}{\partial t} + k \cdot \nabla\phi^\delta - \frac{1}{\sqrt{\delta}}\nabla V\left(\frac{x}{\delta}\right) \cdot \nabla_k \phi^\delta = 0. \tag{4.41}$$

Let us see how far the starting points (x_1, k_1) and (x_2, k_2) can be from each other so that we may expect some correlation from the physical point of view. First, if $|\tilde{x}_1 - \tilde{x}_2| \gg O(\delta)$ then the particles would experience essentially two different random media. Hence, we should take $x_2 = x_1 + \delta y$, with $|y| = O(1)$. Similarly, if k_1 and k_2 are $O(1)$ apart then the particles will start moving into two different directions, which will quickly create separation of order much larger than $O(\delta)$, which, once again, will lead to their independent behavior. Hence, we should take $k_2 = k_1 + p$, with $|p| = O(\delta^\mu)$, and the exponent $\mu > 0$ to be determined, and look for a non-trivial behavior.

Accordingly, we define

$$U(t, x, k, y, p) = \phi^\delta(t, x, k)\phi^\delta(t, x + \delta y, k + p). \tag{4.42}$$

This function satisfies

$$\frac{\partial U}{\partial t} + k\cdot\nabla_x U + \frac{p}{\delta}\cdot\nabla_y U - \frac{1}{\sqrt{\delta}}\nabla V\left(\frac{x}{\delta}\right)\cdot\nabla_k U - \frac{1}{\sqrt{\delta}}\left[\nabla V\left(\frac{x}{\delta} + y\right) - \nabla V\left(\frac{x}{\delta}\right)\right]\cdot\nabla_p U = 0. \tag{4.43}$$

The corresponding characteristics are

$$
\begin{aligned}
\frac{dX}{dt} &= K(t), \quad \frac{dY}{dt} = \frac{1}{\delta}P, \\
\frac{dK}{dt} &= -\frac{1}{\sqrt{\delta}}\nabla V\left(\frac{x}{\delta}\right), \quad \frac{dP}{dt} = -\frac{1}{\delta^{1/2}}\left[\nabla V\left(\frac{X}{\delta}+Y\right) - \nabla V\left(\frac{X}{\delta}\right)\right].
\end{aligned}
\tag{4.44}
$$

Consider first, roughly, what happens on a short time scale $t \sim O(\delta^\alpha)$: let $t = \delta^\alpha s$. Then $K(s) \approx \tilde{k}_1$, and $X(s) \approx \tilde{x}_1 + \delta^\alpha \tilde{k}_1 s$, since the Fokker-Planck diffusion kicks in only on the time scale $t \sim O(1)$. Therefore, we have, approximately:

$$
\frac{dY}{ds} = \delta^{\alpha-1}P, \ \frac{dP}{ds} = -\frac{1}{\delta^{1/2-\alpha}}\left[\nabla V\left(\frac{\tilde{x}_1 + \delta^\alpha \tilde{k}_1 s}{\delta} + Y\right) - \nabla V\left(\frac{\tilde{x}_1 + \delta^\alpha \tilde{k}_1 s}{\delta}\right)\right],
\tag{4.45}
$$

with the initial data $Y(0) = y$, $P(0) = \delta^\mu p_0$. It is easy to see that when $\alpha \geq 1$, $P(s)$ stays essentially constant for $s \sim O(1)$, and thus nothing interesting happens to $Y(s)$, as it remains close to its initial value y. When $\alpha < 1$ the argument in ∇V in (4.45) is rapidly oscillating in s, hence we can expect a non-trivial limit. Then $|P(s)| \sim O(\delta^\mu + \delta^{\alpha/2})$ and $|Y(s) - y| \sim O(\delta^{\alpha+\mu-1} + \delta^{3\alpha/2-1})$, with the first contribution in both expressions coming from the initial condition $P(0) = \delta^\mu p_0$. Therefore, $|Y(s) - y|$ tends to infinity at times $t \gg O(\delta^{2/3})$, meaning that the particles are separated by distance of the order much larger than $O(\delta)$ and their trajectories become decorrelated. Note that this result does not depend on how small $P(0)$ is – even two particles that have the same momentum initially but are separated by distance $O(\delta)$ become uncorrelated on time scales beyond $t \sim O(\delta^{2/3})$.

Without discussing the details we mention that this time scale can be interpreted in terms of the emergence of a random caustic discussed in [79] that appears on the same time scale $O(\delta^{2/3})$. Slightly more precisely, as we have discussed in Section 3.2, the random Liouville equations in a weakly random medium are nothing but the phase space version of the ray equations. Hence, it is not surprising that the decorrelation of various rays in the phase space occurs on the same time scale a caustic appears in the physical space.

A formal asymptotic analysis of the trajectory decorrelation. Let us now address the question of the decorrelation of two trajectories in terms of the formal asymptotic expansions. We will consider the situation when initially the particles are separated by distance $O(\delta)$ and their momenta are $O(\delta^\mu)$ apart. Hence, we redefine slightly (4.42) as

$$
U(t, x, k, y, p) = \phi^\delta(t, x, k)\phi^\delta(t, x + \delta y, k + \delta^\mu p).
\tag{4.46}
$$

The function ϕ^δ is the solution of (4.41), while U satisfies

$$
\begin{aligned}
0 = \frac{\partial U}{\partial t} &+ k \cdot \nabla_x U + \frac{p}{\delta^{1-\mu}} \cdot \nabla_y U - \frac{1}{\sqrt{\delta}}\nabla V\left(\frac{x}{\delta}\right) \cdot \nabla_k U \\
&- \frac{1}{\delta^{1/2+\mu}}\left[\nabla V\left(\frac{x}{\delta} + y\right) - \nabla V\left(\frac{x}{\delta}\right)\right] \cdot \nabla_p U.
\end{aligned}
\tag{4.47}
$$

We are interested in the question of how long the correlations between the two trajectories persist, that is, what is the time scale $t \sim O(\delta^\alpha)$ (with $\alpha > 0$ determined by μ) on which solution of (4.47) behaves in a non-trivial fashion. Hence, we rescale time $t = \delta^\alpha s$, and also correspondingly the spatial variable $x = \delta^\alpha z$. and obtain

$$
\begin{aligned}
0 = \frac{\partial U}{\partial s} &+ k \cdot \nabla_z U + \frac{p}{\delta^{1-\mu-\alpha}} \cdot \nabla_y U - \frac{1}{\delta^{1/2-\alpha}}\nabla V\left(\frac{z}{\delta^{1-\alpha}}\right) \cdot \nabla_k U \\
&- \frac{1}{\delta^{1/2+\mu-\alpha}}\left[\nabla V\left(\frac{x}{\delta^{1-\alpha}} + y\right) - \nabla V\left(\frac{x}{\delta^{1-\alpha}}\right)\right] \cdot \nabla_p U.
\end{aligned}
$$

Let us first consider the range $1/3 \le \mu < 1/2$, then, as we will see, a non-trivial behavior is observed if $\alpha = 2\mu$, which gives

$$0 = \frac{\partial U}{\partial s} + k \cdot \nabla_z U + \tfrac{p}{\delta^{1-3\mu}} \cdot \nabla_y U - \tfrac{1}{\delta^{1/2-2\mu}} \nabla V\left(\tfrac{x}{\delta^{1-2\mu}}\right) \cdot \nabla_k U \\ - \tfrac{1}{\delta^{1/2-\mu}} \left[\nabla V\left(\tfrac{x}{\delta^{1-2\mu}} + y\right) - \nabla V\left(\tfrac{x}{\delta^{1-2\mu}}\right)\right] \cdot \nabla_p U.$$

Now, if $1/3 < \mu < 1/2$ then the non-trivial terms above are

$$\frac{\partial U}{\partial s} + k \cdot \nabla_z U - \frac{1}{\delta^{1/2-\mu}} \left[\nabla V\left(\frac{x}{\delta^{1-2\mu}} + y\right) - \nabla V\left(\frac{x}{\delta^{1-2\mu}}\right)\right] \cdot \nabla_p U = 0. \quad (4.48)$$

A formal multiple scales asymptotic expansion, as the one in Section 4.1.2, gives in the limit, for $\bar{U} = \lim_{\delta\downarrow 0} \mathbb{E}(U)$:

$$\frac{\partial \bar{U}}{\partial s} + k \cdot \nabla_z \bar{U} = \sum_{i,j=1}^{d} \frac{\partial}{\partial p_i} \left(D_2^{ij}(y,k) \frac{\partial \bar{U}}{\partial p_j} \right), \quad (4.49)$$

with

$$D_2^{ij}(y,k) = \int_{-\infty}^{\infty} \left[\frac{\partial^2 R(y+ks)}{\partial y_i \partial y_j} - \frac{\partial^2 R(ks)}{\partial y_i \partial y_j} \right] ds. \quad (4.50)$$

Therefore, the momentum separation behaves as a diffusion while the physical space separation y does not change yet. On the other hand, when $\mu = 1/3$ the the limiting equation is

$$\frac{\partial \bar{U}}{\partial s} + k \cdot \nabla_z \bar{U} + p \cdot \nabla_y \bar{U} = \sum_{i,j=1}^{d} \frac{\partial}{\partial p_i} \left(D_2^{ij}(y,k) \frac{\partial \bar{U}}{\partial p_j} \right), \quad (4.51)$$

Hence, on the time scale $O(\delta^{2/3})$ the particle separation y also starts to evolve, after which the particle are, on the macro-scale separated by distance much larger than $O(\delta)$ and the trajectories become decorrelated. We will return to this result later when we discuss the self-averaging properties for waves in random medium.

4.4. **Spatial diffusion.** Consider the Fokker-Planck equation (4.8):

$$\frac{\partial W}{\partial t} + k \cdot \nabla_x W = \sum_{n,m=1}^{d} \frac{\partial}{\partial k_n} \left(D_{nm}(k) \frac{\partial W}{\partial k_m} \right). \quad (4.52)$$

Solutions of (4.52) converge in the long time limit to the solutions of the spatial diffusion equation [57]. Let us discuss briefly this result. Suppose that $W_\gamma(t,x,k) = W(t/\gamma^2, x/\gamma, k)$, where $\gamma > 0$ and W satisfies (4.8) with the initial data $W_\gamma(0,t,x,k) = \phi_0(\gamma x, k)$. For a fixed $\ell > 0$ we let $w(t,x,\ell)$ be the solution of the Cauchy problem for the spatial diffusion equation:

$$\frac{\partial w}{\partial t} = \sum_{m,n=1}^{d} a_{mn}^{(\ell)} \frac{\partial^2 w}{\partial x_n \partial x_m}, \quad (4.53)$$
$$w_\ell(0,x) = \bar{\phi}_0(x,\ell)$$

with the averaged initial data

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

Here $d\Omega(\hat{k})$ is the surface measure on the unit sphere $\mathbb{S}^{d-1}$ and Γ_{d-1} is the surface area of the sphere. The diffusion matrix $A^{(\ell)} := [a_{nm}^{(\ell)}]$ in (4.53) is given explicitly as

$$a_{nm}^{(\ell)} = \frac{1}{\Gamma_{d-1}} \int_{\mathbb{S}^{d-1}} \hat{k}_n \chi_m(\ell\hat{k}) d\Omega(\hat{k}). \tag{4.54}$$

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

$$\sum_{m,n=1}^{d} \frac{\partial}{\partial k_m} \left(D_{mn}(k) \frac{\partial \chi_j}{\partial k_n} \right) = -k_j. \tag{4.55}$$

Note that equations (4.55) for χ_m are elliptic on each sphere $\mathbb{S}_\ell^{d-1} := [|k| = \ell]$ and the existence of solution follows from an application of Fredholm alternative. It can be checked, by a direct calculation, that the matrix $A^{(\ell)}$ is positive definite [57]. Indeed, let $\mathbf{c} = (c_1, \ldots, c_d) \in \mathbb{R}^d$ be a fixed vector and let $\chi_{\mathbf{c}}(k) := \sum_{m=1}^{d} c_m \chi_m(k)$. Since the matrix D is non-negative we have

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

The last equality holds after integration by parts because $D(k)k = 0$. Moreover, the inequality appearing in the last line of (4.56) is strict. This can be seen as follows. Since the null-space of the matrix $D(k)$ is one-dimensional for each k and consists of the vectors parallel to k, in order for $(A^{(\ell)}\mathbf{c}, \mathbf{c})_{\mathbb{R}^d}$ to vanish one needs that the gradient $\nabla \chi_{\mathbf{c}}(\ell\hat{k})$ is parallel to $\hat{k}$ for all $\hat{k} \in \mathbb{S}^{d-1}$. This, however, together with (4.55) would imply that $\hat{k} \cdot \mathbf{c} = 0$ for all $\hat{k}$, which is impossible.

The following theorem holds [57].

Theorem 4.3. *For every $0 < T_* < T < +\infty$ the re-scaled solution $W_\gamma(t, x, k) = W(t/\gamma^2, x/\gamma, k)$ of (4.8) converges as $\gamma \to 0$ in $C([T_*, T]; L^\infty(\mathbb{R}^{2d}))$ to $w_{|k|}(t, x)$. Moreover, there exists a constant $C > 0$ so that we have*

$$\sup_{(t,x,k) \in [T_*,T] \times \mathcal{A}(M)} |w_{|k|}(t, x) - W_\gamma(t, x, k)| \leq C \left(\gamma T + \sqrt{\gamma}\right) \|\phi_0\|_{1,1}. \tag{4.56}$$

The proof of Theorem 4.3 is based on classical asymptotic expansions and is quite straightforward. As an immediate corollary of Theorems 4.1 and 4.3 we obtain the following result.

Theorem 4.4. *Suppose that $d \geq 3$. Let W_δ be solution of (4.7) with the initial data $W_\delta(0, x, k) = \phi_0(\delta^{1+\alpha} x, k)$ and let $\bar{w}_{|k|}(t, x)$ be the solution of the diffusion equation (4.53) with the initial data $w_{|k|}(0, x) = \bar{\phi}_0(x, |k|)$. Then, for any $T > T_* > 0$ and $M > 0$ there exist $\alpha_0 > 0$ and a constant $C > 0$ so that for all $0 \leq \alpha < \alpha_0$*

$$\sup_{(t,x,k) \in [T_*,T] \times \mathcal{A}(M)} \left| w_{|k|}(t, x) - \mathbb{E} W_\delta\left(t/\delta^{1+2\alpha}, x/\delta^{1+\alpha}, k\right) \right| \leq C T \delta^{\alpha_0 - \alpha}. \tag{4.57}$$

Theorem 4.4 shows that the movement of a particle in a weakly random quenched Hamiltonian is approximated by a Brownian motion in the long time-large space limit, at least for times $T \ll \delta^{\alpha-\alpha_0}$. It would be quite interesting to understand what happens on longer time scales, if anything different.

4.5. **General Hamiltonians.** So far we have described the results for the Hamiltonian of classical mechanics $H(x,k) = k^2/2 + V(x)$. However, the Fokker-Planck limit, as well as the passage to the spatial diffusion limit can be obtained for a large class of weakly random Hamiltonians, see [57]. Consider a particle that moves in an isotropic weakly random Hamiltonian flow with the Hamiltonian of the form $H_\delta(x,k) = H_0(x,k) + \sqrt{\delta} H_1(x,k)$, with $x, k \in \mathbb{R}^d$ and $d \geq 3$:

$$\frac{dX^\delta}{dt} = \nabla_k H_\delta,\ \frac{dK^\delta}{dt} = -\nabla_x H_\delta,\ X^\delta(0) = x_0,\ K^\delta(0) = k_0. \tag{4.58}$$

Here $H_0(x,k)$ is a deterministic function, the so called *background Hamiltonian* and $H_1(x,k)$ is a random perturbation. The special case,

$$H_\delta(x,k) = (c_0 + \sqrt{\delta} c_1(x))|k|, \tag{4.59}$$

arises in the geometrical optics limit for acoustic waves, and is of a particular interest to us. Here c_0 is the background sound speed, and $c_1(x)$ is a random perturbation. We assume that the background Hamiltonian $H_0(x,k)$ is isotropic, that is, it depends only on $|k|$, and is uniform in space, i.e. independent of x. Moreover, we assume that $H_0 : [0,+\infty) \to \mathbb{R}$ 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 $k > 0$.

Assumptions on the random perturbation. Our assumptions on the random perturbation are very similar to what we have assumed about the random potential in the classical case $H_\delta(x) = k^2/2 + \sqrt{\delta} V(x)$. We assume that the random perturbation $H_1(x,|k|)$ is a random field stationary in x Here, this means that for any $(x,\ell) \in \mathbb{R}^d \times [0,+\infty)$ and a collection of points $x_1,\ldots,x_n \in \mathbb{R}^d$ the laws of $(H_1(x_1+x,\ell),\ldots,H_1(x_n+x,\ell))$ and $(H_1(x_1,\ell),\ldots,H_1(x_n,\ell))$ are identical. In addition, we assume that $\mathbb{E}H_1(x,\ell) = 0$, $\mathbb{P}$–a.s. realizations of $H_1(x,\ell)$ are C^2-smooth in $(x,\ell) \in \mathbb{R}^d \times (0,+\infty)$, with the respective derivatives that are deterministically bounded.

We suppose further that the random field is strongly mixing in the uniform sense, as we did for the classical Hamiltonian, and that the two-point spatial correlation function of the random field H_1 is $R(y,\ell) := \mathbb{E}[H_1(y,\ell)H_1(0,\ell)]$ is sufficiently rapidly decaying in y. Thanks to the smoothness assumptions made on the realizations of the field it is C^4 smooth in (y,ℓ). We also need an analog of assumption (4.23)

$$\hat{R}(k,\ell) \text{ does not vanish identically on any hyperplane } H_p = \{k :\ k \cdot p = 0\},\ p \in \mathbb{R}^d \tag{4.60}$$

The above hypothesis is made to ensure that the Fokker-Planck diffusion matrix is of rank $d-1$. Here $\hat{R}(k,\ell) = \int R(x,\ell)\exp(-ik\cdot x)dx$ 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, and in particular, for the wave Hamiltonian (4.59).

Under the above assumptions we have exactly the same results on the convergence of the solutions of the weakly random Hamiltonian system first to a Fokker-Planck limit, and then to a Brownian motion, as for the mechanical Hamiltonian. Let the function $\phi_\delta(t,x,k)$ satisfy the Liouville equation

$$\begin{aligned}&\frac{\partial\phi^\delta}{\partial t}+\nabla_x H_\delta(x,k)\cdot\nabla_k\phi^\delta-\nabla_k H_\delta(x,k)\cdot\nabla_x\phi^\delta=0,\\&\qquad\phi^\delta(0,x,k)=\phi_0(\delta x,k).\end{aligned}\tag{4.61}$$

The Fokker-Planck diffusion matrix $D_{mn}(\hat{k},\ell)$ is now

$$\begin{aligned}D_{mn}(\hat{k},\ell)&=-\tfrac{1}{2}\int_{-\infty}^{\infty}\tfrac{\partial^2 R(H_0'(\ell)s\hat{k},\ell)}{\partial x_n\partial x_m}ds\\&=-\tfrac{1}{2H_0'(\ell)}\int_{-\infty}^{\infty}\tfrac{\partial^2 R(s\hat{k},\ell)}{\partial x_n\partial x_m}ds,\qquad m,n=1,\dots,d.\end{aligned}\tag{4.62}$$

We have the following generalization of Theorem 4.1.

Theorem 4.5. *Suppose that $d\ge 3$ and ϕ^δ is the solution of (4.61) and $\bar{\phi}$ satisfies*

$$\begin{aligned}&\frac{\partial\bar{\phi}}{\partial t}=\sum_{m,n=1}^{d}\frac{\partial}{\partial k_m}\left(D_{mn}(\hat{k},\ell)\frac{\partial\bar{\phi}}{\partial k_n}\right)+H_0'(\ell)\,\hat{k}\cdot\nabla_x\bar{\phi}\\&\qquad\bar{\phi}(0,x,k)=\phi_0(x,k).\end{aligned}\tag{4.63}$$

Then, for any $M,T_0>0$ there exist constants C, $\alpha_0>0$ such that for all $T\ge T_0$, $\delta\in(0,1]$

$$\sup_{(t,x,k)\in[0,T]\times\mathcal{A}(M)}\left|\mathbb{E}\phi^\delta\left(\frac{t}{\delta},\frac{x}{\delta},k\right)-\bar{\phi}(t,x,k)\right|\le CT(1+\|\phi_0\|_{1,4})\delta^{\alpha_0}.\tag{4.64}$$

It is easy to check that we still have the property $D(\hat{k},\ell)\hat{k}=0$, thus the K-process generated by (4.63) is indeed a diffusion process on a sphere $[|k|=\ell]$, or, equivalently, equations (4.63) for different values of k are decoupled.

As in the mechanical Hamiltonian case, solutions of (4.63) converge in the long time limit to the solutions of the spatial diffusion equation. Let $\bar{\phi}_\gamma(t,x,k)=\bar{\phi}(t/\gamma^2,x/\gamma,k)$, where $\bar{\phi}$ satisfies (4.63) with an initial data $\bar{\phi}_\gamma(0,t,x,k)=\phi_0(\gamma x,k)$, and $w^{(\ell)}(t,x)$ be the solution of the spatial diffusion equation (4.53). The diffusion matrix $A^{(\ell)}:=[a_{nm}^{(\ell)}]$ is

$$a_{nm}^{(\ell)}=\frac{H_0'(\ell)}{\Gamma_{d-1}}\int_{\mathbb{S}^{d-1}}\hat{k}_n\chi_m(\ell\hat{k})d\Omega(\hat{k}).\tag{4.65}$$

Here χ_j are the mean-zero solutions of

$$\sum_{m,n=1}^{d}\frac{\partial}{\partial k_m}\left(D_{mn}(\hat{k},k)\frac{\partial\chi_j}{\partial k_n}\right)=-H_0'(|k|)\hat{k}_j.\tag{4.66}$$

Then Theorem 4.3 remains in force, the only difference is that instead of solutions $W_\gamma(t,x,k)$ of (4.8) the statement concerns $\bar{\phi}_\gamma(t,x,k)=\bar{\phi}(t/\gamma^2,x/\gamma,k)$ the solutions of (4.63). Finally from this result and Theorem 4.27 we obtain an analogue of Theorem 4.4 for $\mathbb{E}\bar{\phi}_\delta(t,x,k)$.

4.6. **From waves to diffusion and self-averaging.**

The Wigner transform of mixtures of states. What does this have to do with waves? If we think of the Liouville equations as the ray equations of geometric optics (see Section 3.2), and its relation to the limit Wigner transform, we can interpret solution of (4.61) with the wave Hamiltonian (4.59) as the high frequency limit of the phase space wave energy density in a weakly random medium with a correlation length δ which is much larger than the wave length. Then functionals of the form

$$I^\delta(t) = \int \phi^\delta(t,x,k)\eta(x,k)dxdk,$$

with a smooth, rapidly decaying test function $\eta(x,k)$ are the local phase space averages of the wave energy density. Theorem 4.5 implies that $\mathbb{E}(I^\delta(t)) \to \bar{I}(t)$ as $\delta \to 0$, with

$$\bar{I}(t) = \int \bar{W}(t,x,k)\eta(x,k)dxdk. \tag{4.67}$$

Here $\bar{W}(t,x,k)$ is the solution of the Fokker-Planck equation (4.8), while on an even long time scale it would be the solution of the spatial diffusion equation. This process involves several approximations: wave energy by the solutions of the Liouville equations, Liouville by Fokker-Planck, and, lastly, Fokker-Planck to spatial diffusion. Can we have a direct estimate for the approximation of the wave energy by the solution of the spatial diffusion equation? For that we would need each of the approximations above to come with an error bound. And, indeed, the last two do have error bounds as described in the preceding sections. On the other hand, the the Liouville equations for the Wigner transform come without error bounds, and, actually, the limit is just weak.

In order to overcome this problem we introduce the Wigner transform of a mixture of states. It is defined as follows: consider a family of functions $f_\varepsilon(x,\zeta)$ and set

$$W_\varepsilon(x,k) = \int e^{ik\cdot y} f_\varepsilon(x - \frac{\varepsilon y}{2};\zeta) f_\varepsilon^*(x + \frac{\varepsilon y}{2};\zeta)\frac{dyd\mu(\zeta)}{(2\pi)^d}.$$

The family f_ε depends on an additional "state" parameter $\zeta \in S$, where S is a state space equipped with a non-negative bounded measure $d\mu(\zeta)$. One should think of μ as having a non-trivial support so that the "mixture of states" is, indeed, a mixture. Typically this amounts to introducing random initial data for f_ε at $t = 0$ and taking the expectation of Winger transform with respect to this randomness. The remarkable fact is that the Wigner transform of a mixture of states may be much more regular than a "pure" Wigner transform, and converge to its limit in a much stronger sense [60, 74]. Another context where the mixtures of states arise naturally is in the time reversal applications.

Let us now be a little more precise on the regime we are talking about. Start with the wave equation in dimension $d \geq 3$

$$\frac{1}{c^2(x)}\frac{\partial^2\phi}{\partial t^2} - \Delta\phi = 0 \tag{4.68}$$

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. Rescaling the spatial and temporal variables $x = x'/\delta$ and $t = t'/\delta$ we obtain (after dropping the primes) equation (4.68) with rapidly fluctuating wave speed

$$c_\delta(x) = c_0 + \sqrt{\delta}c_1\left(\frac{x}{\delta}\right). \tag{4.69}$$

It is convenient to re-write (4.68) as the acoustic system $p = \frac{1}{c}\phi_t$ and $u = -\nabla\phi$:

$$\frac{\partial u}{\partial t} + \nabla\left(c_\delta(x)p\right) = 0 \qquad (4.70)$$
$$\frac{\partial p}{\partial t} + c_\delta(x)\nabla\cdot u = 0.$$

We will denote for brevity $v = (u,p)$ and assume that the initial data $v_0(x;\zeta) = (-\varepsilon\nabla\phi_0^\varepsilon, 1/c_\delta\dot{\phi}_0^\varepsilon)$ is an ε-oscillatory and compact at infinity family of functions uniformly bounded in $L^2(\mathbb{R}^d)$ for each "realization" ζ of the initial data. The scale ε of oscillations is much smaller than the correlation length δ of the medium: $\varepsilon \ll \delta \ll 1$. The $(d+1)\times(d+1)$ Wigner matrix of a mixture of solutions of is

$$W_\varepsilon^\delta(t,x,k) = \int\limits_{\mathbb{R}^d\times S} e^{ik\cdot y} v_\varepsilon^\delta(t,x-\frac{\varepsilon y}{2};\zeta) v_\varepsilon^{\delta*}(t,x+\frac{\varepsilon y}{2};\zeta)\frac{dyd\mu(\zeta)}{(2\pi)^d}.$$

The non-negative measure $d\mu$ has bounded total mass: $\int_S d\mu(\zeta) < \infty$. As we have discussed in Section 3 for each fixed $\delta > 0$ (and even without introduction of a mixture of states) one may pass to the limit $\varepsilon \to 0$ and show that W_ε^δ converges weakly in $\mathcal{S}'(\mathbb{R}^d\times\mathbb{R}^d)$ to

$$\bar{W}^\delta(t,x,k) = u_+^\delta(t,x,k)b_+(k)\otimes b_+(k) + u_-^\delta(t,x,k)b_-(k)\otimes b_-(k), \qquad (4.71)$$

with $b_\pm(k) = (\hat{k}/\sqrt{2}, \pm 1/\sqrt{2})$. The scalar amplitudes $u_\pm^\delta$ satisfy the Liouville equations:

$$\frac{\partial u_\pm^\delta}{\partial t} + \nabla_k H_\pm^\delta\cdot\nabla_x u_\pm^\delta - \nabla_x H_\pm^\delta\cdot\nabla_k u_\pm^\delta = 0, \qquad (4.72)$$

with $H^\pm(x,k) = \pm(c_0 + \sqrt{\delta}c - 1(x/\delta))|k|$. Furthermore, Theorem 4.5 implies that one may pass to the limit $\delta\to 0$ in (4.72) and conclude that $\mathbb{E}\left\{u_\pm^\delta\right\}$ converge to the solution of

$$\frac{\partial \bar{u}_\pm}{\partial t} \pm c_0\hat{k}\cdot\nabla_x\bar{u}_\pm = \sum_{m,n=1}^d \frac{\partial}{\partial k_m}\left(|k|^2 D_{mn}(\hat{k})\frac{\partial\bar{u}_\pm}{\partial k_n}\right). \qquad (4.73)$$

Here, the diffusion matrix $D(\hat{k}) = [D_{mn}(\hat{k})]$ is given by

$$D_{mn}(\hat{k}) = \frac{1}{2}\int_{-\infty}^{\infty}\frac{\partial^2 R(c_0 s\hat{k})}{\partial x_n\partial x_m}ds, \qquad (4.74)$$

where $R(x)$ is the covariance function of c_1: $\mathbb{E}\left\{c_1(y)c_1(x+y)\right\} = R(x)$.

When can we justify the direct passage from W_ε^δ to the Fokker-Planck limit $\bar{u}_{pm}$? For that we need an error bound in the approximation of W_ε^δ by W^δ. We can do this in the following regime: let $\mathcal{K}_\mu = \left\{(\varepsilon,\delta):\ \delta \geq |\ln\varepsilon|^{-2/3+\mu}\right\}$, with $0<\mu<2/3$ and assume that $(\varepsilon,\delta)\in\mathcal{K}_\mu$ for some $\mu\in(0,2/3)$. From now on, μ is a given fixed number in $(0,2/3)$.

We assume that the initial Wigner transform $W_\varepsilon^\delta(0,x,k)$ is uniformly bounded in $L^2(\mathbb{R}^d\times\mathbb{R}^d)$ and

$$W_\varepsilon^\delta \to W_0 \text{ strongly in } L^2(\mathbb{R}^d\times\mathbb{R}^d) \text{ as } \mathcal{K}_\mu \ni (\varepsilon,\delta)\to 0. \qquad (4.75)$$

This is possible because we are considering mixtures of states, We also assume that $W_0 \in C_c(\mathbb{R}^d\times\mathbb{R}^d)$ with a support that satisfies

$$\text{supp } W_0(x,k) \subseteq X = \left\{(x,k):\ |x|\leq C,\ C^{-1}\leq |k|\leq C\right\}. \qquad (4.76)$$

Note that (4.75) may not hold for a pure state since $\|\tilde{W}_\varepsilon\|_2 = (2\pi\varepsilon)^{-d/2}\|f_\varepsilon\|_2^2$ (see [60]). We also assume that W_0 has the form

$$W_0(x,k) = u_+^0 b_+ \otimes b_+ + u_-^0 b_- \otimes b_- \tag{4.77}$$

We have the following approximation theorem.

Theorem 4.6. *Under the above assumptions, we have*

$$\|W_\varepsilon^\delta(t) - W^\delta(t)\|_2 \le C(\delta)\varepsilon\|W_0\|_{H^3(\mathbb{R}^{2d})} e^{Ct/\delta^{3/2}} + \|W_\varepsilon^\delta(0) - W_0\|_2, \tag{4.78}$$

where $C(\delta)$ is a rational function of δ with deterministic coefficients that may depend on the constant $C > 0$ in the bound (4.76) on the support of W_0.

This theorem allows us to pass from the wave energy for mixture of states to the Fokker-Planck limit, and then to the diffusive limit, using Theorems 4.5 and 4.4. For instance, we have the following result. Let

$$\bar{W}(t,x,k) = \bar{u}_+(t,x,k)b_+(k) \otimes b_+(k) + \bar{u}_-(t,x,k)b_-(k) \otimes b_-(k). \tag{4.79}$$

The functions $\bar{u}_\pm$ satisfy the Fokker-Planck equation (4.73) with initial data $u_\pm^0$ as in (4.77).

Theorem 4.7. *Let $S(\cdot) \in L^2(\mathbb{R}^d)$ be a test function, and define the moments*

$$s_\varepsilon^\delta(t,x) = \int W_\varepsilon^\delta(t,x,k)S(k)dk \qquad \text{and} \qquad \bar{s}(t,x) = \int \bar{W}(t,x,k)S(k)dk,$$

where $\bar{W}$ is given by (4.79). Then for each $t > 0$ we have

$$\mathbb{E}\left\{\int |s_\varepsilon^\delta(t,x) - \bar{s}(t,x)|^2 dx\right\} \to 0 \tag{4.80}$$

as $\mathcal{K}_\mu \ni (\varepsilon,\delta) \to 0$.

Theorem 4.7 contains two results: first, it gives an approximation of the phase space wave energy by the solution of the Fokker-Planck equation. Second, it means that the moments s_ε^δ converge in probability to a deterministic limit. That is, the locally averaged wave energy density is, actually, not random in the regime of random geometric optics after propagation over long distances. The effect of the random medium is not at all small – it is reflected in the diffusion coefficient in the Fokker-Planck equation, but the energy density still does not depend on the particular details of the realization of the random medium.

5. **Radiative transport regime for the Schrödinger equation.**

5.1. **The radiative transport limit.**

The radiative transport regime for the weakly random Schrödinger equation arises as follows. Consider the Schrödinger equation

$$i\frac{\partial\phi}{\partial t} + \frac{1}{2}\Delta\phi - \sqrt{\varepsilon}V(x)\phi = 0. \tag{5.1}$$

Here $V(x)$ is a mean-zero spatially statistically homogeneous random field, and $\varepsilon \ll 1$ is a small parameter measuring the strength of random fluctuations. As in the case of a particle in a random velocity field, and for weakly random Hamiltonian systems that we have considered in the previous sections, weak randomness will produce a non-trivial effect on time scales of the order $t \sim O(\varepsilon^{-1})$, and distances of the order $l \sim O(\varepsilon^{-1})$. We rescale accordingly (5.1) and arrive at

$$i\varepsilon\frac{\partial\phi_\varepsilon}{\partial t} + \frac{\varepsilon^2}{2}\Delta\phi_\varepsilon - \sqrt{\varepsilon}V\left(\frac{x}{\varepsilon}\right)\phi_\varepsilon = 0. \tag{5.2}$$

The initial data is now of the form $\phi_\varepsilon(0,x) = \phi_0(x/\varepsilon)$ but we will assume more generally that $\phi_\varepsilon(0,x)$ is an ε-oscillatory family, as we have discussed in Section 3. We will discuss in Section 5.2 what kind of behavior one should expect for ϕ_ε itself but here we first concentrate on what happens to the wave energy density. As in the previous situations we have considered, a convenient tool to analyze the wave energy density is by means of the Wigner transform

$$W_\varepsilon(t,x,k) = \int e^{ik\cdot y}\phi_\varepsilon(t,x-\frac{\varepsilon y}{2})\bar{\phi}_\varepsilon(t,x+\frac{\varepsilon y}{2})\frac{dy}{(2\pi)^d}. \tag{5.3}$$

A formal asymptotic analysis. In order to understand what we can expect as a limit for $W_\varepsilon(t,x,k)$ consider the evolution equation for the Wigner transform:

$$\frac{\partial W_\varepsilon}{\partial t} + k\cdot\nabla_x W_\varepsilon = \frac{1}{i\sqrt{\varepsilon}}\int e^{ip\cdot x/\varepsilon}\hat{V}(p)\left[W_\varepsilon(k-\frac{p}{2}) - W_\varepsilon(k+\frac{p}{2})\right]\frac{dp}{(2\pi)^d}. \tag{5.4}$$

As usual, we introduce a formal asymptotic expansion

$$W_\varepsilon(t,x,k) = \bar{W}(t,x,k) + \sqrt{\varepsilon}W_1(t,x,\frac{x}{\varepsilon},k) + \varepsilon W_2(t,x,\frac{x}{\varepsilon},k) + \dots,$$

with a deterministic leading order term $\bar{W}(t,x,k)$ and $W_k(t,x,z,k)$ random and stationary in z variable for all $k \geq 1$. Inserting this expansion into (5.4) gives in the leading order, with the fast variable $z = x/\varepsilon$, and a small parameter $\theta \ll 1$ that we will send to zero later,

$$k\cdot\nabla_z W_1 + \theta W_1 = \frac{1}{i}\int e^{ip\cdot z}\hat{V}(p)\left[\bar{W}(k-\frac{p}{2}) - \bar{W}(k+\frac{p}{2})\right]\frac{dp}{(2\pi)^d}.$$

Since $\bar{W}$ is independent of z, we have, for the Fourier transform of W_1 in z:

$$\hat{W}_1(t,x,p,k) = \frac{\hat{V}(p)}{(-k\cdot p + i\theta)}\left[\bar{W}(k-\frac{p}{2}) - \bar{W}(k+\frac{p}{2})\right]. \tag{5.5}$$

The next order terms give

$$\frac{\partial \bar{W}}{\partial t} + k\cdot\nabla_x\bar{W} + k\cdot\nabla_z W_2 = \frac{1}{i}\int e^{ip\cdot z}\hat{V}(p)\left[W_1(k-\frac{p}{2}) - W_1(k+\frac{p}{2})\right]\frac{dp}{(2\pi)^d}. \tag{5.6}$$

We take the expectation above, with the assumption that $\mathbb{E}(k\cdot\nabla_z W_2) = 0$, that is consistent with stationarity of W_2 in the z variable, to get

$$\frac{\partial \bar{W}}{\partial t} + k\cdot\nabla_x\bar{W} = \frac{1}{i}\int e^{ip\cdot z}\mathbb{E}\left\{\hat{V}(p)\left[W_1(k-\frac{p}{2}) - W_1(k+\frac{p}{2})\right]\right\}\frac{dp}{(2\pi)^d}. \tag{5.7}$$

After substituting from (5.5) we conclude that the right hand side of (5.7) equals

$$\frac{1}{i}\int e^{i(p+q)\cdot z}\mathbb{E}\left[\hat{V}(p)\hat{V}(q)\right]\frac{dpdq}{(2\pi)^{2d}}$$
$$\times\left[\frac{\bar{W}(k-p/2-q/2) - \bar{W}(k-p/2+q/2)}{(-(k-p/2)\cdot q + i\theta)} - \frac{\bar{W}(k+p/2-q/2) - \bar{W}(k+p/2+q/2)}{(-(k+p/2)\cdot q + i\theta)}\right].$$

Using the relation $\mathbb{E}\left[\hat{V}(p)\hat{V}(q)\right] = (2\pi)^d \hat{R}(p)\delta(p+q)$ we transform this expression into

$$
\begin{aligned}
&\tfrac{1}{i}\int \hat{R}(p)\left[\frac{\bar{W}(k)-\bar{W}(k-p)}{(k-p/2)\cdot p+i\theta}+\frac{\bar{W}(k)-\bar{W}(k+p)}{(k+p/2)\cdot p+i\theta}\right]\frac{dp}{(2\pi)^d}\\
=\;&\tfrac{1}{i}\int \hat{R}(p)[\bar{W}(k)-\bar{W}(k-p)]\left[\frac{1}{(k-p/2)\cdot p+i\theta}-\frac{1}{(k-p/2)\cdot p-i\theta}\right]\frac{dp}{(2\pi)^d}\\
=\;&\int \hat{R}(p)[\bar{W}(k-p)-\bar{W}(k)]\frac{2\theta}{[(k-p/2)\cdot p]^2+\theta^2}\frac{dp}{(2\pi)^d}\\
\to\;&\int \hat{R}(k-p)[\bar{W}(p)-\bar{W}(k)]\delta\left(\frac{|k|^2}{2}-\frac{|p|^2}{2}\right)\frac{dp}{(2\pi)^{d-1}},
\end{aligned}
$$

as $\theta \to 0$. We obtain therefore that $\bar{W}(t,x,k)$ satisfies the radiative transport equation

$$
\frac{\partial \bar{W}}{\partial t}+k\cdot\nabla_x\bar{W}=\int \hat{R}(k-p)[\bar{W}(p)-\bar{W}(k)]\delta\left(\frac{|k|^2}{2}-\frac{|p|^2}{2}\right)\frac{dp}{(2\pi)^{d-1}}. \tag{5.8}
$$

Rigorous convergence result. The rigorous justification of the radiative transport limit was first obtained by H. Spohn in [74], for short (but independent of ε) times by using diagrammatic expansions. This method was improved by L. Erdös and H.T. Yau in [37]. The result is as follows. Assume that $V(x)$ is a spatially statistically homogeneous random field of mean zero, and with a sufficiently smooth and rapidly decaying two-point correlation function $R(x)$. Let $\phi_\varepsilon(0,x)$ be an ε-oscillatory, compact at infinity family. Then $\mathbb{E}(W_\varepsilon(t,x,k))$ converges as $\varepsilon \to 0$, weakly in $\mathcal{S}'(\mathbb{R}^{2d})$, to the solution of the radiative transport equation (5.8).

We will discuss some aspects of the diagrammatic expansions in Section 5.2 but only in the simpler situation of time-dependent potentials. This avoid most of the technical difficulties of the proof in [37], and we refer the reader to that paper for details.

5.2. **Limits for the wave function.** Let us now consider solutions of the Schrödinger equation

$$
\begin{aligned}
&i\frac{\partial\phi(t,x)}{\partial t}+\frac{1}{2}\Delta\phi(t,x)-\sqrt{\varepsilon}V(t,x)\phi(t,x)=0,\;\; x\in\mathbb{R}^d,\\
&\phi(0,x)=\phi_0(x),
\end{aligned} \tag{5.9}
$$

with a time-dependent random potential $V(t,x)$ in the spatial dimension $d \geq 2$. (Most of the results presented in this paper in the setting of time dependent potentials also extend to dimension $d=1$ while the results obtained for time-independent potentials do not because of wave localization effects. To avoid confusion, we therefore restrict ourselves to the case $d \geq 2$.) The goal of the present section is to understand the behavior of $\phi(t,x)$ itself after propagation over long distances, rather than for the Wigner transform.

As usual, we recast (5.9) as an equation for the rescaled function $\phi_\varepsilon(t,x) = \phi(t/\varepsilon, x/\varepsilon)$:

$$
\begin{aligned}
&i\varepsilon\frac{\partial\phi_\varepsilon}{\partial t}+\frac{\varepsilon^2}{2}\Delta\phi_\varepsilon-\sqrt{\varepsilon}V(\frac{t}{\varepsilon},\frac{x}{\varepsilon})\phi_\varepsilon=0,\\
&\phi_\varepsilon(0,x)=\phi_0(x/\varepsilon).
\end{aligned} \tag{5.10}
$$

In particular, we have $\hat{\phi}_\varepsilon(0,\xi) = \varepsilon^d \hat{\phi}_0(\varepsilon\xi)$. We assume that the spatial power spectrum has the form

$$\tilde{R}(t,k) = e^{-\mathfrak{g}(k)|t|}\hat{R}(k), \tag{5.11}$$

where $\hat{R}(k) \in L^1(\mathbb{R}^d)$, and

$$\tilde{R}(t,k) = \int e^{-ik\cdot x} R(t,x)dx.$$

The space-time power energy spectrum is then

$$\hat{R}(\omega,k) = \frac{2\mathfrak{g}(k)\hat{R}(k)}{\omega^2 + \mathfrak{g}^2(k)}. \tag{5.12}$$

In order to formulate the main result let

$$D(p,\xi) = \frac{2\hat{R}(p)}{(2\pi)^d[\mathfrak{g}(p) - i(\xi\cdot p - |p|^2/2)]} \tag{5.13}$$

and

$$D(\xi) = \int D(p,\xi)dp = 2\int \frac{\hat{R}(p)}{\mathfrak{g}(p) - i(\xi\cdot p - |p|^2/2)}\frac{dp}{(2\pi)^d}. \tag{5.14}$$

Let also

$$\mathcal{L}F(\xi) := \int D(p,\xi)[F(p) - F(\xi)]dp, \tag{5.15}$$

and $\widehat{W}(t,\xi)$ be the solution of the equation

$$\begin{cases} \partial_t \widehat{W}(t,\xi) = \mathcal{L}\widehat{W}(t,\xi), \\ \widehat{W}(0,\xi) = |\hat{\phi}_0(\xi)|^2. \end{cases} \tag{5.16}$$

Note that (5.16) is simply the integrated in x form of the radiative transport equation.

Theorem 5.1. *Assume that $V(t,x)$ is a spatially homogeneous mean-zero Gaussian random field with the two-point correlation function $R(t,x)$ and the spatial power spectrum $\tilde{R}(t,k)$ of the form (5.11) with*

$$\int \frac{\hat{R}(p)dp}{\mathfrak{g}(p)} < +\infty. \tag{5.17}$$

Define

$$\hat{\zeta}_\varepsilon(t,\xi) = \frac{1}{\varepsilon^d}\hat{\phi}_\varepsilon(t,\xi/\varepsilon)e^{i|\xi|^2 t/(2\varepsilon)}, \tag{5.18}$$

where $\phi_\varepsilon(t,x)$ is the solution of (5.10). Then, for each $t\in\mathbb{R}$ and $\xi\in\mathbb{R}^d$ fixed, $\hat{\zeta}_\varepsilon(t,\xi)$ converges in law, as $\varepsilon\to 0$, to

$$\hat{\zeta}(t,\xi) = e^{-tD_\xi/2}\hat{\phi}_0(\xi) + Z(t,\xi) \tag{5.19}$$

Here $Z(t,\xi)$ is a centered, complex valued Gaussian random variable, whose variance equals

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

Let us recall that a random variable $Z = X + iY$ is a centered complex Gaussian if X and Y are mean-zero Gaussian independent random variables with $\mathbb{E}(X^2) = \mathbb{E}(Y^2)$.

Note that Theorem 5.1 implies, in particular, that $\mathbb{E}\left\{|\hat{\zeta}(t,\xi)|^2\right\} = \widehat{W}(t,\xi)$ is the solution of (5.16), as would be expected from the usual kinetic theory for waves.

However, this theorem gives a much more precise information on the limit of the whole random field $\hat{\zeta}_\varepsilon(t,\xi)$ and not just its second absolute moment. It would be very interesting to obtain a similar result for time-independent potentials, but that would be a much more difficult problem.

5.3. **Convergence of the expectation of the wave function.** We will not present the full proof of Theorem 5.1 here but rather describe the proof of the convergence

$$\mathbb{E}(\hat{\zeta}_\varepsilon(t,\xi)) \to e^{-tD_\xi}\hat{\phi}_0(\xi), \tag{5.20}$$

as this will allow us to introduce at least some (simpler) aspects of the diagrammatic techniques.

The Duhamel expansion. We re-write (5.10) as an integral in time equation

$$\begin{aligned}\hat{\phi}_\varepsilon(t,\xi) = \hat{\phi}_0(\xi)e^{-i\varepsilon|\xi|^2 t/2} + \tfrac{1}{i\sqrt{\varepsilon}}\int_0^t \int \tfrac{\hat{V}(s_1/\varepsilon, dp_1)}{(2\pi)^d} \\ \times\hat{\phi}_\varepsilon(s_1, \xi - \tfrac{p_1}{\varepsilon})e^{-i\varepsilon|\xi|^2(t-s_1)/2}ds_1.\end{aligned}$$

Hence, the function $\hat{\zeta}_\varepsilon(t,\xi)$ solves

$$\hat{\zeta}_\varepsilon(t,\xi) = \hat{\phi}_0(\xi) + \frac{1}{i\sqrt{\varepsilon}}\int_0^t \int \frac{\hat{V}(s_1/\varepsilon, dp_1)}{(2\pi)^d}\hat{\zeta}_\varepsilon(s_1, \xi - p_1)e^{i(|\xi|^2 - |\xi-p_1|^2)s_1/(2\varepsilon)}ds_1, \tag{5.21}$$

as $\hat{\zeta}(0,\xi) = \hat{\phi}_0(\xi)$. Iterating (5.21) leads to an infinite series expansion for $\hat{\zeta}_\varepsilon(t,\xi)$:

$$\hat{\zeta}_\varepsilon(t,\xi) = \sum_{n=0}^\infty \hat{\zeta}_n^\varepsilon(t,\xi), \tag{5.22}$$

with the individual terms of the form

$$\begin{aligned}\hat{\zeta}_n^\varepsilon(t,\xi) = \left[\tfrac{1}{i\sqrt{\varepsilon}(2\pi)^d}\right]^n \int\limits_{\Delta_n(t)} d\mathbf{s}^{(n)} \int \hat{V}(\tfrac{s_1}{\varepsilon}, dp_1)\ldots\hat{V}(\tfrac{s_n}{\varepsilon}, dp_n) \\ \times\hat{\phi}_0(\xi - p_1 - \cdots - p_n)e^{iG_n(\mathbf{s}^{(n)},\mathbf{p}^{(n)})/\varepsilon},\end{aligned} \tag{5.23}$$

with the phase

$$\begin{aligned}G_n(\mathbf{s}^{(n)},\mathbf{p}^{(n)}) &= \sum_{k=1}^n (|\xi - p_1 - \cdots - p_{k-1}|^2 - |\xi - p_1 - \cdots - p_k|^2)\tfrac{s_k}{2} \\ &= A_n(\mathbf{s}^{(n)},\mathbf{p}^{(n)}) - B_n(\mathbf{s}^{(n)},\mathbf{p}^{(n)}).\end{aligned}$$

Here we use the notation $p_0 = 0$, $\mathbf{s}^{(n)} = (s_1,\ldots,s_n) \in \mathbb{R}^n$, $\mathbf{p}^{(n)} = (p_1,\ldots,p_n) \in \mathbb{R}^{nd}$, so that $d\mathbf{s}^{(n)} = ds_1 ds_2\ldots ds_n$. We have also split the phase into

$$\begin{aligned}A_n(\mathbf{s}^{(n)},\mathbf{p}^{(n)}) = \sum_{m=1}^n (\xi\cdot p_m)s_m,\ B_n(\mathbf{s}^{(n)},\mathbf{p}^{(n)}) \\ = \sum_{m=1}^n s_m p_m \cdot \left(\textstyle\sum_{j=1}^{m-1} p_j\right) + \tfrac{1}{2}\sum_{m=1}^n s_m|p_m|^2.\end{aligned} \tag{5.24}$$

Finally, $\Delta_n(t)$ denotes the time simplex

$$\Delta_n(t) = \{(s_1, s_2, \ldots, s_n) :\ 0 \le s_n \le s_{n-1} \le \cdots \le s_1 \le t\}.$$

The next proposition shows that the series (5.22) converges almost surely and, moreover, one can take the expectation term-wise for $\varepsilon > 0$ fixed. This allows us to work with term-wise estimates for each $\mathbb{E}(\hat{\zeta}_\varepsilon^n)$ separately.

Proposition 5.2. *(i) The series* (5.22) *for the function* $\hat{\zeta}_\varepsilon(t,\xi)$ *converges almost surely for all values of* $\gamma, \varepsilon \in (0,1]$ *and* $\phi_0 \in C_c^\infty(\mathbb{R}^d)$. *(ii) Moreover, for each* $(t,\xi) \in \mathbb{R}^{1+d}$ *fixed, we have*

$$\mathbb{E}\hat{\zeta}_\varepsilon(t,\xi) = \sum_{n=0}^{\infty} \mathbb{E}\hat{\zeta}_n^\varepsilon(t,\xi). \tag{5.25}$$

Proof. We may assume without loss of generality that $\varepsilon = 1$. Using an elementary result from analysis the conclusion of the lemma follows, provided we can show that $\sum_{n=0}^{\infty}[\mathbb{E}|\hat{\zeta}_n^\varepsilon(t,\xi)|^2]^{1/2} < +\infty$.

Note that

$$\mathbb{E}|\hat{\zeta}_n^\varepsilon(t,\xi)|^2 = \tfrac{1}{(2\pi)^{2nd}} \int\limits_{\Delta_n(t)} d\mathbf{s}^{(n)} \int\limits_{\Delta_n(t)} d\tilde{\mathbf{s}}^{(n)} \int \mathbb{E}\left[\prod_{k=1}^{n} \hat{V}(s_k, dp_k) \prod_{k=1}^{n} \hat{V}^*(\tilde{s}_k, d\tilde{p}_k)\right]$$
$$\times \hat{\phi}_0(\xi - \sum_{j=1}^{n} p_j)\hat{\phi}_0^*(\xi - \sum_{j=1}^{n} \tilde{p}_j) e^{iG_n(\mathbf{s}^{(n)},\mathbf{p}^{(n)})} e^{-iG_n(\tilde{\mathbf{s}}^{(n)},\tilde{\mathbf{p}}^{(n)})}$$

for some constant $C > 0$ independent of n. The random elements $\{\hat{V}(s_k, dp_k), \hat{V}^*(\tilde{s}_l, d\tilde{p}_l),\ k.l = 1\ldots, d\}$ are jointly Gaussian. The moment of a product of an even number of Gaussians $\{X_i,\ i = 1, \ldots, 2n\}$ can be computed according to the formula

$$\mathbb{E}\left[\prod_{i=1}^{2n} X_i\right] = \sum_{(i,j)} \mathbb{E}\left[X_i X_j\right].$$

The summation extends over all partitions of the set $\{1, \ldots, 2n\}$ into two element subsets (i,j) - the so called *pairings*. We use the relations

$$\mathbb{E}\left[\hat{V}(s_k, dp_k)\hat{V}(\tilde{s}, d\tilde{p}_l)\right] = e^{-\mathfrak{g}(p_k)|s_k - s_l|}\hat{R}(p_k)\delta(p_k + p_l)dp_k dp_l \tag{5.26}$$

and $\hat{V}^*(s, dp) = \hat{V}(s, -dp)$ (because the potential is real valued).

Since there are $(2n-1)!!$ pairings of $\{1, \ldots, 2n\}$ and

$$|\mathbb{E}\left[\hat{V}(s_k, dp_k)\hat{V}(\tilde{s}, d\tilde{p}_l)\right]| \leq \hat{R}(p_k)\delta(p_k + p_l)dp_k dp_l$$

we conclude that the right hand side of (5.26) is estimated by

$$\frac{t^n(2n-1)!!}{n!^2(2\pi)^{2nd}}\left[\int \hat{R}(dp)\right]^n \|\hat{\phi}_0\|_\infty^2 \leq \frac{C^n}{n!}$$

for some constant $C > 0$ independent of n and the conclusion of the proposition follows. □

Convergence of the expectation $\mathbb{E}(\hat{\zeta}_\varepsilon(t,\xi))$. We now prove (5.20). The initial step in the proof is the following uniform bound for the individual terms of (5.25).

Proposition 5.3. *For all* $T > 0$, $n \geq 0$ *and all* $\xi \in \mathbb{R}^d \setminus \{0\}$ *there exists a constant* $C(T)$ *such that*

$$\sup_{t\in[0,T]} |\mathbb{E}\hat{\zeta}_n^\varepsilon(t,\xi)| \leq \frac{C^n(T;\xi)}{n!} \tag{5.27}$$

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

As a consequence, we may interchange the limit $\varepsilon \downarrow 0$ and the summation in n.

Corollary 5.4. *We have*

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\hat{\zeta}_\varepsilon(t,\xi)=\sum_{n=0}^{\infty}\lim_{\varepsilon\downarrow 0}\mathbb{E}\hat{\zeta}_n^\varepsilon(t,\xi), \tag{5.28}$$

for all $t\in\mathbb{R}$ and $\xi\in\mathbb{R}^d\setminus\{0\}$.

Next, we identify the limit of the individual terms in the right side of (5.28).

Proposition 5.5. *We have $\mathbb{E}\hat{\zeta}_n^\varepsilon(t,\xi)=0$ when n is odd and*

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\hat{\zeta}_{2n}^\varepsilon(t,\xi)=\frac{1}{n!}\left(\frac{-tD(\xi)}{2}\right)^n\hat{\phi}_0(\xi) \tag{5.29}$$

for all $n\ge 0$, $t\in\mathbb{R}$ and $\xi\neq 0$.

This together with (5.28) implies convergence of the expectation in (5.20). We will only present the proof of Proposition 5.5 that shows how the diagrams are summed.

Proof of Proposition 5.5. Let us introduce some terminology: the pairing $(1,2),\dots,(2n-1,2n)$ shall be called *a time-ordered pairing*. For a given pairing $\mathcal{F}$ we let

$$\mathcal{I}_\varepsilon(t;\mathcal{F}):=\int\limits_{\Delta_{2n}(t)}d\mathbf{s}^{(2n)}\int d\mathbf{p}^{(2n)}\prod_{(k,l)\in\mathcal{F}}e^{-\mathfrak{g}(p_k)|s_k-s_l|/\varepsilon}\delta(p_k+p_l)\hat{R}(p_k). \tag{5.30}$$

One can verify that

$$\mathcal{I}(\mathcal{F})=\limsup_{\varepsilon\downarrow 0}\sup_{t\in[0,T]}\varepsilon^{-n}\mathcal{I}_\varepsilon(t;\mathcal{F})<+\infty, \tag{5.31}$$

for any pairing $\mathcal{F}$. We will now show that $\mathcal{I}(\mathcal{F})=0$ if $\mathcal{F}$ is not a time-ordered pairing, and then identify the actual limit of $\varepsilon^{-n}\mathcal{I}_\varepsilon(\mathcal{F})$ for the time-ordered pairings completing the proof of Proposition 5.5. We start with non time-ordered pairings.

Lemma 5.6. *Suppose that $\mathcal{F}$ is not a time-ordered pairing. Then,*

$$\lim_{\varepsilon\downarrow 0}\sup_{t\in[0,T]}\varepsilon^{-n}\mathcal{I}_\varepsilon(t;\mathcal{F})=0, \tag{5.32}$$

for any $T>0$.

Proof. We will verify the statement of the lemma only for $n=2$, the general case is done by induction [9]. We have to consider then two pairings $\mathcal{F}_1=\{(1,3),(2,4)\}$ and $\mathcal{F}_2=\{(1,4),(2,3)\}$. Start with the first one. Suppose that $\kappa\in(0,1)$ and consider the sets of the following times: $A_1=[|s_1-s_3|\ge\varepsilon^\kappa]$ and $A_2=[|s_2-s_4|\ge\varepsilon^\kappa]$, as well as $A_3=A_1^c\cup A_2^c$. Consider the expressions

$$I_i(\varepsilon)=\int\limits_{\Delta_4(t)\cap A_i}ds_1\dots ds_4\int dp_1dp_2\exp-[\mathfrak{g}(p_1)(s_1-s_3)$$
$$+\mathfrak{g}(p_2)(s_2-s_4)]/\varepsilon\hat{R}(p_1)\hat{R}(p_2),$$

for $i=1,2,3$, then

$$\mathcal{I}_\varepsilon(t;\mathcal{F}_1)\le\sum_{i=1}^3 I_i(\varepsilon).$$

We will see that $I_1(\varepsilon)$ and $I_2(\varepsilon)$ are small because the integrand is exponentially small in ε, while $I_3(\varepsilon)$ vanishes because the domain of integration is small. Indeed, observe that

$$I_1(\varepsilon) \le \int_0^t \int_0^t ds_1 ds_3 \int_{\mathbb{R}} \int_{\mathbb{R}} ds_2 ds_4 \int dp_1 dp_2 e^{-\varepsilon^{\kappa-1}\mathfrak{g}(p_1)/2} \\ e^{-[\mathfrak{g}(p_1)|s_1-s_3|+\mathfrak{g}(p_2)|s_2-s_4|]/(2\varepsilon)} \hat{R}(p_1)\hat{R}(p_2)$$

$$= (2t\varepsilon)^2 \int e^{-\varepsilon^{\kappa-1}\mathfrak{g}(p_1)/2} \frac{\hat{R}(p_1)dp_1}{\mathfrak{g}(p_1)} \int \frac{\hat{R}(p_2)dp_2}{\mathfrak{g}(p_2)}$$

and it follows from the Lebesgue dominated convergence theorem that

$$\lim_{\varepsilon\downarrow 0} \sup_{t\in[0,T]} \varepsilon^{-2} I_1(\varepsilon) = 0. \tag{5.33}$$

Similarly one can prove that (5.33) holds for $I_2(\varepsilon)$. On the other hand, we note that if $0 \le s_1 - s_3 \le \varepsilon^\kappa$ and $0 \le s_2 - s_4 \le \varepsilon^\kappa$ (so that $(s_1, s_2, s_3, s_4) \in A_3$) then (since $0 \le s_3 \le s_2$), we have $0 \le s_1 - s_4 \le 2\varepsilon^\kappa$ as well. Hence,

$$I_3(\varepsilon) \le Ct\varepsilon^{3\kappa}$$

and (5.33) follows for $I_3(\varepsilon)$, provided that $\kappa > 2/3$. We have shown in this way that

$$\lim_{\varepsilon\downarrow 0} \sup_{t\in[0,T]} \varepsilon^{-2}\mathcal{I}(t;\mathcal{F}_1) = 0.$$

A similar argument also yields an analogous statement for $\mathcal{I}(t;\mathcal{F}_2)$. □

The contribution of the time-ordered pairings. The last step in the proof of Proposition 5.5 is to consider the contribution of the time-ordered pairings. We have shown so far that

$$\lim_{\varepsilon\downarrow 0} \mathbb{E}\hat{\zeta}^\varepsilon_{2n}(t,\xi) = J_n(t,\xi), \tag{5.34}$$

where

$$J_n(t,\xi) = \hat{\phi}_0(\xi) \lim_{\varepsilon\downarrow 0} \frac{(-1)^n}{[\varepsilon(2\pi)^d]^n} \int_{\Delta_{2n}(t)} d\mathbf{s}^{(2n)} \int d\mathbf{p}^{(2n)} \prod_{k=1}^n \hat{R}(p_{2k-1})\delta(p_{2k-1}+p_{2k}) \\ \times e^{-\mathfrak{g}(p_{2k-1})(s_{2k-1}-s_{2k})/\varepsilon} \exp\left\{iG_n(\mathbf{s}^{(2n)},\mathbf{p}^{(2n)})/\varepsilon\right\}$$

where $G_n(\mathbf{s}^{(2n)},\mathbf{p}^{(2n)})$ is given by (5.24). For the time-ordered pairing, taking into account the delta-functions, we have

$$G_n(\mathbf{s}^{(2n)},\mathbf{p}^{(2n)}) = \sum_{m=1}^n \left[\xi\cdot p_{2m-1} - \frac{1}{2}|p_{2m-1}|^2\right](s_{2m-1}-s_{2m}).$$

Hence, (5.35) can be written as

$$J_n(t,\xi) = \hat{\phi}_0(\xi) \lim_{\varepsilon\downarrow 0} \frac{(-1)^n}{[\varepsilon(2\pi)^d]^n} \int_{\Delta_{2n}(t)} d\mathbf{s}^{(2n)} \int d\mathbf{p}^{(2n)} \prod_{k=1}^n \hat{R}(p_{2k-1}) \\ \times\delta(p_{2k-1}+p_{2k}) e^{-Q(p_{2k-1})(s_{2k-1}-s_{2k})/\varepsilon}, \tag{5.35}$$

with

$$Q(p) = \mathfrak{g}(p) - i\left(\xi\cdot p - \frac{1}{2}|p|^2\right).$$

Changing variables $s'_{2m} = (s_{2m-1} - s_{2m})/\varepsilon$ we obtain, after dropping the primes:

$$\begin{aligned} J_n(t,\xi) = \hat{\phi}_0(\xi) \lim_{\varepsilon\downarrow 0} \frac{(-1)^n}{(2\pi)^{nd}} \int_0^t ds_1 \int_0^{s_1/\varepsilon} ds_2 \int_0^{s_1-\varepsilon s_2} ds_3 \dots \int_0^{s_{2n-3}-\varepsilon s_{2n-2}} ds_{2n-1} \\ \int_0^{s_{2n-1}/\varepsilon} ds_{2n} \int \dots \int \prod_{k=1}^n \hat{R}(p_{2k-1}) dp_{2k-1} \prod_{k=1}^n e^{-Q(p_{2k-1})s_{2k}}. \end{aligned} \tag{5.36}$$

One can now compute the limit in (5.36):

$$\begin{aligned} J_n(t,\xi) &= \hat{\phi}_0(\xi) \frac{(-1)^n}{(2\pi)^{nd}} \int_0^t ds_1 \int_0^{s_1} ds_3 \dots \int_0^{s_{2n-3}} ds_{2n-1} \\ &\quad \times \int \dots \int \prod_{k=1}^n \frac{\hat{R}(p_{2k-1})}{Q(p_{2k-1})} dp_{2k-1} \\ &= \hat{\phi}_0(\xi) \frac{(-1)^n t^n}{(2\pi)^{nd} n!} \left(\int \frac{\hat{R}(p)}{Q(p)} dp \right)^n \\ &= \hat{\phi}_0(\xi) \tfrac{(-tD(\xi))^n}{2^n n!}, \end{aligned}$$

where $D(\xi)$ is given by (5.14). This completes the proof of Proposition 5.5. □

Remark. This is the simplest example of summation of diagrams in such a context. Rapid time decorrelation made the contribution of non-time-ordered diagrams small in a particularly simple way. In the more difficult case when the medium is time-independent they are still small, but only because of the oscillatory phase, that we have completely discarded here. A careful estimation of the contribution of the oscillatory phase is much more delicate [37] than in the case discussed here.

5.4. **A simplified model: Itô-Schrödinger.** The simplest route to radiative transfer models starts with (heavily) simplifying the wave model. In the paraxial approximation to acoustic wave equations and under some additional assumptions, we model wave propagation as the following stochastic partial differential equation, called the *Itô-Schrödinger* equation[1]:

$$d\psi_\eta(z,x) = \frac{1}{2}\big(i\eta\Delta_x - R(0)\big)\psi_\eta(z,x)dz + i\psi_\eta(z,x)B\Big(\frac{x}{\eta}, dz\Big). \tag{5.37}$$

Here, $\eta > 0$ and $B(x,z)$ is a Wiener process, defined over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, whose covariance function equals

$$\mathbb{E}\{B(x,z)B(y,z')\} = R(x-y) z \wedge z', \tag{5.38}$$

where $\mathbb{E}$ is mathematical expectation corresponding to $\mathbb{P}$, $z \wedge z' = \min(z,z')$ and $R(x)$ is the covariance function of the random medium. A rigorous passage from the wave equation to (5.37) can be found in [1] when $d = 2$ and in stratified media. Here, we simply adopt it as a simplified model. Note that $z \in \mathbb{R}$ and $x \in \mathbb{R}^{d-1}$ are both spatial variables (this is a one-frequency equation) but z being the predominant direction of propagation plays the role of "time" in the Schrödinger equation, hence the random potential in (5.37) is "white in time".

[1]Here we will denote the small parameter by η rather than ε, to avoid confusion with the symbol ϵ that will be reserved for quantities taking values $\pm$.

5.4.1. *A kinetic model.* The radiative transfer equations in the Itô-Schrödinger regime are obtained in the high-frequency asymptotics of (5.37). The appropriate tool in the analysis of such equations is, as usual, the Wigner transform of the wave function, defined as

$$W_\eta[\psi_\eta](z,x,k) = W_\eta(z,x,k) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{ik\cdot y}\psi_\eta\Big(z,x-\frac{\eta y}{2}\Big)\overline{\psi_\eta}\Big(z,x+\frac{\eta y}{2}\Big)dy. \tag{5.39}$$

The main result is the following: under appropriate conditions on the initial condition ψ_η^0, the *ensemble average* of the Wigner transform $a_\eta := \mathbb{E}\{W_\eta\}$ converges weakly to the solution a of the following radiative transfer equation (or linear Boltzmann equation):

$$\Big(\frac{\partial}{\partial z} + k\cdot\nabla_x + R_0 - \mathcal{Q}\Big)a(z,x,k) = 0, \qquad a(0,x,k) = a_0(x,k), \tag{5.40}$$

where a_0 is the limit of the ensemble average of the Wigner transform of the initial condition ψ_η^0, $R_0 := (2\pi)^d R(\mathbf{0})$ and the scattering operator $\mathcal{Q}$ acts as

$$(\mathcal{Q}a)(z,x,k) = \int_{\mathbb{R}^d}\hat{R}(k-k')a(z,x,k')dk'.$$

Here, $\hat{R}$ denotes the Fourier transform of R with the convention

$$\hat{R}(k) = \int_{\mathbb{R}^d} e^{-ik\cdot x}R(x)dx.$$

Since $R(x)$ is a correlation function, $\hat{R}(k)$ is non-negative by Bochner's theorem. The derivation of (5.40) from the Itô-Schrödinger equation (5.37) is immediate since moments of the wavefunction satisfy closed-form equations – this is a huge advantage of the Itô-Schrödinger model. Starting from (5.37) and writing the stochastic equation for the Wigner transform, an application of the Itô formula yields that a_η solves (5.40) with an initial condition $a_{\eta 0} := \mathbb{E}\{W_\eta[\psi_\eta^0]\}$, see for instance [69]. It then suffices to pass to the limit in the initial condition to obtain the convergence of a_η to a. This eliminates the complexity of the derivation of the kinetic limit and allows us to consider other non-trivial issues related to the kinetic limit: its statistical stability, error estimates and so on.

The above kinetic equation is similar to the radiative transfer equation we have discussed for a time-independent random potential for the Schrödinger equation but the scattering operator is now replaced by $R_0 - \mathcal{Q}$. The main difference is that scattering is not *elastic* as $|k|$ is not preserved through scattering. This leads to a very different qualitative behavior for solutions in the long time limit.

5.4.2. *Self-averaging for the Itô-Schrödinger model.* In the Itô-Schrödinger regime, the convergence of W_η to its average can be made precise so as to obtain information on the rate of convergence or on the size of the averaging domain that is needed to obtain statistical stability (typically the size of the support of the test function φ). This is rendered possible by the fact that the *scintillation function* J_η (or covariance function), defined as

$$J_\eta(z,x,k,y,p) = \mathbb{E}\{W_\eta(z,x,k)W_\eta(z,y,p)\} - \mathbb{E}\{W_\eta(z,x,k)\}\mathbb{E}\{W_\eta(z,y,p)\}, \tag{5.41}$$

solves the closed-form equation

$$\Big(\frac{\partial}{\partial z} + \mathcal{T}_2 + 2R_0 - \mathcal{Q}_2 - \mathcal{K}_\eta\Big)J_\eta = \mathcal{K}_\eta a_\eta \otimes a_\eta, \tag{5.42}$$

equipped with vanishing initial conditions $J_\eta(0,x,k,y,p)=0$ when the initial condition of the Schrödinger equation is deterministic. This is, once again, obtained by an application of the Itô formula, see [2]. Here, we have defined

$$\begin{aligned}
\mathcal{T}_2 &= k\cdot\nabla_x + p\cdot\nabla_y,\\
\mathcal{Q}_2 h &= \int_{\mathbb{R}^{2d}}\Big[\hat{R}(k-k')\delta(p-p')+\hat{R}(p-p')\delta(k-k')\Big]h(x,k',y,p')dk'dp',\\
\mathcal{K}_\eta h &= \sum_{\epsilon_i,\epsilon_j=\pm 1}\epsilon_i\epsilon_j\int_{\mathbb{R}^{2d}}\hat{R}(u)e^{i\frac{(x-y)\cdot u}{\eta}}h\left(x,k+\epsilon_i\frac{u}{2},y,p+\epsilon_j\frac{u}{2}\right)du.
\end{aligned} \tag{5.43}$$

Above, δ is the Dirac delta-function. The analysis of (5.42) and of the highly oscillating operator $\mathcal{K}_\eta$ shows that J_η converges weakly to zero, which implies convergence of W_η in probability thanks to the Chebyshev inequality

$$\mathbb{P}\Big(|\langle W_\eta(z),\varphi\rangle-\langle a_\eta(z),\varphi\rangle|\geq\varepsilon\Big)\leq\frac{1}{\varepsilon^2}\langle J_\eta(z),\varphi\otimes\varphi\rangle,$$

with $\varphi\otimes\varphi(x,k,y,p):=\varphi(x,k)\varphi(y,p)$. Note that convergence is obtained weakly in space. We do not expect to obtain convergence point-wise in space as the energy density needs to be averaged over an area that is large compared to the wavelength. Precisely, over how large an array the energy needs to be averaged to be self-averaging also depends on the structure of the initial condition and is treated in detail in the following section.

5.4.3. *Main results on self-averaging.* To be consistent with the usual notation for the time-dependent Schrödinger equation, we relabel the variable z as t. We assume that the initial condition ψ_η^0 is deterministic (i.e., independent of the random medium) and uniformly bounded with respect to η in $L^2(\mathbb{R}^d)$. We assume that our random medium has sufficiently short range correlations so that $\hat{R}\in L^1(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d)$. In such a setting, it is proved in [33] that (5.37) admits a unique solution $\psi_\eta(t,x,\omega)\in\mathcal{C}^0([0,\infty),L^2(\mathbb{R}^d))$, $\mathbb{P}$ a.e., such that

$$\|\psi_\eta(t,\cdot)\|_{L^2(\mathbb{R}^d)}\leq\|\psi_\eta^0\|_{L^2(\mathbb{R}^d)}\leq C,\quad\forall\, t\geq 0,$$

with probability one for some constant C independent of η. Moreover, we assume that ψ_η^0 is deterministic so, in consequence, ψ_η admits moments of arbitrary order so that its Wigner transform and related scintillation function are well-defined. Also $a_{\eta 0}:=\mathbb{E}\{W_\eta[\psi_\eta^0]\}=W_\eta[\psi_\eta^0]$, where W_η is defined in (5.39).

Let $\mathcal{F}a_{\eta 0}$ be the Fourier transform of $a_{\eta 0}$ in both variables x and k, and $\mathcal{F}_x a_{\eta 0}$ (resp. $\mathcal{F}_k a_{\eta 0}$) be its partial Fourier transform with respect to x (resp. k). Two important quantities are the L^1 norms of $\mathcal{F}_x a_{\eta 0}$ and $\mathcal{F}_k a_{\eta 0}$. Denoting by $a\lesssim b$ the inequality $a\leq Cb$, where $C>0$ is some universal constant, this leads us to make the following hypotheses on $a_{\eta 0}$:

Hypotheses H: $\mathcal{F}\nabla_x^p a_{\eta 0}\in L^\infty(\mathbb{R}^{2d})$, $\mathcal{F}_x\nabla_x^p a_{\eta 0}\in L^1(\mathbb{R}^{2d})$, $\mathcal{F}_k\nabla_x^p a_{\eta 0}\in L^1(\mathbb{R}^{2d})$, for $p=0$ or 1 (with the convention that $\nabla_x^0 a_{\eta 0}:=a_{\eta 0}$) with the following estimates, for $0\leq\alpha\leq 1$ and $0\leq\beta\leq 1$:

$$\|\mathcal{F}\nabla_x a_{\eta 0}\|_{L^\infty(\mathbb{R}^{2d})}\lesssim\eta^{-\alpha},$$

$$\|\mathcal{F}_x\nabla_x^p a_{\eta 0}\|_{L^1(\mathbb{R}^{2d})}\lesssim\eta^{-(d+p)\alpha}\quad\text{ and }\quad\|\mathcal{F}_k\nabla_x^p a_{\eta 0}\|_{L^1(\mathbb{R}^{2d})}\lesssim\eta^{-d\beta-p\alpha}.$$

For instance, when $\psi_\eta^0 \in \mathcal{S}(\mathbb{R}^d)$, it follows from

$$\mathcal{F}_x a_{\eta 0}(u,p) = \frac{1}{\eta^d}\mathcal{F}\psi_{\eta 0}\left(\frac{p}{\eta}+\frac{u}{2}\right)\overline{\mathcal{F}\psi_{\eta 0}}\left(\frac{p}{\eta}-\frac{u}{2}\right),$$
$$\mathcal{F}_k a_{\eta 0}(x,\xi) = \psi_{\eta 0}\left(x+\frac{\eta}{2}\xi\right)\overline{\psi_{\eta 0}}\left(x-\frac{\eta}{2}\xi\right),$$

that $\mathcal{F}\nabla_x^p a_{\eta 0} \in L^\infty(\mathbb{R}^{2d})$, $\mathcal{F}_x\nabla_x^p a_{\eta 0} \in L^1(\mathbb{R}^{2d})$, and $\mathcal{F}_k\nabla_x^p a_{\eta 0} \in L^1(\mathbb{R}^{2d})$ for $p=0$ or 1, though the norms are not bounded uniformly in η. The relevance of the above hypothesis is better explained by looking at the following examples.

Typical initial conditions. Let us consider initial conditions $\psi_\eta(x,0)$ oscillating at frequencies of order η^{-1} and with a spatial support of size η^α for $0 \le \alpha \le 1$. The parameter α quantifies the macroscopic concentration of the initial condition. The simplest example is a modulated plane wave of the form:

$$\psi_\eta^{(1)}(x) = \frac{1}{\eta^{\frac{d\alpha}{2}}}\chi\Big(\frac{x-x_0}{\eta^\alpha}\Big)e^{i\frac{(x-x_0)\cdot k_0}{\eta}}, \tag{5.44}$$

where $\chi \in \mathcal{S}(\mathbb{R}^d)$. The direction of propagation is given by k_0. Note that the above sequence of initial conditions is indeed uniformly bounded in $L^2(\mathbb{R}^d)$, and that the corresponding Wigner transform is

$$a_{\eta 0}(x,k) = \frac{1}{\eta^d}a_0\left(\frac{x-x_0}{\eta^\alpha},\frac{k-k_0}{\eta^{1-\alpha}}\right), \tag{5.45}$$

where $a_0(x,k)$ is the Wigner transform of the rescaled initial condition $\psi_1^{(1)}$. Such an initial condition then satisfies hypotheses **H** with $\beta = 1-\alpha$. The parameter α measures the concentration of the initial conditions in the spatial variables while β measures that in the momentum variables. We restrict α and β to be less than one to ensure that η^{-1} is the highest frequency in the problem. Allowing for higher frequencies while still considering a Wigner transform at the frequency η^{-1} will lead to vanishing limiting Wigner transforms and would be of little interest for then energy is lost when passing to the limit, see e.g. [43, 60].

As another example of initial conditions, consider

$$\psi_\eta^{(2)}(x) = \frac{1}{\eta^{\frac{(d-1)\alpha+1}{2}}}\chi\Big(\frac{x}{\eta^\alpha}\Big)J_0\Big(\frac{|k_0||x|}{\eta}\Big), \tag{5.46}$$

where J_0 is the order 0 Bessel function of the first kind. Such an initial condition is supported in the Fourier domain in the vicinity of wavenumbers k such that $|k| = |k_0|$ so that $\psi_\eta^{(2)}$ emits radiation isotropically at wavenumber $|k_0|$; see [13, 14] for more details. We again verify that the above sequence of initial conditions is indeed uniformly bounded in $L^2(\mathbb{R}^d)$ and satisfies **H** with $\alpha = 1-\beta$. For this, we use that $J_0(z) = (2/\pi z)^{1/2}\cos(z-\frac{\pi}{4}) + \mathcal{O}(z^{-3/2})$ (p. 227 of [46]) and the fact that $\nabla_x a_{\eta 0}$ is the Wigner transform of

$$\frac{1}{\eta^{\frac{(d-1)\alpha+3}{2}}}(\nabla\chi)\Big(\frac{x}{\eta^\alpha}\Big)J_0\Big(\frac{|k_0||x|}{\eta}\Big),$$

since $\overline{J_0}(|x|) = J_0(-|x|)$ so that the gradients of $\overline{J_0}(|x|)$ and $J_0(|x|)$ cancel in the computation.

Since the scintillation function J_η is itself oscillatory, the limit depends at which scale it is measured. We thus define localized test functions of the form:

$$\varphi_{\eta,s_1,s_2}(x,k) = \frac{1}{\eta^{d(s_1+s_2)}}\varphi\Big(\frac{x}{\eta^{s_1}},\frac{k-k_1}{\eta^{s_2}}\Big), \tag{5.47}$$

where $(s_1, s_2) \in \mathbb{R}^2$ and $k_1 \in \mathbb{R}^d$ and $\varphi \in \mathcal{S}(\mathbb{R}^{2d})$. We do not optimize the convergence rates as a function of s_1 and s_2 so as to obtain statistical stability for averaging domains as small as possible. We refer to [15] for such results, where it is shown for instance that for initial conditions with large support, that is for $\alpha = 0$, then we only need $s_1 < 1$ to obtain statistical stability, which amounts to averaging the energy density over a domain of typical size $\eta^{1-\delta}$, with $\delta > 0$.

Denote

$$\gamma_d := d(1-\alpha) - 2d(s_1 + s_2), \quad \kappa := 2(1-\alpha) - s_1 - s_1 \vee s_2 + (\alpha - \beta) \vee 0,$$
$$\chi_d := 1 - \beta + ((\alpha - \beta) \vee 0) \wedge ((d-1)(1-\alpha-\beta) + \alpha), \quad \chi_2 := 1 + \alpha - 2\beta.$$

We shall also write that $f_1(\eta) \lesssim f_2(\eta)$ if there exists $C > 0$ such that $f_1(\eta) \le C f_2(\eta)$ for all $\eta \in (0, 1]$. Our first result is the following:

Theorem 5.7. *Let $d \ge 2$ and assume that hypotheses **H** are satisfied. Then, the scintillation function J_η verifies the following estimate, uniformly on compact intervals:*

$$|\langle J_\eta(t), \varphi_{\eta,s_1,s_2} \otimes \varphi_{\eta,s_1,s_2}\rangle| \lesssim g_d(\eta),$$

$$g_d(\eta) = \eta^{\gamma_d + \kappa} \vee \eta^{\chi_d}, \; d \ge 3,$$
$$g_2(\eta) = \eta^{\gamma_2 + \kappa} \vee \left[\eta^{\chi_2}(1 + |\log \eta^{\alpha-\beta}|) \wedge 1\right].$$

Here, $\langle \cdot, \cdot \rangle$ denotes the $S' - S$ duality pairing and $a \vee b = \max(a, b)$.

Theorem 5.7 is a refined version of the result of [15]. It can be shown, see Theorem 5.8 below, that the rate of convergence of J_η is optimal when the test function φ is smooth ($s_1 = s_2 = 0$) and for initial conditions of the form (5.44). Since the proof of Theorem 5.7 does not depend on the particular form of the initial conditions, we expect the rate to be optimal for any initial conditions satisfying hypotheses **H**, although we do not have a rigorous proof of such a statement.

Our second result on the convergence of scintillation requires that we first define:

$$\begin{aligned}
j^1_\alpha(t,x,k,y,p) &= \delta(x - x_0 - tk)\,\delta(y - x_0 - tp)(\nabla\delta)^T(k - k_0)\,M^\alpha(t)\,(\nabla\delta)(p - k_0),\\
(M^{\frac12}(t))_{ij} &= \hat R(\mathbf{0}) \int_{\mathbb{R}^d} \mathcal{F}\partial_{x_i} a_0 \otimes \partial_{y_j} a_0(w, tw, -w, -tw)\,dw,\\
(M^\alpha(t))_{ij} &= M_{ij} = (M^{\frac12}(0))_{ij}, \qquad 0 \le \alpha < \frac12,\\
(M^\alpha(t))_{ij} &= \int_0^\infty (M^{\frac12}(t))_{ij}\,dt, \qquad \frac12 < \alpha < 1.
\end{aligned}$$

The above matrices are well-defined and for $0 \le \alpha < 1$, we have

$$|(M^\alpha)_{ij}| \;\le\; \hat R(\mathbf{0}) \|\mathcal{F}\partial_{y_j} a_0\|_{L^\infty(\mathbb{R}^{2d})} \left(\|\mathcal{F}_x \partial_{x_i} a_0\|_{L^1(\mathbb{R}^{2d})} + \|\mathcal{F}_k \partial_{x_i} a_0\|_{L^1(\mathbb{R}^{2d})}\right).$$

We also need to define

$$\begin{aligned}
j^2_\alpha(t,x,k,y,p) = {}& 2\,\delta(x-y)\Big(\sigma_\alpha(t,x,k-k_0)\delta(p-k) - \sigma_\alpha(t,x,p-p_0)\delta(k-k_0)\\
&-\sigma_\alpha(t,x,k-k_0)\delta(p-p_0) + \delta(k-k_0)\delta(p-p_0)\int_{\mathbb{R}^d} \sigma_\alpha(t,x,k)dk\Big),
\end{aligned}$$

where the cross section σ_α depends on the value of α and on the spatial dimension:

$$\begin{aligned}
\sigma_0(t,x,p) &= (2\pi)^d\,\hat R^2(p)\int_0^t d\tau e^{-2R_0(t-\tau)}|\mathcal{F}_k a_0(x-x_0-k_0t-(t-\tau)\frac{1}{2}p,-\tau p)|^2,\\
\sigma_\alpha(t,x,k) &= \delta(x-x_0-tk_0)\,\sigma_\alpha(t,k), \qquad \alpha>0,\\
\sigma_{\frac12}(t,k) &= \hat R^2(k)\int_0^\infty\int_{\mathbb{R}^d}|\mathcal{F}a_0(w,tw-\tau k)|^2dwd\tau,\\
\sigma_\alpha(t,k) &= \sigma(k)=\sigma_{\frac12}(0,k), \qquad 0<\alpha<\frac12,\\
\sigma_\alpha(t,k) &= \int_0^\infty \sigma_{\frac12}(t,k)dt, \qquad \frac12<\alpha<1, \quad d\geq 3,\\
\sigma_\alpha(t,k) &= \hat R^2(k)\int_0^\infty\int_{\mathbb{R}^d}|\mathcal{F}a_0(\tau k,w)|^2dwd\tau, \qquad \frac12<\alpha<1, \quad d=2.
\end{aligned}$$

Moreover, $\sigma_0\in\mathcal{C}^0([0,T],L^1(\mathbb{R}^d\times\mathbb{R}^d))$, $\sigma_\alpha(t,k)\in L^1(\mathbb{R}_+\times\mathbb{R}^d)$ for $0<\alpha\leq\frac12$ and $\sigma_\alpha(0,k)\in L^1(\mathbb{R}^d)$ for $\frac12<\alpha<1$.

We define the functional spaces X_p (for $1\leq p\leq\infty$), and Z the spaces of tempered distributions h in $\mathcal{S}'(\mathbb{R}^{4d})$ such that

$$\begin{aligned}
\|h\|_{X_p}^p &= \sup_{v,\zeta\in\mathbb{R}^d}\int_{\mathbb{R}^d}\sup_{\xi\in\mathbb{R}^d}|\mathcal{F}h(u,\xi,v,\zeta)|^pdu<\infty, \qquad 1\leq p<\infty\\
\|h\|_{X_\infty} &= \sup_{u,\zeta,v,\xi\in\mathbb{R}^d}|\mathcal{F}h(u,\xi,v,\zeta)|<\infty,\\
\|h\|_Z &= (2\pi)^{-4d}\int_{\mathbb{R}^{4d}}\omega(u,\xi,v,\zeta)|\mathcal{F}h(u,\xi,v,\zeta)|d\xi dudvd\zeta<\infty,\\
\omega(u,\xi,v,\zeta) &= (1+|\xi|+|\xi||u|+|u|^2)(1+|\zeta|+|\zeta||v|+|v|^2).
\end{aligned}$$

Here $|u|$ is the Euclidean norm of the vector u. We denote by Z' the dual of Z. Above, we identified the Fourier transform of the distribution h with the function $\mathcal{F}h$.

Then we have the following result for the convergence of the scintillation:

Theorem 5.8. *Assume the initial condition ψ_η^0 has the form (5.44). Then under the assumptions and notations of Theorem 5.7, we have, for $0<\alpha<1$,*

$$J_\eta=\eta^{(d+2)(1-\alpha)+(2\alpha-1)\vee 0}\,J_\alpha^1+\eta^{d(1-\alpha)+\alpha}\big(\big[\eta^{2\alpha-1}f_d(\eta)\big]\wedge 1\big)\,J_\alpha^2+r_\eta,$$

where $f_d=1$ when $d\geq3$ and $f_2=1+|\log\eta^{\alpha-\beta}|$, where r_η is negligible compared to the first two terms in the $L^\infty((0,T),\mathcal{S}'(\mathbb{R}^{4d}))-*$ topology, and where we have defined

$$J_\eta=\eta^d\,J_0^2+r_\eta \quad\text{when } \alpha=0, \qquad \text{and} \qquad J_\eta=\eta\,J_1^1+r_\eta \quad\text{when } \alpha=1.$$

Here, $J_\alpha^1\in\mathcal{C}^0([0,T],Z')$ when $\alpha<1$ and $J_1^1\in\mathcal{C}^0([0,T],X_\infty)$ and $J_\alpha^2\in\mathcal{C}^0([0,T],X_\infty)$ are distributional solutions to the following 4-transport equations,

$$\Big(\frac{\partial}{\partial t}+\mathcal{T}_2+2R_0-\mathcal{Q}_2\Big)J_\alpha^i=S_\alpha^i, \qquad J_\alpha^i(t=0,\cdot)=J_\alpha^{i,0}. \tag{5.48}$$

For $i=1,2$, we have $S_\alpha^i=0$ when $\alpha>\frac12$ and $J_\alpha^{i,0}=0$ when $\alpha\leq\frac12$, and

$$S_\alpha^i=j_\alpha^i \quad\text{when}\quad 0\leq\alpha\leq\frac12 \qquad\text{and}\qquad J_\alpha^{i,0}=j_\alpha^i(0,\cdot) \quad\text{when}\quad \frac12<\alpha<1.$$

Theorem 5.8 indicates how the statistical instabilities propagate. Depending on the value of α, either the first term or the second term dominates in the decomposition of J_η. When $d \geq 3$, the critical value of α is $\alpha^\star = \frac{2}{3}$: when $\alpha < \alpha^\star$, then the term involving J^2_α is the leading one, while the term involving J^1_α dominates when $\alpha > \alpha^\star$; when $\alpha = \alpha^\star$, both terms are of the same order. Both J^1_α and J^2_α satisfy a 4-transport equation. Depending on whether $\alpha \leq \frac{1}{2}$ or $\alpha > \frac{1}{2}$, the instabilities are created either by a source term or by an initial condition. J^1_α is the most singular term as the corresponding data in the transport equation are proportional to delta distributions both in space and momentum (when $\alpha < 1$) whereas the data corresponding to J^2_α are more regular in the momentum variables. This should be related to the fact that J^1_α is linear with respect to the power spectrum $\hat{R}$ while J^2_α is proportional to $\hat{R}^2$ so that J^1_α corresponds to the simple scattering contribution to the scintillation while J^2_α corresponds to the double scattering and is therefore more regular. Moreover, when $\alpha < \alpha^\star$, the double scattering contribution gives the leading order, while it is given by the simple scattering when $\alpha > \alpha^\star$. It can also be noticed that higher order scattering terms are negligible in the limit. Let us now examine the different scenarios depending on the value of α.

Case $0 < \alpha \leq \frac{1}{2}$. The initial condition $a_{\eta 0}$ is more singular in the momentum variables than in the spatial variables, with comparable singularities when $\alpha = \frac{1}{2}$. The instabilities are created by the ballistic part of the wave through the source term j^2_α supported at the spatial points $x = y = x_0 - tk_0$ with four configurations for the momentum k and p: (i) $k = p$, the amplitude of k is given by $\sigma_{\frac{1}{2}}(0, k - k_0)$ when $\alpha < \frac{1}{2}$ and by $\sigma_{\frac{1}{2}}(t, p - p_0)$ when $\alpha = \frac{1}{2}$; (ii) $k = k_0$, the amplitude of p is given by $\sigma_{\frac{1}{2}}(0, p - p_0)$; (iii) $p = p_0$, the amplitude of k is given by $\sigma_{\frac{1}{2}}(0, k - k_0)$; (iv) $k = p = k_0$. Instabilities are thus created along the wave propagation in the direction of the initial condition k_0 but also in other directions.

Case $\frac{1}{2} < \alpha < 1$. The initial condition $a_{\eta 0}$ is more singular in the spatial variables than in the momentum variables. This results in a stronger localization of the instabilities, which undergo more scattering and decrease exponentially with time. They are generated by an initial condition given by $j^1_\alpha(0, \cdot)$ when $\alpha > \alpha^\star$ and $j^2_\alpha(0, \cdot)$ when $\alpha < \alpha^\star$. When $\alpha < \alpha^\star$, instabilities are created at $x = y = x_0$ with the same momentum configuration as the case $0 < \alpha \leq \frac{1}{2}$. When $\alpha > \alpha^\star$, instabilities are still created at $x = y = x_0$ but with momentum $k = p = k_0$. Note that these instabilities are fairly singular since they are defined in this case by gradients of delta distributions.

Case $\alpha = 1$. This the *most unstable* case since instabilities are of order η. Since in this configuration the initial condition $a_{\eta 0}$ is regular with respect to k, instabilities are created at $x = y = x_0$ in all directions, which can be seen from the following expression of $J^{1,0}_1$, which is more regular in the momentum variables than $J^{1,0}_\alpha$ for $\alpha < 1$:

$$
\begin{aligned}
&J^{1,0}_1(x,k,y,p) \\
&= \Big(\pi \int_{\mathbb{R}^d} dw \hat{R}(w) \delta(w \cdot (k-p)) G(w, k - k_0, p - k_0) \\
&\quad + i \,\text{p.v.} \int_{\mathbb{R}^d} dw \hat{R}(w) \frac{1}{w \cdot (k-p)} G(w, k - k_0, p - k_0) \Big) \delta(x-y)\delta(x-x_0)
\end{aligned}
$$

$$G(w,k,p) = \Big[\mathcal{F}_x a_0(-w,k+\frac{w}{2}) - \mathcal{F}_x a_0(-w,k-\frac{w}{2})\Big]\Big[\mathcal{F}_x a_0(w,p+\frac{w}{2}) - \mathcal{F}_x a_0(w,p-\frac{w}{2})\Big].$$

Besides, $J_1^{1,0}$ belongs to X_∞, is real-valued, and the principal value contribution vanishes when a_0 is even with respect to the variable x.

Case $\alpha = 0$. This is the most stable case since instabilities are of order η^d. The initial condition is regular with respect to the spatial variables so that the source term j_0^2 is also regular. The situation is essentially the same as the case $0 < \alpha \leq \frac{1}{2}$. The main difference is that the instabilities are created not only at the ballistic position at time t (that is at $x = x_0 - kt$), but on a larger domain related to the spatial support of a_0.

Finally, we remark that in the most stable configurations (when $\alpha < \frac{1}{2}$), the instabilities persist with time while they decrease for more unstable configurations (when $\alpha > \frac{1}{2}$).

When $d = 2$, the situation is similar: only the values of $\alpha^\star$ and σ_α change.

5.5. **Transport equations for time-dependent Schrödinger.** In the preceding section, the random potential was replaced by a time dependent potential with extremely rapid oscillations (modeled as the white noise limit of a very rapidly oscillatory function). In this section, we consider the more realistic model where the fluctuations in the time domain are comparable or slow compared to the fluctuations in the spatial domain. The temporal fluctuations (even without the white noise assumption considered in the previous section) significantly simplify the analysis of convergence compared to the time-independent case considered in [37]. Heuristically, since time does not loop back to previously visited positions in the state space, random mixing is much more efficient when the potential is allowed to vary in time. However, since the fluctuations are no longer modeled as white noise, the spatial fluctuations also become important and the analysis is more involved than in the Itô-Schrödinger setting.

Once again, we turn to the parabolic wave equation, and denote by t the direction of propagation, and by x the transverse directions (or one may simply think of this as a Schrödinger equation with a time-dependent random potential, regardless of its murky wave origins)

$$\begin{gathered} i\varepsilon \frac{\partial \psi_\varepsilon}{\partial t} + \frac{\varepsilon^2}{2}\Delta\psi_\varepsilon - \sqrt{\varepsilon} V\left(\frac{t}{\varepsilon}, \frac{x}{\varepsilon}\right)\psi_\varepsilon = 0 \\ \psi_\varepsilon(0,x) = \psi_\varepsilon^0(x;\zeta). \end{gathered} \tag{5.49}$$

Here, the initial data depend on an additional random variable ζ defined over a probability space (S, Σ, μ), so that we consider a mixture of states.

We follow the presentation in [10, 11, 20]. For a related derivation of the kinetic equation in the time-dependent setting, we refer the reader to [70].

5.5.1. *Equation for the Wigner transform.* We want to analyze the energy density of the solution to the paraxial wave equation in the limit $\varepsilon \to 0$. As in the preceding chapter, the Wigner transform is a useful tool. Let us define the Wigner transform as the usual Wigner transform of the field ψ_ε averaged over the parameter $\zeta \in S$:

$$W_\varepsilon(t,x,k) = \int_{\mathbb{R}^d \times S} e^{ik\cdot y}\psi_\varepsilon\left(t, x - \frac{\varepsilon y}{2};\zeta\right)\bar{\psi}_\varepsilon\left(t, x + \frac{\varepsilon y}{2};\zeta\right)\frac{dy}{(2\pi)^d}d\mu(\zeta). \tag{5.50}$$

We *assume* that the initial data $W_\varepsilon(0,x,k)$ converges strongly in $L^2(\mathbb{R}^d \times \mathbb{R}^d)$ to a limit $W_0(x,k)$. This is possible thanks to the introduction of a mixture of states, i.e., an integration against the measure $\mu(d\zeta)$. This is the main reason why the space (S, Σ, μ) is introduced, as we have previously discussed.

Using the calculus introduced earlier we verify that the Wigner transform satisfies the following evolution equation

$$\frac{\partial W_\varepsilon}{\partial t} + k \cdot \nabla_x W_\varepsilon = \frac{1}{i\sqrt{\varepsilon}} \int_{\mathbb{R}^d} e^{ip\cdot x/\varepsilon} \Big[W_\varepsilon(k - \frac{p}{2}) - W_\varepsilon(k + \frac{p}{2}) \Big] \tilde{V}\Big(\frac{t}{\varepsilon}, p\Big) \frac{dp}{(2\pi)^d}. \quad (5.51)$$

Here, $\tilde{V}(t,p)$ is the partial Fourier transform of $V(t,x)$ in the variable x. The above evolution equation *preserves* the $L^2(\mathbb{R}^d \times \mathbb{R}^d)$ norm of $W_\varepsilon(t,\cdot,\cdot)$:

Lemma 5.9. *Let $W_\varepsilon(t,x,k)$ be the solution of (5.51) with initial conditions $W_\varepsilon(0,x,k)$. Then we have*

$$\|W_\varepsilon(t,\cdot,\cdot)\|_{L^2(\mathbb{R}^d\times\mathbb{R}^d)} = \|W_\varepsilon(0,\cdot,\cdot)\|_{L^2(\mathbb{R}^d\times\mathbb{R}^d)}, \qquad \textit{for all } t > 0. \quad (5.52)$$

Proof. This can be obtained by integrations by parts in (5.51), in a way that is similar to showing that (5.49) preserves the L^2 norm. □

5.5.2. *Hypotheses on the randomness.* We describe here the construction of the random potential $V(t,x)$. Our main hypothesis is that $V(t,x)$ is a Markov process in the t variable. The Markovian hypothesis is crucial to simplify the mathematical analysis because it allows us to treat the process $t \mapsto (V(t/\varepsilon, x/\varepsilon), W_\varepsilon(t,x,k))$ as jointly Markov.

In addition to being Markovian, $V(t,x)$ is assumed to be stationary jointly in (t,x), mean zero, and is constructed as follows (this construction is quite standard). Denote by $\mathcal{V} = C_b^1(\mathbb{R}^d)$ the space of $\mathbb{R}-$valued functions that are bounded and continuous together with the first derivatives. It is equipped with the standard supremum norm that shall be denoted by $\|\cdot\|_{\mathcal{V}}$. Suppose that $\{V_t,\ t \ge 0\}$ is a Markovian process with the state space $\mathcal{V}$ and for $v \in \mathcal{V}$, define $V(v) := v(0)$. We assume that π is the invariant distribution under the dynamics of the process. Moreover it is also spatially homogeneous, i.e. $\pi\tau_x = \pi$, for all $x \in \mathbb{R}^d$. Here $\tau_x : \mathcal{V} \to \mathcal{V}$ is defined by $\tau_x v(\cdot) := v(x + \cdot)$. The random field is defined as $V(t,x) := V(\tau_x V_t)$. The correlation function $R(t,x)$ of the field is defined as

$$R(t,x) = \mathbb{E}\left\{V(s,y)V(t+s,x+y)\right\} \qquad \text{for all } (t,x),(s,y) \in \mathbb{R}^{1+d}. \quad (5.53)$$

Its Fourier transform $\hat{R}(\omega,p)$ as

$$\hat{R}(\omega,p) = \int_{\mathbb{R}^{1+d}} e^{-i\omega t - ip\cdot x} R(t,x)\,dt\,dx, \quad (5.54)$$

is called the space-time power spectrum of the field. We shall also use the partial Fourier transform in the spatial variable and denote it by

$$\tilde{R}(t,p) = \int_{\mathbb{R}^d} e^{-ip\cdot x} R(t,x)\,dx. \quad (5.55)$$

Moreover the $L^2(\pi)$ generator Q of the process is supposed to satisfy the spectral gap property, i.e. there exists $\alpha > 0$ such that

$$\langle f, (-Q)f\rangle_\pi \ge \alpha(\|f\|^2_{L^2(\pi)} - \langle f, 1\rangle_\pi^2) \quad f \in L^2(\pi).$$

With this assumption we can guarantee that the Fredholm alternative holds for the Poisson equation. Namely, for any $f \in L^2(\pi)$ such that $\langle f, 1\rangle_\pi = 0$, we have

$$\|e^{rQ} f\|_{L^2(\pi)} \le C\|f\|_{L^2(\pi)} e^{-(\alpha/2)r}. \quad (5.56)$$

Given the above hypotheses, the Fredholm alternative holds for the Poisson equation

$$Qg = f. \tag{5.57}$$

It has a unique solution g with $\langle g, 1\rangle_\pi = 0$ and $g \in L^2(\pi)$. The solution g is given explicitly by

$$f(v) = -\int_0^\infty \left(e^{rQ} g(v)\right) dr, \tag{5.58}$$

and the integral converges absolutely thanks to (5.56).

The unitary representation of the group of shifts $\{\tau_x, x \in \mathbb{R}^d\}$ on $L^2(\pi)$ is given by $T_x f := f \circ \tau_x$, $x \in \mathbb{R}^d$ for $f \in L^2(\pi)$. We assume that it is strongly continuous, so, by the spectral theorem, there exists its spectral resolution. It can be used to construct $\hat{V}(dp)$, the stochastic measure, defined on $(\mathcal{V}, \mathcal{B}(\mathcal{V}), \pi)$, that corresponds to $T_x V := V \circ \tau_x$. This measure is $L^2(\pi)$-valued and satisfies: $\hat{V}(-dp) = \hat{V}^*(dp)$ (because V is real valued) and

$$\hat{R}(p)\delta(p-q)dpdq = \langle \hat{V}(dp)\hat{V}^*(dq)\rangle_\pi, \tag{5.59}$$

where $\langle\cdot\rangle_\pi$ we denote the expectation with respect to π,

$$R(0,x) = \int e^{ip\cdot x}\hat{R}(p)dp.$$

We shall also use the notation $\tilde{V}(t; dp) := \hat{V}(dp; V_t)$, $t \geq 0$. We can write then

$$V(t,x) = \int e^{ip\cdot x}\tilde{V}(t;dp),$$

where the right hand side is interpreted as a composition of the appropriate stochastic integral and the process V_t. The equality is understood in the L^2 sense. In fact, the stochastic integral $\int \psi(p)\tilde{V}(t;dp)$ can be defined for any complex valued function that satisfies

$$\|\psi\|^2_{\mathcal{H}} := \int |\psi(p)|^2 \hat{R}(p)dp < +\infty. \tag{5.60}$$

Denote by $\mathcal{H}$ the space consisting of complex even function $\psi(p)$ (i.e. $\psi(-p) = \psi^*(p)$) such that the above condition is satisfied. It is Hilbert when equipped with the norm $\|\cdot\|_{\mathcal{H}}$.

We assume also that

$$V_* := \text{ess-sup}\|v\|_{\mathcal{V}} < +\infty, \tag{5.61}$$

where the essential supremum is taken in π measure.

5.5.3. *Main result of convergence to kinetic model.* Let us summarize the hypotheses. We define $W_\varepsilon(t,x,k)$ in (5.50) as a mixture of states of solutions to the paraxial wave equation (5.49). The mixture of state is such that $W_\varepsilon(0,x,k)$, whence $W_\varepsilon(t,x,k)$ for all $t > 0$ is uniformly bounded in $L^2(\mathbb{R}^{2d})$. We assume that the initial conditions $W_\varepsilon(0,x,k)$ converge *strongly* in $L^2(\mathbb{R}^{2d})$ to its limit $W_0(0,x,k)$. We further assume that the random field $V(t,x)$ satisfies the hypotheses described above. By $L^2_w(\mathbb{R}^{2d})$ we denote the space $L^2(\mathbb{R}^{2d})$ equipped with the weak topology.

Then we have the following convergence result.

Theorem 5.10. *Under the above assumptions, the processes described by the Wigner distribution $\{W_\varepsilon(t),\, t \geq 0\}$ converge, as $\varepsilon \to 0$, in probability in the topology of*

$C([0,+\infty);L^2_w(\mathbb{R}^{2d}))$ to the deterministic process corresponding to the solution $\overline{W}$ of the following transport equation

$$\frac{\partial \overline{W}}{\partial t} + k\cdot\nabla_x \overline{W} = \mathcal{L}\overline{W}, \tag{5.62}$$

where the scattering kernel has the form

$$\mathcal{L}W(x,k) = \int_{\mathbb{R}^d} \hat{R}\Big(\frac{|p|^2-|k|^2}{2}, p-k\Big)\Big(W(x,p)-W(x,k)\Big)\frac{dp}{(2\pi)^d}. \tag{5.63}$$

The above statement means that for any test function $\lambda \in L^2(\mathbb{R}^{2d})$ the process $\{\langle W_\varepsilon(t),\lambda\rangle,\ t\ge 0\}$ converges, as $\varepsilon\to 0$, to $\{\langle \overline{W}(t),\lambda\rangle,\ t\ge 0\}$ in probability as $\varepsilon\to 0$, uniformly on all finite time intervals.

Note that the whole process $\{W_\varepsilon(t),\ t\ge 0\}$, and not only its average $\mathbb{E}\{W_\varepsilon(t)\}$ converges to the (deterministic) limit $\overline{W}(t)$. This means that the process is *statistically stable* in the limit $\varepsilon\to 0$. The process $W_\varepsilon(t,x,k)$ does not converge pointwise to the deterministic limit: averaging against a test function $\lambda(x,k)$ is necessary.

We now summarize the main ingredients of the proof before a detailed proof is presented in Section 5.5.4. Recall that the main assumption is that $V(t,x)$ is Markov in the t variable (but this assumption may be greatly relaxed – see [39]). Let us set $T>0$ and consider $t\in[0,T]$. The Markov assumption on $\{V_t,\ t\ge 0\}$ allows us to show that $(V(t/\varepsilon,x/\varepsilon),W_\varepsilon(t,x,k))$ is jointly Markov with the phase space $\mathcal{V}\times B_W$, where $B_W=\{\|W\|_2\le C\}$ is an appropriate ball in $L^2(\mathbb{R}^{2d})$.

Evolution equation and random process.

Recall that W_ε satisfies the Cauchy problem

$$\frac{\partial W_\varepsilon}{\partial t} + k\cdot\nabla_x W_\varepsilon = \mathcal{L}_\varepsilon W_\varepsilon, \tag{5.64}$$

with $W_\varepsilon(0,x,k)=W^0_\varepsilon(x,k)$, where formally

$$\mathcal{L}_\varepsilon W_\varepsilon = \frac{1}{i(2\pi)^d\sqrt{\varepsilon}}\int_{\mathbb{R}^d}\tilde{V}(\frac{t}{\varepsilon},dp)e^{ip\cdot x/\varepsilon}\Big[W_\varepsilon(x,k-\frac{p}{2})-W_\varepsilon(x,k+\frac{p}{2})\Big]. \tag{5.65}$$

The solution to the above Cauchy problem is understood as a process, adapted to the filtration $\{\mathcal{F}_{t/\varepsilon},\ t\ge 0\}$ whose trajectories belong to $C([0,+\infty);L^2(\mathbb{R}^{2d})\cap L^\infty(\mathbb{R}^{2d}))$ and such that for every test function $\lambda(t,x,k)$ belonging to $C^1([0,+\infty);L^2(\mathbb{R}^d)\otimes\mathcal{H})$, it satisfies

$$\langle W_\varepsilon(t),\lambda(t)\rangle - \langle W^0_\varepsilon,\lambda(0)\rangle = \int_0^z \langle W_\varepsilon(s),\Big(\frac{\partial}{\partial s}+k\cdot\nabla_x+\mathcal{L}_\varepsilon\Big)\lambda(s)\rangle ds.$$

Here, we have used that $\mathcal{L}_\varepsilon$ is a self-adjoint operator for $\langle\cdot,\cdot\rangle$.

Tightness of the family of $\varepsilon-$measures. The above construction defines the process $W_\varepsilon(t)$ in $L^2(\mathbb{R}^{2d})$ and generates the corresponding measure P_ε on the space $C([0,+\infty);L^2(\mathbb{R}^{2d}))$ of functions continuous in time and with values in $L^2(\mathbb{R}^d)$. We denote by $\{\mathcal{M}_t,\ ,t\ge 0\}$ the filtration of σ-algebras generated by coordinate maps $t\mapsto W(t)$. The σ algebra $\mathcal{M}$ generated by the filtration coincides with the Borel σ-algebra on the space. We recall that intuitively, the filtration renders the past, i.e. events described at times $s\le t$ measurable, i.e., "known", and the future $t>s$ non-measurable, i.e., not known yet.

The family P_ε will be shown to be *tight as* $\varepsilon\to 0$, i.e. for any sequence $\varepsilon_n\to 0$ one can choose a weakly convergent subsequence of measures from $\{P_{\varepsilon_n},\ n\ge 1\}$.

More precisely, we can extract a subsequence still denoted by P_{ε_n}, such that for any continuous and bounded function f defined on $\mathcal{C}_T := C([0,T]; L^2(\mathbb{R}^{2d}))$, we have

$$\mathbb{E}^{P_{\varepsilon_n}}\{f\} \equiv \int_{\mathcal{C}_T} f(\omega)dP_{\varepsilon_n}(\omega) \to \int_{\mathcal{C}_T} f(\omega)dP(\omega) \equiv \mathbb{E}^P\{f\}, \qquad \text{as} \qquad n \to +\infty. \tag{5.66}$$

The limiting measure P does not depend on the subsequence – it is the δ-type measure representing the law of the deterministic process $\{\overline{W}(t),\, t \geq 0\}$. This in turn implies that P_ε converges weakly to P, as $\varepsilon \to 0$.

Construction of the first approximate martingale. Once tightness is ensured, the proof of convergence of the processes $W_\varepsilon(t)$ to its deterministic limit is obtained in two steps. Let us fix a deterministic test function $\lambda \in L^2(\mathbb{R}^{2d})$. We use the Markovian property of the random field $V(t,x)$ to construct the first functional $G_\lambda : \mathcal{C}_L \to C[0,T]$ by

$$G_\lambda[W](t) = \langle W(t), \lambda(t)\rangle - \int_0^t \langle W(s), \frac{\partial \lambda}{\partial s} + k\cdot\nabla_x\lambda + \mathcal{L}\lambda(s)\rangle ds. \tag{5.67}$$

Here, $\mathcal{L}$ is the limiting scattering kernel defined in (5.63). We will show that $\{G_\lambda(t),\, t \geq 0\}$ is an approximate P_ε-martingale (with respect to the filtration $\mathcal{M}_s$). More precisely the above means that there exists a constant $C_{\lambda,T} > 0$ such that

$$\left|\mathbb{E}^{P_\varepsilon}\{G_\lambda[W](t)|\mathcal{M}_s\} - G_\lambda[W](s)\right| \leq C_{\lambda,T}\sqrt{\varepsilon} \tag{5.68}$$

uniformly for all $W \in \mathcal{C}_L$ and $0 \leq s < t \leq T$. Choosing $s = 0$ above, the two convergences (5.66) and (5.68) (weak against strong) show that

$$\mathbb{E}^P\{G_\lambda[W](t)\} = 0. \tag{5.69}$$

We thus obtain the transport equation (5.62) for $\overline{W}(t) := \mathbb{E}^P\{W(t)\}$ in its weak formulation.

Construction of the second approximate martingale and the convergence of the full family of $\varepsilon-$measures. So far, we have characterized the convergence of the first moment of P_ε. We now consider the convergence of the second moment and show that the variance of the limiting process vanishes, whence the convergence to a deterministic process.

We will show that for every test function $\lambda(t,x,k)$, the new functional

$$G_{2,\lambda}[W](t) = \langle W,\lambda\rangle^2(t) - 2\int_0^t \langle W,\lambda\rangle(\zeta)\langle W, \frac{\partial\lambda}{\partial t} + k\cdot\nabla_x\lambda + \mathcal{L}\lambda\rangle(\zeta)d\zeta \tag{5.70}$$

is also an approximate P_ε-martingale. We then obtain that

$$\mathbb{E}^{P_\varepsilon}\left\{\langle W,\lambda\rangle^2\right\} \to \langle\overline{W},\lambda\rangle^2. \tag{5.71}$$

This crucial convergence implies convergence in probability. It follows that the limit measure P is unique and deterministic, and that the whole sequence P_ε converges.

5.5.4. *Proof of Theorem 5.10.* The proof of tightness of the family of measures P_ε is postponed to the end of the section as it requires estimates that are developed in the proofs of convergence of the approximate martingales. We thus start with the latter proofs.

Convergence in expectation. To obtain the approximate martingale property (5.68), one has to consider the conditional expectation of functionals $F \in C([0,+\infty); C_b(B_W \times \mathcal{V}))$, with respect to the (joint) probability measure $\tilde{P}_\varepsilon$ that is the law of $\{(W_\varepsilon(t), V_{t/\varepsilon}),\, t\ge 0\}$ on $C([0,T];B_W\times\mathcal{V})$. In fact the only functionals we need to consider are those of the form $F(t,W,v)=\langle W,\lambda_v(t)\rangle$ with $\lambda\in L^\infty(\mathcal{V};C^1([0,T];\mathcal{S}(\mathbb{R}^{2d})))$. Given F us above let us define the conditional expectation

$$\mathbb{E}^{\tilde P_\varepsilon}_{W,v,t}\left\{F(\tau,W(\tau),V(\tau))\right\}:=\mathbb{E}^{\tilde P_\varepsilon}\left\{F(\tau,W(\tau),V(\tau))|\, W(t)=W,V(t)=v\right\},\ \tau\ge t.$$

Using Markov property of the process $\{V_t,\, t\ge 0\}$ and the fact that $W_\varepsilon(t)$ solves (5.64) we obtain

$$\begin{aligned}&\tfrac{d}{dh}\mathbb{E}^{\tilde P_\varepsilon}_{W,v,t}\left\{F(t+h,W(t+h),V(t+h))\right\}\Big|_{h=0}\\&=\tfrac{1}{\varepsilon}\langle W,Q\lambda_v(t)\rangle+\left\langle W,\left(\frac{\partial}{\partial t}+k\cdot\nabla_x+\tfrac{1}{\sqrt{\varepsilon}}\mathcal{K}[v,\tfrac{x}{\varepsilon}]\right)\lambda_v(t)\right\rangle.\end{aligned}$$

Here, for any $(v,z,\lambda)\in\mathcal{V}\times\mathbb{R}^d\times\mathcal{S}(\mathbb{R}^{2d})$ we let

$$\begin{aligned}\mathcal{K}[v,z]\lambda(x,k)&:=\sum_{\sigma=\pm1}\sigma\int_{\mathbb{R}^d}e^{ik\cdot z}v(\eta+\frac{\sigma z}{2})\hat\lambda(x,z)dz\\&=\tfrac{i}{(2\pi)^d}\sum_{\sigma=\pm1}\sigma\int_{\mathbb{R}^d}e^{ip\cdot z}\lambda(x,k+\frac{\sigma p}{2})\widehat V(dp;v),\end{aligned}$$

where $\hat\lambda(x,z)$ denotes the partial Fourier transform of λ with respect to the second variable. Equality (5.72) implies that

$$G^\varepsilon_\lambda(t)=\langle W(t),\lambda_{V(t)}(t)\rangle-\int_0^t\left\langle W(s),\mathcal{L}^\varepsilon_{s,V(s)}\lambda_{V(s)}(s)\right\rangle ds,\tag{5.72}$$

where $\mathcal{L}^\varepsilon_{s,v}:=\varepsilon^{-1}Q+\partial_s+k\cdot\nabla_x+\varepsilon^{-1/2}\mathcal{K}[v,x/\varepsilon]$, is a $\tilde P_\varepsilon$-martingale since the drift term has been subtracted.

Given a test function $\lambda(t,x,k)\in C^1([0,L];\mathcal{S})$ we construct a function

$$\lambda_\varepsilon(t,x,k)=\lambda(t,x,k)+\sqrt{\varepsilon}\lambda^\varepsilon_{1,V(t)}(t,x,k)+\varepsilon\lambda^\varepsilon_{2,V(t)}(t,x,k),\tag{5.73}$$

with $\lambda^\varepsilon_{i,v}(t)$, $i=1,2$ bounded in $C([0,L);\mathcal{V}\otimes L^2(\mathbb{R}^{2d}))$. This is the method of **perturbed test function**. Rather than performing asymptotic expansions on the Wigner transform itself, which is not sufficiently smooth to justify Taylor expansions, we perform the expansion on smooth test functions.

The functions λ_i^ε, $i=1,2$ will be chosen to remove all high-order terms in the definition of the martingale (5.72), i.e., so that

$$\|G^\varepsilon_{\lambda_\varepsilon}(t)-G_\lambda(t)\|_{L^2(\pi)}\le C_\lambda\sqrt{\varepsilon}\tag{5.74}$$

for all $t\in[0,L]$. Here $G^\varepsilon_{\lambda_\varepsilon}$ is defined by (5.72) with λ replaced by λ_ε, and G_λ is defined by (5.67). The approximate martingale property (5.68) follows from this.

The functions λ_i^ε, $i=1,2$ are as follows. Let $\lambda_{1,v}(t,x,\eta,k)$ be the π-mean zero solution of the Poisson equation

$$k\cdot\nabla_\eta\lambda_{1,v}(t)+Q\lambda_{1,v}(t)=-\mathcal{K}[v,\eta]\lambda(t).\tag{5.75}$$

It is given explicitly by

$$\lambda_{1,v}(t,x,\eta,k)=\int_0^\infty e^{rQ}\left\{\mathcal{K}[\cdot,\eta+kr]\lambda(t,x,k)\right\}dr.\tag{5.76}$$

The improper integral on the right hand side exists because

$$\left\langle\{\mathcal{K}[\cdot,\eta+kr]\lambda(t,x,k)\}^2\right\rangle_\pi=\left\langle\{\mathcal{K}[\cdot,\eta]\lambda(t,x,k)\}^2\right\rangle_\pi<+\infty,\quad\forall\, r\ge0.\tag{5.77}$$

Then we let $\lambda^\varepsilon_{1,v}(t,x,k) := \lambda_{1,v}(t,x,x/\varepsilon,k)$. A direct calculation, invoking (5.59), yields

$$\langle \mathcal{K}[v,\eta]\lambda_{1,v}\rangle_\pi = \mathcal{L}\lambda. \tag{5.78}$$

To define the second order corrector we introduce

$$\lambda^{(\mathcal{L})}_v(t,x,\eta,k) := \mathcal{L}\lambda_v(t,x,k) - (\mathcal{K}[v,\eta]\lambda_{1,v}(t))(x,k).$$

Using the first formula in (5.72) together with (5.61) we conclude that

$$\left\langle \left[\lambda^{(\mathcal{L})}_v(t,x,\eta,k)\right]^2\right\rangle_\pi < +\infty.$$

Therefore we can define $\lambda^\varepsilon_{2,v}(t,x,k) := \lambda_{2,v}(t,x,x/\varepsilon,k)$ where $\lambda_{2,v}(t,x,\eta,k)$ is the mean-zero solution of

$$k\cdot\nabla_\eta\lambda_{2,v} + Q\lambda_{2,v} = \lambda^{(\mathcal{L})}_v(t,x,\eta,k), \tag{5.79}$$

which exists thanks to (5.78).

The explicit expression for $\lambda_{2,v}$ is given by

$$\lambda_{2,v}(t,x,\eta,k) = -\int_0^\infty e^{rQ}[\lambda^{(\mathcal{L})}(t,x,\eta+kr,k)]dr.$$

Using (5.75) and (5.79) we have

$$\begin{aligned}
&\tfrac{d}{dh}\mathbb{E}^{\tilde P_\varepsilon}_{W,v,t}\left\{\langle W(t+h),\lambda_\varepsilon(t+h)\rangle\right\}\Big|_{h=0}\\
&= \left\langle W, \mathcal{L}^\varepsilon_{t,v}\left(\lambda(t)+\sqrt{\varepsilon}\lambda^\varepsilon_{1,v}(t)+\varepsilon\lambda^\varepsilon_{2,v}(t)\right)\right\rangle\\
&= \left\langle W, \left(\frac{\partial}{\partial t}+k\cdot\nabla_x+\mathcal{L}\right)\lambda(t)\right\rangle\\
&\quad + \left\langle W, \left(\frac{\partial}{\partial t}+k\cdot\nabla_x\right)\left(\sqrt{\varepsilon}\lambda^\varepsilon_{1,v}(t)+\varepsilon\lambda^\varepsilon_{2,v}(t)\right)+\sqrt{\varepsilon}\mathcal{K}[v,\tfrac{x}{\varepsilon}]\lambda^\varepsilon_{2,v}(t)\right\rangle\\
&= \left\langle W, \left(\frac{\partial}{\partial t}+k\cdot\nabla_x+\mathcal{L}\right)\lambda(t)\right\rangle + \sqrt{\varepsilon}\langle W,\zeta^\lambda_{\varepsilon,v}(t)\rangle,
\end{aligned}$$

with

$$\zeta^\lambda_{\varepsilon,v}(t) := (\partial_t+k\cdot\nabla_x)\,\lambda^\varepsilon_{1,v}(t)+\sqrt{\varepsilon}\,(\partial_t+k\cdot\nabla_x)\,\lambda^\varepsilon_{2,v}(t)+\mathcal{K}[v,\frac{x}{\varepsilon}]\lambda^\varepsilon_{2,v}(t).$$

The terms $k\cdot\nabla_x\lambda^\varepsilon_{i,v}(t)$ above are understood as the differentiation with respect to the slow variable x only, and not with respect to $\eta = x/\varepsilon$. Let $\zeta^\lambda_\varepsilon(t) := \zeta^\lambda_{\varepsilon,V(t)}(t)$. It follows, see (5.72), that

$$G^\varepsilon_{\lambda_\varepsilon}(t) = \langle W(t),\lambda_\varepsilon(t)\rangle - \int_0^t\left\langle W(s),\left(\frac{\partial}{\partial s}+k\cdot\nabla_x+\mathcal{L}\right)\lambda(s)\right\rangle ds - \sqrt{\varepsilon}\int_0^t\langle W(s),\zeta^\lambda_\varepsilon(s)\rangle ds \tag{5.80}$$

and is a martingale with respect to the Borel measure $\tilde P_\varepsilon$ on $C([0,T];B_W\times\mathcal{V})$. The estimate (5.68) follows from the following two lemmas.

Lemma 5.11. *Let $\lambda\in C^1([0,T];\mathcal{S}(\mathbb{R}^{2d}))$. Then the correctors $\lambda^\varepsilon_{i,v}(t)$, $i=1,2$ satisfy the following bounds*

$$C_i := \sup_{t\in[0,T]}\sup_{\varepsilon\in(0,1]}\left\{\langle\|\lambda^\varepsilon_{i,v}(t)\|^2_{L^2(\mathbb{R}^{2d})}\rangle_\pi + \left\langle\left\|(\partial_t+k\cdot\nabla_x)\lambda^\varepsilon_{i,v}(t)\right\|^2_{L^2(\mathbb{R}^{2d})}\right\rangle_\pi\right\} < +\infty. \tag{5.81}$$

We stress here that the derivative $k\cdot\nabla_x$ concerns only the slow variable.

Let $K_{\varepsilon,v}\lambda(x,k) := \mathcal{K}[v, x/\varepsilon]\lambda(x,k)$ for any $\lambda \in L^2(\mathbb{R}^{2d})$.

Lemma 5.12. *We have*

$$K_* := \operatorname*{ess\text{-}sup}_{v\in\mathcal{V}} \sup_{\varepsilon\in(0,1]} \|K_{\varepsilon,v}\|_{L^2(\mathbb{R}^{2d})\to L^2(\mathbb{R}^{2d})} < +\infty.$$

Lemma 5.12 follows immediately from the definition of $\mathcal{K}$ (see the first equality in (5.72)) and bound (5.61).

With these lemmas we conclude easily that there exists a constant $C > 0$ such that for all $\varepsilon \in (0,1]$ we have

$$\left|\mathbb{E}^{\tilde{P}_\varepsilon}\langle W(t), \lambda(t)\rangle - \mathbb{E}^{\tilde{P}_\varepsilon}\langle W(t), \lambda_\varepsilon(t)\rangle\right| \le C\sqrt{\varepsilon}$$

and

$$\mathbb{E}^{\tilde{P}_\varepsilon}\|\zeta_\varepsilon^\lambda(t)\|^2_{L^2(\mathbb{R}^{2d})} \le C, \quad \forall\, t \in [0, L] \tag{5.82}$$

so that (5.68) follows.

Proof of Lemma 5.11. Suppose $\lambda \in L^2(\mathbb{R}^{2d})$. Using (5.56) and (5.77) we obtain that

$$\begin{aligned}
\|\lambda^\varepsilon_{1,v}(t,x,k)\|_{L^2(\pi)} &\le \int_0^\infty \left\|e^{rQ}\mathcal{K}[\cdot, \frac{x}{\varepsilon} + kr]\lambda(t,x,k)\right\|_{L^2(\pi)} dr \\
&\le C\int_0^\infty e^{-(\alpha/2)r}\left\|\mathcal{K}[\cdot, \frac{x}{\varepsilon} + kr]\lambda(t,x,k)\right\|_{L^2(\pi)} dr \\
&= \frac{2C}{\alpha}\left\|\mathcal{K}[\cdot, \frac{x}{\varepsilon}]\lambda(t,x,k)\right\|_{L^2(\pi)}.
\end{aligned}$$

Hence, from Lemma 5.12 we get

$$\langle\|\lambda_1^\varepsilon(t)\|^2_{L^2(\mathbb{R}^{2d})}\rangle_\pi \le C\,\|\lambda(t)\|^2_{L^2(\mathbb{R}^{2d})}$$

for some constant independent of $\varepsilon \in (0,1]$ The bounds on λ_2^ε, as well as on the partials, are very similar in spirit and we omit them.

Lemma 5.11 and Lemma 5.12 together with (5.80) imply the bound (5.74). The tightness of measures P_ε, as $\varepsilon \to 0$, claimed in Lemma 5.14 below implies that the expectation $\mathbb{E}\{W_\varepsilon(t,x,k)\}$ converges weakly in $L^2(\mathbb{R}^{2d})$ to the solution $\overline{W}(t,x,k)$ of the transport equation for each $t \in [0,T]$.

Convergence in probability. We now prove that for any test function λ the second moment $\mathbb{E}\left\{\langle W_\varepsilon(t), \lambda\rangle^2\right\}$ converges to $\langle\overline{W}(t), \lambda\rangle^2$. This will imply the convergence in probability claimed in Theorem 5.10. The proof is similar to that for $\mathbb{E}\{\langle W_\varepsilon(t), \lambda\rangle\}$ and is based on constructing an appropriate approximate martingale for the functional $\langle W(t)\otimes W(t), \mu(t)\rangle$, where $\mu(t,x_1,k_1,x_2,k_2)$ is a test function, and $W\otimes W(t,x_1,k_1,x_2,k_2) = W(t,x_1,k_1)W(t,x_2,k_2)$. We need to consider the action of the infinitesimal generator on functions of $W \in L^2(\mathbb{R}^{2d})$ and $v \in \mathcal{V}$ of the form

$$F(W,v) = \langle W\otimes W, \mu_v(t)\rangle,$$

where $\mu_v(\cdot) \in C^1([0,+\infty); \mathcal{S}(\mathbb{R}^{2d}))$ is a given function. The infinitesimal generator acts on such functions as

$$\begin{aligned}
&\tfrac{d}{dh}\mathbb{E}^{\tilde{P}_\varepsilon}_{W,v,t}\left\{\langle W\otimes W(t+h), \mu_v(t+h)\rangle\right\}\Big|_{h=0} \\
&= \frac{1}{\varepsilon}\langle W\otimes W, Q\mu_v(t)\rangle + \langle W\otimes W, \mathcal{H}_2^\varepsilon\mu_v(t)\rangle,
\end{aligned} \tag{5.83}$$

where for any function $\mu(x_1,k_1,x_2,k_2)$ from $\mathcal{S}(\mathbb{R}^{2d})$ we let

$$\mathcal{H}_2^\varepsilon\mu = \sum_{j=1}^2 \frac{1}{\sqrt{\varepsilon}}\mathcal{K}_j\left[v,\frac{x_j}{\varepsilon}\right]\mu + k_j\cdot\nabla_{x_j}\mu, \tag{5.84}$$

with $\mathcal{K}_j$ acting on variable k_j by formula (5.72). Therefore the functional

$$\begin{aligned} &G_\mu^{2,\varepsilon}(t)\\ &= \langle W\otimes W(t),\mu_{V(t)}(t)\rangle - \int_0^t \left\langle W\otimes W(s),\mathcal{L}^{2,\varepsilon}_{s,V(s)}\mu_{V(s)}(s)\right\rangle ds, \end{aligned} \tag{5.85}$$

with

$$\mathcal{L}^{2,\varepsilon}_{s,v} := \frac{1}{\varepsilon}Q + \frac{\partial}{\partial s} + k_1\cdot\nabla_{x_1} + k_2\cdot\nabla_{x_2} + \frac{1}{\sqrt{\varepsilon}}(\mathcal{K}_1[v,\frac{x_1}{\varepsilon}] + \mathcal{K}_2[v,\frac{x_2}{\varepsilon}]),$$

is a $\tilde{P}^\varepsilon$ martingale. We let $\mu\in\mathcal{S}(\mathbb{R}^{2d}\times\mathbb{R}^{2d})$ be a test function independent of v. Denote $x=(x_1,x_2)$, and $k=(k_1,k_2)$. We will also use the following notation $\mu_j^\varepsilon(t,x,k)=\mu_{j,V(t)}(t,x,x/\varepsilon,k)$, $j=1,2$, with $\mu_{j,v}$ are to be determined later,

$$Q_K := Q + \sum_{j=1}^2 k_j\cdot\nabla_{\eta_j},\quad \mathcal{K}^{1,2} := \sum_{j=1}^2 \mathcal{K}_j\left[v,\eta_j\right]$$

and $\nabla_{t,1,2} := \partial_t + \sum_{j=1}^2 k_j\cdot\nabla_{x_j}$.

We define an approximation

$$\mu_\varepsilon(t,x,k) = \mu(t,x,k) + \sqrt{\varepsilon}\mu_1^\varepsilon(t,x,k) + \varepsilon\mu_2^\varepsilon(t,x,k).$$

We now use (5.83) to get

$$\begin{aligned} D_\varepsilon(t) &:= \frac{d}{dh}\Big|_{h=0}\mathbb{E}_{W,v,t}\langle W\otimes W(t+h),\mu_\varepsilon(t+h)\rangle\\ &= \frac{1}{\varepsilon}\left\langle W\otimes W, Q_K\mu(t)\right\rangle + \frac{1}{\sqrt{\varepsilon}}\left\langle W\otimes W, Q_K\mu_{1,v}(t)+\mathcal{K}^{1,2}\mu(t)\right\rangle\\ [2mm] &+ \left\langle W\otimes W, Q_K\mu_{2,v}(t)+\mathcal{K}^{1,2}\mu_{1,v}(t)+\nabla_{t,1,2}\mu(t)\right\rangle\\ [2mm] &+\sqrt{\varepsilon}\left\langle W\otimes W,\mathcal{K}^{1,2}\mu_{2,v}(t)+\nabla_{t,1,2}(\mu_{1,v}(t)+\sqrt{\varepsilon}\mu_{2,v}(t))\right\rangle. \end{aligned} \tag{5.86}$$

The above expression is evaluated at $\eta_j = x_j/\varepsilon$. The term of order ε^{-1} in $D_\varepsilon(t)$ vanishes since μ is independent of v and the fast variable η. We cancel the term of order $\varepsilon^{-1/2}$ in the same way as before by defining $\mu_{1,v}$ as the unique π-mean-zero (in the variables v and $\eta=(\eta_1,\eta_2)$) solution of

$$Q_K\mu_{1,v}(t,x,\eta,k) + \mathcal{K}^{1,2}\mu(t,x,\eta,k) = 0. \tag{5.87}$$

It is given explicitly by

$$\mu_{1,v}(t,x,\eta,k) = \int_0^\infty e^{rQ}\left[\mathcal{K}^{1,2}\mu(t,x,\eta_1+k_1r,\eta_2+k_2r,k)\right]dr.$$

When μ has the form $\mu(t)=\lambda(t)\otimes\lambda(t)$, then $\mu_{1,v}(t)$ has the form $\mu_{1,v}=\lambda_{1,v}(t)\otimes\lambda(t)+\lambda(t)\otimes\lambda_{1,v}(t)$ with the corrector $\lambda_{1,v}(t)$ given by (5.76). Let us also define $\mu_{2,v}(t)$ as the mean zero, with respect to $\pi(dv)$, solution of

$$Q_K\mu_2(t,x,\eta,k) + \mathcal{K}^{1,2}\mu_1(t,x,\eta,k) = \langle\mathcal{K}^{1,2}\mu_1(t,x,\eta,k)\rangle_\pi. \tag{5.88}$$

The function $\mu_{2,v}(t)$ is given by

$$\mu_{2,v}(t,x,\eta,k) = \int_0^\infty e^{rQ}\{\langle \mathcal{K}^{1,2}\mu_1(t,x,\eta_1+k_1r,\eta_2+k_2r,k)\rangle_\pi \\ -\mathcal{K}^{1,2}\mu_1(t,x,\eta_1+k_1r,\eta_2+k_2r,k)\}dr.$$

Unlike the first corrector $\mu_{1,v}(t)$, the second corrector $\mu_{2,v}(t)$ may not be written as an explicit sum of tensor products even if $\mu(t)$ has the form $\mu(t)=\lambda(t)\otimes\lambda(t)$ because $\mu_{1,v}(t)$ depends on v.

The $\tilde{P}^\varepsilon$-martingale $G^{2,\varepsilon}_{\mu_\varepsilon}(t)$ is given by

$$G^{2,\varepsilon}_{\mu_\varepsilon}(t) = \langle W\otimes W(t),\mu_\varepsilon(t)\rangle - \int_0^t \Big\langle W\otimes W(s), \Big(\nabla_{s,1,2}+\tilde{\mathcal{L}}^\varepsilon_2\Big)\mu(s)\Big\rangle\, ds \\ -\sqrt{\varepsilon}\int_0^t \langle W\otimes W(s),\zeta^\mu_\varepsilon(s)\rangle ds, \tag{5.89}$$

where $\zeta^\mu_\varepsilon(t) := \mathcal{K}^{1,2}\mu^\varepsilon_2(t) + \nabla_{s,1,2}(\mu^\varepsilon_1+\sqrt{\varepsilon}\mu^\varepsilon_2)$ and the operator $\tilde{\mathcal{L}}^\varepsilon_2$ is defined by

$$\tilde{\mathcal{L}}^\varepsilon_2\mu(t,x,k) := -\textstyle\sum_{j=1}^2 \mathcal{A}_j\mu(t,x,k) + \mathcal{B}_\varepsilon\mu(t,x,k) \tag{5.90}$$

where for any function $f(x,k)$, with $x=(x_1,x_2)$, $k=(k_1,k_2)$,

$$\mathcal{B}_\varepsilon f(x,k) := \frac{1}{(2\pi)^d}\sum_{\sigma_1,\sigma_2=\pm 1}\sigma_1\sigma_2\int_0^\infty\int_{\mathbb{R}^d}\tilde{R}(r,p) \\ \times\Big(e^{ip\cdot\frac{x_2-x_1}{\varepsilon}+irk_2\cdot p} + e^{i\frac{x_1-x_2}{\varepsilon}\cdot p+irk_1\cdot p}\Big) f(x,k_{\sigma_1,\sigma_2}(p))drdp,$$

$$\mathcal{A}_1 f(x,k) := -\frac{1}{(2\pi)^d}\sum_{\sigma=\pm1}\int_0^\infty\int_{\mathbb{R}^d}\tilde{R}(r,p)e^{irk^{(1)}_\sigma(p)\cdot p}[f(x,k)-f(x,k^{(1)}_\sigma(p))]drdp,$$

$\mathcal{A}_2$ acts analogously but on the variable k_2. Here for any $\sigma=\pm1$, we let $k^{(1)}_\sigma(p) := (k_1+\sigma p/2,k_2)$ and $k^{(2)}_\sigma(p) := (k_1,k_2+\sigma p/2)$. For any $\sigma_1,\sigma_2=\pm1$ we also let $k_{\sigma_1,\sigma_2}(p) := (k_1+\sigma_1 p/2, k_2+\sigma_2 p/2)$.

Recall that $\tilde{R}(r,p)$ is given by (5.55). We have used in the calculation of $\tilde{\mathcal{L}}^\varepsilon_2$ that for a sufficiently regular function f, we have

$$\Big\langle \int_{\mathbb{R}^d}\frac{\hat{V}(dq)}{(2\pi)^d}\int_0^\infty dr\, e^{rQ}\int_{\mathbb{R}^d}\hat{V}(dp)f(r,p,q)\Big\rangle = \int_0^\infty dr\int_{\mathbb{R}^d}\tilde{R}(r,p)f(r,p,-p)dp.$$

The bound on $\zeta^\mu_\varepsilon(t)$ is similar to that on $\zeta^\lambda_\varepsilon(t)$ obtained previously as the correctors $\mu^\varepsilon_j(t)$ satisfy the same kind of estimates as the correctors $\lambda^\varepsilon_j(t)$:

Lemma 5.13. *There exists a constant $C_\mu>0$ so that the functions $\mu^\varepsilon_{1,2}$ obey the uniform bounds*

$$C^{(2)}_j := \sup_{t\in[0,L]}\sup_{\varepsilon\in(0,1]}\Big\{\Big\langle\|\mu^\varepsilon_j(t)\|^2_{L^2(\mathbb{R}^{2d})}\Big\rangle_\pi \\ +\Big\langle\|\nabla_{t,1,2}\mu^\varepsilon_j(t)\|^2_{L^2(\mathbb{R}^{2d})}\Big\rangle_\pi\Big\} < +\infty,\quad j=1,2. \tag{5.91}$$

The proof of this lemma is very similar to that of Lemma 5.11 and is therefore omitted. Analogously to what has been done before we can also conclude from this lemma that

$$\Big|\mathbb{E}^{\tilde{P}_\varepsilon}\langle W\otimes W(t),\mu_\varepsilon(t)\rangle - \mathbb{E}^{\tilde{P}_\varepsilon}\langle W\otimes W(t),\mu(t)\rangle\Big| \le C\sqrt{\varepsilon}$$

and $\mathbb{E}^{\tilde{P}_\varepsilon}\|\zeta^\mu_\varepsilon(t)\|^2_{L^2(\mathbb{R}^{2d})} \le C$ for all $t \in [0, L]$. These facts allow us to take the limit of the first and third term on the right hand side of (5.89). Unlike the first moment case, however the averaged operator $\tilde{\mathcal{L}}^\varepsilon_2$ still depends on ε. Therefore in order to prove the convergence we have to deal with this term as well. The a priori bound on W_ε in L^2 allows us to tackle this issue and show strong convergence. This is shown as follows. The terms corresponding to $\mathcal{A}_j$ are independent of ε. For example $\mathcal{A}_1$ gives the following contribution:

$$\begin{aligned}\mathcal{A}_1\mu(t,x,k) = \int_0^\infty dr \int_{\mathbb{R}^d} \frac{dp}{(2\pi)^d} \Big[&\tilde{R}(r, p-k_1)e^{ir\frac{p^2-k_1^2}{2}}(\mu(t,x,p,k_2) - \mu(t,x,k_1,k_2)) \\ &+\tilde{R}(r, k_1-p)e^{ir\frac{k_1^2-p^2}{2}}(\mu(t,x,p,k_2) - \mu(t,x,k_1,k_2))\Big].\end{aligned}$$

Using the fact that $\tilde{R}(-r,-q) = \tilde{R}(r,q)$ and integrating out the r variable we conclude that $\mathcal{A}_1\mu(t,x,k) = \mathcal{L}_{k_1}\mu(t,x,k)$, where $\mathcal{L}_{k_1}$ denotes the operator $\mathcal{L}$ acting in the k_1 variable. Likewise we get $\mathcal{A}_2\mu(t,x,k) = \mathcal{L}_{k_2}\mu(t,x,k)$.

The term corresponding to $\mathcal{B}_\varepsilon$ is oscillatory and its contribution tends to 0 as $\varepsilon \to 0$ for sufficiently smooth test functions.

Since $\mathcal{B}_\varepsilon\mu(t)$ and $\mu(t)$ are real valued quantities, we can take the real part of the above term and, after the change of variables $r \to -r$ and $p \to -p$, obtain from (5.91)

$$\begin{aligned}&\mathcal{B}_\varepsilon\mu(t,x,k) \\ &= \tfrac{1}{(2\pi)^d}\sum_{\sigma_1,\sigma_2=\pm1}\sigma_1\sigma_2\int_0^\infty\int_{\mathbb{R}^d}\tilde{R}(r,p) \\ &\quad\times\cos(p\cdot\tfrac{x_2-x_1}{\varepsilon})\cos(r(k_1+k_2)\cdot p/2)\mu(t,x,k_{\sigma_1,\sigma_2}(p))drdp \\ &= \tfrac{1}{2(2\pi)^d}\sum_{\sigma_1,\sigma_2=\pm1}\sigma_1\sigma_2\int_{\mathbb{R}^d}\hat{R}\left(-(k_1+k_2)\cdot\frac{p}{2},p\right)\exp\left\{ip\cdot\frac{x_2-x_1}{\varepsilon}\right\}\mu(t,x,k_{\sigma_1,\sigma_2}(p))dp.\end{aligned}$$

We have

$$\begin{aligned}\|\mathcal{B}_\varepsilon\mu(t)\|^2_{L^2(\mathbb{R}^{4d})} = C\sum_{\sigma_j=\pm1,j=1,\dots,4}\prod_{j=1}^4\sigma_j\int_{\mathbb{R}^{6d}}&dxdkdpdq\hat{R}\left(-(k_1+k_2)\cdot\frac{p}{2},p\right) \\ &\times\hat{R}\left(-(k_1+k_2)\cdot\tfrac{q}{2},q\right)e^{i(p-q)\cdot\frac{x_2-x_1}{\varepsilon}}\mu(t,x,k_{\sigma_1,\sigma_2}(p))\mu(t,x,k_{\sigma_3,\sigma_4}(p)).\end{aligned} \tag{5.92}$$

Using the density argument we may assume that μ has the form

$$\mu(t,x,k) = \mu_0(t)\mu_1(x_1-x_2)\mu_2(x_1+x_2)\mu_3(k_1)\mu_4(k_2).$$

Then the expression on the right hand side of (5.94) equals

$$\begin{aligned}&C\mu_0^2(t)\sum_{\sigma_j=\pm1,j=1,\dots,4}\prod_{j=1}^4\sigma_j\int_{\mathbb{R}^{6d}}dxdkdpdq\hat{R}\left(-(k_1+k_2)\cdot\frac{p}{2},p\right)\hat{R}\left(-(k_1+k_2)\cdot\frac{q}{2},q\right) \\ &\times e^{-i(p-q)\cdot\frac{x_1}{\varepsilon}}\mu_1^2(x_1)\mu_2^2(x_2)\mu_3(k_1+\tfrac{\sigma_1 p}{2})\mu_4(k_2+\tfrac{\sigma_2 p}{2})\mu_3(k_1+\tfrac{\sigma_3 q}{2})\mu_4(k_2+\tfrac{\sigma_4 q}{2}) \\ &= C\mu_0^2(t)\|\mu_2\|^2_{L^2(\mathbb{R}^d)}\sum_{\sigma_j=\pm1,j=1,\dots,4}\prod_{j=1}^4\sigma_j\int_{\mathbb{R}^{4d}}dkdpdq\hat{R}\left(-(k_1+k_2)\cdot\frac{p}{2},p\right) \\ &\times\hat{R}\left(-(k_1+k_2)\cdot\tfrac{q}{2},q\right)\hat{\nu}(\tfrac{p-q}{\varepsilon})\mu_3(k_1+\tfrac{\sigma_1 p}{2})\mu_4(k_2+\tfrac{\sigma_2 p}{2})\mu_3(k_1+\tfrac{\sigma_3 q}{2})\mu_4(k_2+\tfrac{\sigma_4 q}{2})\end{aligned}$$

where $\nu(x) = \mu_1^2(x)$. We introduce $G(p) = \sup_\omega \hat{R}(\omega, p)$ and use the Cauchy-Schwarz inequality in k_1 and k_2:

$$\|\mathcal{B}_\varepsilon\mu(t)\|^2_{L^2(\mathbb{R}^{4d})} \leq C\mu_0^2(t)\|\mu_2\|^2_{L^2(\mathbb{R}^d)}\|\mu_3\|^2_{L^2(\mathbb{R}^d)}\|\mu_4\|^2_{L^2(\mathbb{R}^d)} \times \int_{\mathbb{R}^{2d}} dpdqG(p)G(q)\left|\hat{\nu}(\frac{p-q}{\varepsilon})\right|.$$

We use again the Cauchy-Schwarz inequality, now in p, to get

$$\begin{aligned}\|\mathcal{B}_\varepsilon\mu(t)\|^2_{L^2(\mathbb{R}^{4d})} &\leq C\|\mu_2\|^2_{L^2(\mathbb{R}^d)}\|\mu_3\|^2_{L^2(\mathbb{R}^d)}\|\mu_4\|^2_{L^2(\mathbb{R}^d)}\|G\|_{L^2(\mathbb{R}^d)} \\ &\quad\times \int_{\mathbb{R}^d} dqG(q)\left(\int_{\mathbb{R}^d} dp\left|\hat{\nu}(\frac{p}{\varepsilon})\right|^2\right)^{1/2} \\ &\leq C\varepsilon^{d/2}\|\mu_2\|^2_{L^2(\mathbb{R}^d)}\|\mu_3\|^2_{L^2(\mathbb{R}^d)}\|\mu_4\|^2_{L^2(\mathbb{R}^d)}\|G\|_{L^2(\mathbb{R}^d)}\|G\|_{L^1(\mathbb{R}^d)}\|\nu\|_{L^2}.\end{aligned}$$

This proves that $\|\mathcal{B}_\varepsilon\mu(t)\|_{L^2(\mathbb{R}^{4d})} \to 0$ as $\varepsilon \to 0$. Note that oscillatory integrals of the form

$$\int_{\mathbb{R}^d} e^{i\frac{p\cdot x}{\varepsilon}}\mu(p)dp, \tag{5.93}$$

are not small in the bigger space $\mathcal{A}'$, which is natural in the context of Wigner transforms. In this bigger space, we cannot control the norm of $\mathcal{B}_\varepsilon\mu(t)$ and actually do not know whether the limit measure P is deterministic; see [10].

We have therefore deduced that

$$G_\mu^2(t) = \langle W\otimes W(t), \mu(t)\rangle - \int_0^t \left\langle W\otimes W(s), \Big(\frac{\partial}{\partial s} + k_1\cdot\nabla_{x_1} + k_2\cdot\nabla_{x_2} + \mathcal{L}_{k_1} + \mathcal{L}_{k_2}\Big)\mu\right\rangle(s)ds$$

is an approximate $\tilde{P}_\varepsilon$ martingale. Suppose that P is a weak limiting measure for the family $\tilde{P}_\varepsilon$, as $\varepsilon \to 0$. The limit of the second moment

$$W_2(t,x,k) = \mathbb{E}^P\left\{W\otimes W(t,x,k)\right\}$$

thus satisfies (weakly) the transport equation

$$\frac{\partial W_2}{\partial t} + (k_1\cdot\nabla_{x_1} + k_2\cdot\nabla_{x_2})W_2 = (\mathcal{L}_{k_1} + \mathcal{L}_{k_2})W_2, \tag{5.94}$$

with the initial data $W_2(0,x,k) = W_0(x_1,k_1)W_0(x_2,k_2)$. This implies that

$$\mathbb{E}^P\left\{W(t,x_1,k_1)W(t,x_2,k_2)\right\} = \mathbb{E}^P\left\{W(t,x_1,k_1)\right\}\mathbb{E}^P\left\{W(t,x_2,k_2)\right\}$$

by uniqueness of the solution to (5.94) with initial conditions given by the product $W_0(x_1,k_1)W_0(x_2,k_2)$. This proves that the limiting measure P is deterministic and unique (because characterized by the transport equation) and thus $W_\varepsilon(t,x,k)$ converges in probability to $W(t,x,k)$, as $\varepsilon \to 0$.

Tightness of P_ε.

Our principal result is the following.

Lemma 5.14. *The family of measures P_ε is weakly compact in $C([0,+\infty); L^2_w(\mathbb{R}^{2d}))$.*

Proof. A theorem of Mitoma and Fouque [65, 40] implies that in order to verify tightness of the family P_ε it is enough to check that for each $\lambda \in C^1([0,T], \mathcal{S}(\mathbb{R}^{2d}))$ the family of laws $\mathcal{P}_\varepsilon$ on $C([0,T];\mathbb{R})$ of the random processes $W_\lambda(t) = \langle W(t), \lambda(s)\rangle$, considered over $(C([0,T]; L^2_w(\mathbb{R}^{2d})), \mathcal{M}_T, P_\varepsilon)$, is tight.

Note that we have

$$\langle W(t),\lambda(t)\rangle = G^\varepsilon_{\lambda_\varepsilon}(t) - \sqrt{\varepsilon}\langle W(t),\lambda_1^\varepsilon(t)\rangle - \varepsilon\langle W(t),\lambda_2^\varepsilon(t)\rangle + \int_0^t \langle W(s),\hat{\mathcal{L}}_s\lambda(s)\rangle ds + \sqrt{\varepsilon}\int_0^t \langle W(s),\zeta^\lambda_\varepsilon(s)\rangle ds,$$

where $\hat{\mathcal{L}}_t := \frac{\partial}{\partial t} + k\cdot\nabla_x + \mathcal{L}$. Denote

$$\begin{aligned} x_\varepsilon^{(1)}(t) &:= G^\varepsilon_{\lambda_\varepsilon}(t)\\ x_\varepsilon^{(2)}(t) &:= \sqrt{\varepsilon}\langle W(t),\lambda_1^\varepsilon(t)\rangle,\\ x_\varepsilon^{(3)}(t) &:= \int_0^t \langle W(s),\hat{\mathcal{L}}_s\lambda(s)\rangle ds,\\ x_\varepsilon^{(4)}(t) &:= \varepsilon\langle W(t),\lambda_2^\varepsilon(t)\rangle,\\ x_\varepsilon^{(5)}(t) &:= \sqrt{\varepsilon}\int_0^t \langle W(s),\zeta^\lambda_\varepsilon(s)\rangle ds. \end{aligned}$$

Tightness of the family $\mathcal{P}_\varepsilon$, as $\varepsilon\to 0$, follows upon proving:

Lemma 5.15. *Each of $\{x_\varepsilon^{(j)}(t),\, t\ge 0\}$, $j=1,\dots,5$ is tight in $C[0,+\infty)$*

Proof. We claim that for $j=2,\dots 5$ the processes $\{x_\varepsilon^{(j)}(t),\, t\ge 0\}$ satisfy the following Kolmogorov-Chentzov moment condition [23]

$$\mathbb{E}^{P_\varepsilon}\left\{\left|x_\varepsilon^{(j)}(t)-x_\varepsilon^{(j)}(u)\right|^\gamma \left|x_\varepsilon^{(j)}(u)-x_\varepsilon^{(j)}(s)\right|^\gamma\right\} \le C_\lambda (t-s)^{1+\beta}, \tag{5.95}$$

for all $0\le s\le u\le t\le T$, with $\gamma,\beta>0$ and C_λ independent of ε. It implies tightness, see the explanation given after (2.29).

To demonstrate (5.95) we shall only discuss the case $j=2$, as the other cases can be done similarly. Let $K_\varepsilon[t,k,x,v] := \partial_t + k\cdot\nabla_x + (1/\sqrt{\varepsilon})\mathcal{K}[v,x/\varepsilon]$. We have

$$\begin{aligned}\mathbb{E}^{\tilde P_\varepsilon}\left\{\left|x_\varepsilon^{(2)}(t)-x_\varepsilon^{(2)}(s)\right|^2\right\} &\le 2\left\{\mathbb{E}^{\tilde P_\varepsilon}\left|\int_s^t \langle\sqrt{\varepsilon}W(\tau),K_\varepsilon[\tau,x,k,V(\tau)]\lambda_1^\varepsilon(t)\rangle d\tau\right|^2 \right.\\ &\left.\quad +2\mathbb{E}^{\tilde P_\varepsilon}\left|\int_s^t \langle W(s),\sqrt{\varepsilon}\partial_\tau\lambda_1^\varepsilon(\tau)\rangle d\tau\right|^2\right\}\\ &\le C(t-s)^2,\end{aligned}$$

thanks to the bounds established in Lemmas 5.11 and 5.12. Estimate (5.95) then follows for $\{x_\varepsilon^{(2)}(t),\, t\ge 0\}$, with $\gamma=\beta=1$, upon an application of Cauchy-Schwartz inequality and (5.96).

For $j=1$ the processes $\{G^\varepsilon_{\lambda_\varepsilon}(t), t\ge 0\}$ are martingales. Their tightness follows from tightness of their quadratic variation $\{\langle G^\varepsilon_{\lambda_\varepsilon}\rangle_t,\, t\ge 0\}$, see Theorem VI.4.13, p. 358 of [50]. We will now compute it explicitly. First,

$$\begin{aligned}&\tfrac{d}{dh}\mathbb{E}^{P_\varepsilon}_{W,v,t}\left\{\langle W(t+h),\lambda_\varepsilon(t+h)\rangle^2\right\}\Big|_{h=0}\\ &= 2\langle W,\lambda_{\varepsilon,v}(t)\rangle\langle W,K_\varepsilon[t,k,x,v]\lambda_{\varepsilon,v}(t)\rangle + \tfrac{1}{\varepsilon}Q\left[\langle W,\lambda_{\varepsilon,v}\rangle^2\right]\end{aligned}$$

so that

$$\langle W(t),\lambda_\varepsilon(t)\rangle^2 - \int_0^t \left\{2\langle W(s),\lambda_\varepsilon(s)\rangle\langle W(s),K_\varepsilon[s,k,x,V(s)]\lambda_\varepsilon(s)\rangle + \frac{1}{\varepsilon}Q\left[\langle W(s),\lambda_\varepsilon(s)\rangle^2\right](V(s))\right\}ds$$

is a martingale. Using (5.72), we conclude that

$$\langle G^{\varepsilon}_{\lambda_\varepsilon}\rangle_t = \int_0^t \left\{\frac{1}{\varepsilon}Q[\langle W(s),\lambda_\varepsilon(s)\rangle^2](V(s)) - \frac{2}{\varepsilon}\langle W(s),\lambda_\varepsilon(s)\rangle\langle W(s),Q\lambda_\varepsilon(s)\rangle\right\}ds$$
$$= \int_0^t \left\{Q\left[\langle W(s),\lambda_1^\varepsilon(s)\rangle^2\right] - \langle W(s),\lambda_1^\varepsilon(s)\rangle\langle W(s),Q\lambda_1^\varepsilon(s)\rangle\right\}ds$$
$$+\sqrt{\varepsilon}\int_0^t H_\varepsilon(s,W(s),V(s))ds$$

with

$$H_\varepsilon(s,W,v) := 2\sqrt{\varepsilon}\left\{Q[\langle W,\lambda_{1,v}^\varepsilon(s)\rangle\langle W,\lambda_{2,v}^\varepsilon(s)\rangle] - \langle W,\lambda_{1,v}^\varepsilon(s)\rangle\langle W,(Q\lambda_2^\varepsilon(s))(v)\rangle\right.$$
$$\left. -\langle W,\lambda_{2,v}^\varepsilon(s)\rangle\langle W,Q\lambda_{1,v}^\varepsilon(s)\rangle\right\} + \varepsilon\left\{Q[\langle W,\lambda_2^\varepsilon(s)\rangle^2]\right.$$
$$\left. -2\langle W,\lambda_{2,v}^\varepsilon(s)\rangle\langle W,(Q\lambda_2^\varepsilon(s))(v)\rangle\right\}.$$

Recall that $\sup_{t\ge 0}\|W(t)\|_{L^2(\mathbb{R}^{2d})}$ remains deterministically bounded $\tilde P_\varepsilon$ a.s. Using bounds on $\|\lambda_j^\varepsilon\|_{L^2(\pi)}$, $j=1,2$ from Lemma 5.11 and the fact that Q is a contraction on both $L^1(\pi)$ and $L^2(\pi)$ we conclude that $\mathbb{E}^{\tilde P_\varepsilon}|H_\varepsilon(s)|\le C$ for all $s\in[0,T]$. This yields $\mathbb{E}^{\tilde P_\varepsilon}\left\{\langle G^\varepsilon_{\lambda_\varepsilon}\rangle_t - \langle G^\varepsilon_{\lambda_\varepsilon}\rangle_s\right\}\le C(t-s)$ whence for a fixed $T>0$ the family of random variables $\langle G^\varepsilon_{\lambda_\varepsilon}\rangle_T$ is tight, as $\varepsilon\to 0$. By gathering of Theorems 3.8 and 3.10 and Remark 3.9 of [51] this demonstrates tightness of the family of increasing processes $\{\langle G^\varepsilon_{\lambda_\varepsilon}\rangle_t,\, t\in[0,T]\}$, as $\varepsilon\to 0$, in $D[0,T]$, thus in our case in $C[0,T]$ as well. This ends the proof of tightness.

5.6. **Fluctuations of the Wigner transform with an OU potential.** Now, that we know that (at least in some situations) the Wigner transform of the solutions of the random Schrödinger equation converges in probability to the deterministic solution of the kinetic radiative transport equation, we would like to understand how its fluctuations behave. This question was partially addressed in Section 5.4 where we have discussed the scintillation of the Wigner transform in the Itô-Schrödinger regime. Here, following [56, 59] we describe what happens for the full fluctuation process (scintillation is just the second moment of the fluctuation) when the initial data for the Wigner transform is localized in space.

As before, we consider the Schrödinger equation with a weakly random potential, in the appropriate long time, large distance scaling:

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

The Wigner transform of the solution (without any consideration of mixtures of states for the moment), defined as

$$W_\varepsilon(t,x,k) = \int e^{ik\cdot y}\phi_\varepsilon(t,x-\frac{\varepsilon y}{2})\bar\phi_\varepsilon(t,x+\frac{\varepsilon y}{2})\frac{dy}{(2\pi)^d},$$

satisfies

$$\frac{\partial W_\varepsilon(t,x,k)}{\partial t} + k\cdot\nabla_x W_\varepsilon(t,x,k) = \frac{i}{\sqrt{\varepsilon}}\sum_{\sigma=\pm 1}\sigma\int\frac{\hat V(t/\varepsilon,dp)}{(2\pi)^d}e^{ip\cdot z}W_\varepsilon\left(t,x,k+\frac{\sigma p}{2}\right). \tag{5.96}$$

As we have seen in the previous section, when the initial data W_0 for (5.96) is in $L^2(\mathbb{R}^{2d})$ the solutions converge in probability, as $\varepsilon\downarrow 0$, to $\bar W(t,x,k)$ the solution of a linear Boltzmann equation

$$\partial_t \bar{W}(t,x,k) + k\cdot\nabla_x \bar{W}(t,x,k) = \mathcal{L}\bar{W}(t,x,k), \qquad (5.97)$$
$$\bar{W}(0,x,k) = W_0(x,k),$$

where the operator $\mathcal{L}$ is given by

$$\mathcal{L}W(x,k) = \int \hat{R}(\frac{p^2-k^2}{2}, p-k)(W(x,p)-W(x,k))\frac{dp}{(2\pi)^{2d}}.$$

It is important, in particular, for inverse problems, to understand the fluctuations of W_ε around this self-averaging limit, as wave energy fluctuations are often large in practice [6, 7, 14] despite being theoretically small. As we have seen in the Itô-Schrödinger case, the size of the fluctuations depends on the regularity of the initial W_0 – both spatially and wave vector localized singularities in W_0 produce stronger fluctuations than smooth initial energy distributions. Here, we consider the fluctuations of the Wigner transform

$$Z_\varepsilon(t,x,k) = \varepsilon^{-1/2}[W_\varepsilon(t,x,k) - \bar{W}(t,x,k)]$$

when $W_0(x,k) = \delta(x)f(k)$ and $f \in \mathcal{S}(\mathbb{R}^d)$, that is, the initial wave energy distribution is spatially localized but smoothly distributed in various directions. The fact that the fluctuations have the size $O(\sqrt{\varepsilon})$ comes from the spatial singularity of the initial data – their size would be smaller were $W_0(x,k)$ more regular in space. However, physically this is a very important case – the localized source.

We assume that the random potential $V(t,x)$ is of the Ornstein-Uhlenbeck type – see [59] for the detailed construction of such potentials. The main result, which we will describe informally to avoid the rather technical but physically irrelevant details, is as follows. Let $\bar{Z}(t)$ be the solution of the deterministic kinetic equation

$$\frac{\partial \bar{Z}}{\partial t} + k\cdot\nabla_x \bar{Z} = \int \hat{R}\left(\frac{|k|^2-|p|^2}{2}, k-p\right)(W(t,x,p)-W(t,x,k))dp, \qquad (5.98)$$

with the random initial data $\bar{Z}(0,x,k) = \delta(x)X(k)$. Here X is a real valued Gaussian distribution that can be written down explicitly in terms of the random potential (see [59] for the explicit formula). Informally, $X(k)$ is obtained as follows, from the initial layer problem for the fluctuation. In the fast variables $s = t/\varepsilon$, $y = x/\varepsilon$ equation (5.96) may be re-written as

$$\frac{\partial W'_\varepsilon}{\partial s} + k\cdot\nabla_y W'_\varepsilon = -i\frac{\sqrt{\varepsilon}}{(2\pi)^d}\sum_{\sigma=\pm 1}\sigma\int_{\mathbb{R}^d} e^{ip\cdot y}\hat{V}(s,dp)W'_\varepsilon(s,y,k+\frac{\sigma p}{2}),$$

where $W'_\varepsilon(s,y,k) := W_\varepsilon(\varepsilon s, \varepsilon y, k)$. We introduce a formal asymptotic expansion

$$W'_\varepsilon(s,y,k) = \bar{W}'(s,y,k) + \sqrt{\varepsilon}Z'(s,y,k) + \ldots$$

The leading order term satisfies the homogeneous transport equation

$$\bar{W}'_s + k\cdot\nabla_y \bar{W}' = 0,\ \bar{W}'(0,y,k) = \varepsilon^{-d}\delta(y)f(k),$$

and is, therefore, given by $\bar{W}'(s,y,k) = \varepsilon^{-d}\delta(y-ks)f(k)$. The equation for $Z'(s,y,k)$ is

$$\partial_s Z'(s,y,k) + k\cdot\nabla_y Z'(s,y,k) = -i\sum_{\sigma=\pm 1}\sigma\int_{\mathbb{R}^d} e^{ip\cdot y}\hat{V}(s,dp)\bar{W}'\left(s,y,k+\frac{\sigma p}{2}\right)$$

with the initial data $Z'(0,y,k)=0$. This gives an explicit formula for $Z(s,y,k)$:

$$Z'(s,y,k) = -i\sum_{\sigma=\pm1}\sigma\int_0^s\int_{\mathbb{R}^d} e^{ip\cdot(y-k(s-\tau))}\bar{W}'(\tau,y-k(s-\tau),k+\frac{\sigma p}{2})\hat{V}(\tau,dp)d\tau$$
$$= -i\varepsilon^{-d}\sum_{\sigma=\pm1}\sigma\int_0^s\int_{\mathbb{R}^d} e^{ip\cdot(k+\sigma p/2)\tau}f\left(k+\frac{\sigma p}{2}\right)\delta(y-ks-\frac{\sigma p}{2}\tau)\hat{V}(\tau,dp)d\tau.$$

We obtain therefore:

$$Z(t,x,k) = Z'(t/\varepsilon,x/\varepsilon,k)$$
$$= -i\varepsilon^{-d}\sum_{\sigma=\pm1}\sigma\int_0^{t/\varepsilon}\int_{\mathbb{R}^d} e^{ip\cdot(k+\sigma p/2)\tau}f\left(k+\frac{\sigma p}{2}\right)$$
$$\times\delta(\varepsilon^{-1}(x-kt+\varepsilon\sigma p\tau/2))\hat{V}(\tau,dp)d\tau$$

and since $\varepsilon^{-d}\delta(z/\varepsilon)=\delta(z)$ we obtain that for small $t\ll 1$ the quantity $\varepsilon\sigma p\tau\le pt\ll 1$ can be neglected, thus

$$Z(0,x,k)\approx -i\delta(x)\sum_{\sigma=\pm1}\sigma\int_0^\infty\int_{\mathbb{R}^d} e^{ip\cdot(k+\sigma p/2)\tau}f\left(k+\frac{\sigma p}{2}\right)\hat{V}(\tau,dp)d\tau = \delta(x)X(k),$$

which gives an explicit expression for $X(k)$.

The physical reason why randomness of Z_ε appears in the limit only as a random initial data for the deterministic radiative transport equation is that after the short initial time layer the leading order term $W(t,x,k)$ is no longer as localized in space as at $t=0$. Therefore, the fluctuation produced at $t>0$ is not of the size $O(\sqrt{\varepsilon})$ but is smaller and it appears only in the higher order terms. Essentially, the main contribution to $W_\varepsilon - W$ at $t>0$ comes only from the evolution of the random fluctuation produced near $t=0$. On the other hand, if the initial data $W_0(x,k)$ would be smooth, we expect the fluctuation of $W_\varepsilon(t,x,k)$ to be of the smaller size and to satisfy a radiative transport equation with a random force. This, however, is currently an open problem.

As the proofs of this fluctuation theorem are somewhat technical, we refer the reader to [56, 59] for the rigorous statements and proofs.

5.7. **Different kinetic regimes for time-dependent Schrödinger.** In the preceding sections, we considered the Schrödinger equation with a random potential whose oscillations in the time domain that were as rapid as its oscillations in the spatial domain. We now discuss how one can generalize the convergence analysis to cases where the temporal random fluctuations are slower than the spatial fluctuations of the potential. We also want to include spatial correlation lengths larger than the wavelength, as in the random Liouville problem (random geometric optics). Finally, we want to be able to pass directly from a wave model to a diffusion model, which is an approximation of the kinetic model in the limit of vanishing mean free path. All of this is included in the following Schrödinger equation with time dependent potential

$$i\varepsilon^{1+\delta}\frac{\partial\psi_\varepsilon}{\partial t}+\frac{\varepsilon^2}{2}\Delta\psi_\varepsilon-\varepsilon^{\frac{\gamma-\delta}{2}}V\Big(\frac{t}{\varepsilon^{\alpha+\delta}},\frac{x}{\varepsilon^\beta}\Big)\psi_\varepsilon=0, \tag{5.99}$$

with the same initial conditions and mixtures of states as in the preceding section. We define $\gamma=\beta+\min(\beta-\alpha,0)$. We consider the case of spatial dimension $d\ge 2$.

Various values of the parameters α, β, γ and δ correspond to various relations between the macroscopic and microscopic scales. The resulting regimes appearing are follows. Set $\delta = 0$ first – this is a "hyperbolic" scaling. Then, $\beta = 1$ corresponds to the radiative transfer regime (wave length is comparable to the spatial correlation length of the medium) with white noise scattering kernel when $\alpha > 1$ (the random medium oscillates faster in time than in space), inelastic scattering when $\alpha = \beta = 1$, and elastic scattering when $\alpha < 1$. Also, $\beta < 1$ corresponds to the Fokker-Planck regime with white noise scattering kernel when $\alpha > \beta$, inelastic scattering when $\alpha = \beta$ (the random media oscillates on comparable temporal and spatial micro-scales), and elastic scattering when $\alpha < \beta$ (the temporal scale of the random medium oscillations is longer than the spatial scale).

When $\delta > 0$ (this is a longer time scale than the radiative transport time scale), we are in the diffusive regime when $\alpha > \beta$. The case of elastic scattering $\alpha < \beta$ should hold for all $\alpha \geq 0$ since we expect the results to hold in the limit of no time-dependent regularization. Here, we describe how the proofs based on the Markovian methods and the perturbed test function method can be developed for the case $\alpha/\beta > 3/4$ and extended to the case $\alpha/\beta > 1/2$ with a reasonable amount of unspecified work. For slower time fluctuations of the medium, other techniques than the Markov regularization considered here presumably need to be developed, and the use of the diagrammatic techniques might be unavoidable.

The equation for the Wigner transform is now

$$\begin{aligned} &\varepsilon^{\delta}\frac{\partial W_\varepsilon}{\partial t} + k\cdot\nabla W_\varepsilon \\ &= \frac{\varepsilon^{\frac{\gamma-\delta}{2}}}{i\varepsilon}\int_{\mathbb{R}^d}\tilde{V}\Big(\frac{t}{\varepsilon^{\alpha+\delta}},p\Big)e^{i\frac{x\cdot p}{\varepsilon^{\beta}}}\Big(W_\varepsilon(t,x,k-\frac{\varepsilon^{1-\beta}p}{2}) - W_\varepsilon(t,x,k+\frac{\varepsilon^{1-\beta}p}{2})\Big)\frac{dp}{(2\pi)^d}. \end{aligned} \tag{5.100}$$

We assume that the above equation is augmented with initial conditions of the form $W_\varepsilon(0,x,k) = W_0(x,|k|)$, which, to simplify, we assume is independent of ε and $\hat{k}$.

5.7.1. *Convergence to various kinetic models.* The main results are as follows

Theorem 5.16 (Transport and Fokker-Planck regimes)**.** *Let $\delta = 0$ and $d \geq 2$. Then W_ε converges weakly in $L^2(\mathbb{R}^{2d})$ and in probability to the solution $W(t,x,k)$ of the following transport equation*

$$\frac{\partial W}{\partial t} + k\cdot\nabla_x W = \mathcal{L}W, \qquad W(0,x,k) = W_0(x,|k|), \tag{5.101}$$

where the scattering kernel has the following form

$$\frac{3}{4} < \alpha < \beta = 1 \qquad \mathcal{L}W = \int_{\mathbb{R}^d}\hat{R}_0(p-k)\Big(W(p)-W(k)\Big)\delta\Big(\frac{|k|^2-|p|^2}{2}\Big)\frac{dp}{(2\pi)^d} \tag{5.102}$$

$$\alpha = \beta = 1 \qquad \mathcal{L}W = \int_{\mathbb{R}^d}\hat{R}(\frac{|k|^2-|p|^2}{2},p-k)\Big(W(p)-W(k)\Big)\frac{dp}{(2\pi)^d}$$

$$\alpha > \beta = 1 \qquad \mathcal{L}W = \int_{\mathbb{R}^d}\hat{R}(0,p-k)\Big(W(p)-W(k)\Big)\frac{dp}{(2\pi)^d}$$

$$\frac{3\beta}{4} < \alpha < \beta < 1 \qquad \mathcal{L}W = \frac{1}{2}\nabla_k \cdot \Big(\int_{\mathbb{R}^d} \delta(k\cdot p)\hat{R}_0(p) p\otimes p \frac{dp}{(2\pi)^d}\Big)\nabla_k W \tag{5.103}$$

$$\alpha = \beta < 1 \qquad \mathcal{L}W = \frac{1}{2}\nabla_k \cdot \Big(\int_{\mathbb{R}^d} \hat{R}(k\cdot p, p) p\otimes p \frac{dp}{(2\pi)^d}\Big)\nabla_k W$$

$$\alpha > \beta < 1 \qquad \mathcal{L}W = \frac{1}{2}\nabla_k \cdot \Big(\int_{\mathbb{R}^d} \hat{R}(0, p) p\otimes p \frac{dp}{(2\pi)^d}\Big)\nabla_k W.$$

Theorem 5.17 (Diffusion regime)**.** *Let $\delta(\alpha,\beta) > 0$ sufficiently small and $d \geq 2$. Then W_ε converges weakly in $L^2(\mathbb{R}^{2d})$ and in probability to the solution $W(t,x,|k|)$ of the following diffusion equation*

$$\frac{\partial W}{\partial t} - \nabla_x \cdot \Big(\frac{1}{\Gamma_d |k|^{d-1}}\int_{\mathbb{R}^d} p\otimes \boldsymbol{\chi}(p)\delta(|p|-|k|)dp\Big)\nabla_x W = 0, \quad W(0,x,|k|) = W_0(x,|k|), \tag{5.104}$$

where Γ_d is the volume S^{d-1} and $\boldsymbol{\chi}$ is the mean-zero solution of the system of equations $k = \mathcal{L}\boldsymbol{\chi}$, where $\mathcal{L}$ is given by (5.103) for $\frac{3\beta}{4} < \alpha < \beta < 1$ and by (5.102) for $\frac{3}{4} < \alpha < \beta = 1$.

We thus observe that the different choices of correlation lengths and mean free paths may yield to all "classical" kinetic regimes of wave propagation in random media, namely, the radiative transfer, Fokker Planck, and diffusion regimes.

The case of kinetic models for Schrödinger equations with time independent potentials would formally correspond to the case $\alpha = 0$ and $\beta = 1$ (and $\delta = 0$). The proofs do not extend to this case. The case of Fokker-Planck models with time independent fluctuations would correspond to the case $\alpha = 0$ and $\beta < 1$ (with $\delta = 0$ still). This is also not accessible by the current method of proof. Note however, that Fokker Planck models of the types seen before occur as soon as $\alpha < \beta < 1$. This means that Fokker Planck models occur as soon as the correlation length is significantly larger than the wavelength. In the case $\alpha = 0$ treated in earlier sections, we need the correlation length to be very large compared to the wavelength (larger than $|\ln\varepsilon|^{-1}$). We do not require such a large gap in the presence of time dependent potentials.

5.7.2. *Sketch of the derivation.* The derivation of the above results is very similar to that of the radiative transfer model obtained in Section 5.5. The main difference is that the approximate martingales we now construct involve an averaged operator $\mathcal{L}_\varepsilon$ that depends on ε and thus needs to be approximated by an $\varepsilon-$independent operator $\mathcal{L}$. We mainly outline here the construction of the approximate martingale to obtain convergence of the expectation and leave to the reader the extension of the proof of convergence of higher moments and tightness of the measures generated by the Wigner transforms.

Convergence in expectation.

We consider the conditional expectation of functionals $F(W,v)$ with respect to the probability measure $\tilde{P}_\varepsilon$ on $D([0,L]; B_W \times \mathcal{V})$, the space of right-continuous paths with left-side limits [23] generated by the process (W,v). Note that W is a continuous function in t thanks to the evolution equation it solves. The process V however need not be continuous. Given a function $F(W,\hat{V})$ let us define the conditional expectation

$$\mathbb{E}^{\tilde{P}_\varepsilon}_{W,\hat{V},t}\Big\{F(W,\hat{V})\Big\}(\tau) = \mathbb{E}^{\tilde{P}_\varepsilon}\Big\{F(W(\tau),\tilde{V}(\tau))|\, W(t) = W, \tilde{V}(t) = \hat{V}\Big\}, \quad \tau \geq t.$$

For a function of the form $F(W,\hat V)=\langle W,\lambda(\hat V)\rangle$ with $\lambda\in L^\infty(\mathcal{V};C^1([0,L];\mathcal{S}(\mathbb{R}^{2d})))$, the weak form of the infinitesimal generator of the Markov process generated by $\tilde P_\varepsilon$ is then given by

$$\begin{aligned}&\tfrac{d}{dh}\mathbb{E}^{\tilde P_\varepsilon}_{W,\hat V,t}\left\{\langle W,\lambda(\hat V)\rangle\right\}(t+h)\Big|_{h=0}\\&=\tfrac{1}{\varepsilon^\alpha}\langle W,Q\lambda\rangle+\left\langle W,\left(\varepsilon^\delta\frac{\partial}{\partial t}+k\cdot\nabla_x+\tfrac{\varepsilon^{\frac{\gamma-\delta}{2}}}{\varepsilon}\mathcal{K}[\hat V,\tfrac{x}{\varepsilon^\beta}]\right)\lambda\right\rangle,\end{aligned}\tag{5.105}$$

where the operator $\mathcal{K}$ is defined as

$$\mathcal{K}[\hat V,y]\psi(x,y,k,\hat V)=\frac{1}{i}\int_{\mathbb{R}^d}\frac{d\hat V(p)}{(2\pi)^d}e^{ip\cdot y}\left[\psi(x,y,k-\frac{\eta p}{2})-\psi(x,y,k+\frac{\eta p}{2})\right],\tag{5.106}$$

where $\eta=\varepsilon^{1-\beta}$. Note that the operator $\mathcal{K}$ applied to smooth functions is of order $O(\eta)$.

Let $\mathcal{F}_t$ be the sub-σ-algebra on $D([0,L];B_W\times\mathcal{V})$ generated by $((V_\varepsilon(s),W_\varepsilon(s)),0\le s\le t)$. Then $(G^\varepsilon_\lambda(t),\mathcal{F}_t)$, where

$$G^\varepsilon_\lambda(t)=\langle W,\lambda(\hat V)\rangle(t)-\int_0^t\left\langle W,\left(\frac{1}{\varepsilon^\alpha}Q+\varepsilon^\delta\frac{\partial}{\partial t}+k\cdot\nabla_x+\frac{\varepsilon^{\frac{\gamma-\delta}{2}}}{\varepsilon}\mathcal{K}[\hat V,\frac{x}{\varepsilon^\beta}]\right)\lambda\right\rangle(s)ds,\tag{5.107}$$

is a $\tilde P_\varepsilon$-martingale.

Given a test function $\lambda(z,x,k)\in C^1([0,L];\mathcal{S})$ we construct a function

$$\lambda_\varepsilon(z,x,k,\hat V)=\lambda(z,x,k)+\lambda_1^\varepsilon(z,x,k,\hat V)+\lambda_2^\varepsilon(z,x,k,\hat V),\tag{5.108}$$

where the functions $\lambda^\varepsilon_{1,2}(t,x,k,\hat V)=\lambda_{1,2}(t,x,\varepsilon^{-\beta}x,k,\hat V)$ will be chosen to remove high-order terms in the definition of the martingale (5.107). For functions of the form $\lambda(t,x,y,k,\hat V)$, the operator in the brackets in (5.107) takes the form

$$\varepsilon^{-\alpha}Q+\varepsilon^\delta\frac{\partial}{\partial t}+\varepsilon^{-\beta}k\cdot\nabla_y+k\cdot\nabla_x+\varepsilon^{-1+\frac{\gamma-\delta}{2}}\mathcal{K}[\hat V,y].$$

The first corrector is defined as the mean-zero (with respect to the invariant measure of V) solution to the Poisson equation

$$\varepsilon^{-\alpha}Q\lambda_1+\varepsilon^{-\beta}k\cdot\nabla_y\lambda_1+\varepsilon^{-1+\frac{\gamma-\delta}{2}}\mathcal{K}[\hat V,y]\lambda=0.\tag{5.109}$$

The latter solution is given explicitly by

$$\lambda_1=\frac{\varepsilon^{\alpha+\frac{\gamma-\delta}{2}}}{i\varepsilon}\int_0^\infty dre^{rQ}\int_{\mathbb{R}^d}\frac{d\hat V(p)}{(2\pi)^d}e^{ir\varepsilon^{\alpha-\beta}k\cdot p+iy\cdot p}\left[\lambda(x,k-\frac{\eta p}{2})-\lambda(x,k+\frac{\eta p}{2})\right].\tag{5.110}$$

For smooth test functions λ, we obtain that λ_1 is at most of order $O(\eta\varepsilon^{\alpha+\frac{\gamma-\delta}{2}-1})=O(\varepsilon^{\alpha-\beta+\frac{\gamma-\delta}{2}})$.

The next-order corrector is given by

$$\varepsilon^{-\alpha}Q\lambda_2+\varepsilon^{-\beta}k\cdot\nabla_y\lambda_2+\varepsilon^{-1+\frac{\gamma-\delta}{2}}\Big(\mathcal{K}[\hat V,y]\lambda_1-\mathbb{E}\{\mathcal{K}[\hat V,y]\lambda_1\}\Big)=0.\tag{5.111}$$

It admits an explicit expression, from which we deduce that it is at most of order $O(\eta^2\varepsilon^{2\alpha+\gamma-\delta-2}) = O(\varepsilon^{2\alpha-2\beta+\gamma-\delta})$. The expression for $\mathbb{E}\{\mathcal{K}[\hat{V},y]\lambda_1\}$ is given by

$$
\begin{aligned}
&\mathbb{E}\{\mathcal{K}[\hat{V},y]\lambda_1\}\\
&= -\frac{\varepsilon^{\alpha+\frac{\gamma-\delta}{2}}}{\varepsilon}\mathbb{E}\Big\{\int_{\mathbb{R}^d}\frac{d\hat{V}(p)}{(2\pi)^d}e^{ip\cdot y}\int_0^\infty dre^{rQ}\int_{\mathbb{R}^d}\frac{d\hat{V}(q)}{(2\pi)^d}e^{ir\varepsilon^{\alpha-\beta}(k-\frac{\eta p}{2})\cdot q}e^{iy\cdot q}\\
&\quad\times\Big(\lambda(x,k-\eta\frac{p+q}{2})-\lambda(x,k-\eta\frac{p-q}{2})\Big) \qquad + c.c.\Big\}\\
&= -\frac{\varepsilon^{\alpha+\frac{\gamma-\delta}{2}}}{\varepsilon}\int_{\mathbb{R}^d}\int_0^\infty\frac{\tilde{R}(r,p)}{(2\pi)^d}e^{ir\varepsilon^{\alpha-\beta}(k+\frac{\eta p}{2})\cdot p}\Big(\lambda(x,k)-\lambda(x,k+\eta p)\Big)dp \qquad + c.c.\\
&= -\frac{\varepsilon^{\alpha+\frac{\gamma-\delta}{2}}}{\varepsilon}\int_{\mathbb{R}^d}\frac{\hat{R}(\varepsilon^{\alpha-\beta}(k+\eta\frac{p}{2})\cdot p,p)}{(2\pi)^d}\Big(\lambda(x,k)-\lambda(x,k+\eta p)\Big)dp\\
&= -\frac{\varepsilon^{\alpha+\frac{\gamma-\delta}{2}}}{\varepsilon}\int_{\mathbb{R}^d}\frac{\hat{R}(\varepsilon^{\alpha-1}\frac{|k|^2-|\eta p|^2}{2},p-\frac{k}{\eta})}{(2\pi)^d}\Big(\lambda(x,k)-\lambda(x,\frac{p}{\eta})\Big)dp.
\end{aligned}
$$

The next-to-last expression is useful when $\beta < 1$, i.e., $\eta \ll 1$. The last expression is useful when $\beta = 1$, i.e., $\eta = 1$.

Up to a lower order term, the drift term in the martingale (5.107) is thus given by

$$\varepsilon^\delta\frac{\partial\lambda}{\partial t} + k\cdot\nabla\lambda + \varepsilon^{-\delta}\mathcal{L}_\varepsilon\lambda,$$

where

$$
\begin{aligned}
\mathcal{L}_\varepsilon\lambda &= \varepsilon^{\alpha+\gamma-2}\int_{\mathbb{R}^d}\frac{\hat{R}(\varepsilon^{\alpha-\beta}(k+\eta\frac{p}{2})\cdot p,p)}{(2\pi)^d}\Big(\lambda(x,k+\eta p)-\lambda(x,k)\Big)dp\\
&= \varepsilon^{\alpha+\gamma-2}\int_{\mathbb{R}^d}\frac{\hat{R}(\varepsilon^{\alpha-1}\frac{|k|^2-|\eta p|^2}{2},p-\frac{k}{\eta})}{(2\pi)^d}\Big(\lambda(x,\frac{p}{\eta})-\lambda(x,k)\Big)dp.
\end{aligned}
\tag{5.112}
$$

The remainder is given by

$$\zeta_\varepsilon = (\varepsilon^\delta\frac{\partial}{\partial t}+k\cdot\nabla_x)(\lambda_1^\varepsilon+\lambda_2^\varepsilon)+\varepsilon^{-1+\frac{\gamma-\delta}{2}}\Big(\mathcal{K}[\hat{V},y]\lambda_2\Big)_{|y=\varepsilon^{-\beta}x}.$$

The two main contributions are λ_1 and $\mathcal{K}\lambda_2$ and are of order

$$O(\varepsilon^{\alpha-\beta+\frac{\gamma-\delta}{2}}) \quad \text{and} \quad O(\varepsilon^{2\alpha-3\beta+\frac{3}{2}(\gamma-\delta)}).$$

Consider the case $\alpha < \beta$. There, we verify with $\gamma = \beta$ that the error is the minimum of

$$\alpha<\beta: \qquad \zeta_\varepsilon = O(\varepsilon^{\alpha-\frac{1}{2}(\beta+\delta)}\wedge\varepsilon^{2\alpha-\frac{3}{2}(\beta+\delta)}).$$

This requires that $\alpha > \frac{3}{4}(\beta+\delta)$ in order for the error to be $\ll 1$. When $\alpha \geq \beta$, so that $\delta = 0$ since no diffusion is possible then, we verify that the error, with $\gamma = 2\beta-\alpha$, is

$$\alpha\geq\beta: \qquad \zeta_\varepsilon = O(\varepsilon^{\frac{1}{2}\alpha}).$$

It remains to evaluate the limit of $\mathcal{L}_\varepsilon$ in the various cases of interest. Let us consider the case $\alpha > \beta$. Then by simple Taylor expansion, we find using $\gamma = 2\beta-\alpha$ that

$$
\begin{aligned}
\lim_{\varepsilon\to0}\mathcal{L}_\varepsilon\lambda &= \frac{1}{2}\Big(\int_{\mathbb{R}^d}\hat{R}(0,p)(p\cdot\nabla_k)^2\frac{dp}{(2\pi)^d}\Big)\lambda(x,k), \qquad \beta<1,\\
&= \int_{\mathbb{R}^d}\frac{\hat{R}(0,p-k)}{(2\pi)^d}\Big(\lambda(x,k)-\lambda(x,p)\Big)dp, \qquad \beta=1.
\end{aligned}
\tag{5.113}
$$

The first contribution explicitly uses the fact that $\hat{R}$ and its derivative with respect to the first variable are even with respect to all variables. Consider now the case $\alpha = \beta$. We verify that

$$\begin{aligned} \lim_{\varepsilon\to 0} \mathcal{L}_\varepsilon \lambda &= \frac{1}{2}\Big(\int_{\mathbb{R}^d} \hat{R}(k\cdot p, p)(p\cdot\nabla_k)^2 \frac{dp}{(2\pi)^d}\Big)\lambda(x,k), && \beta < 1, \\ &= \int_{\mathbb{R}^d} \frac{\hat{R}(\frac{1}{2}|k|^2 - \frac{1}{2}|p|^2, p-k)}{(2\pi)^d}\Big(\lambda(x,k) - \lambda(x,p)\Big)dp, && \beta = 1. \end{aligned} \tag{5.114}$$

We finally consider the more delicate case $\alpha < \beta$ with $\gamma = \beta$. The limits are given by

$$\begin{aligned} \lim_{\varepsilon\to 0} \mathcal{L}_\varepsilon \lambda &= \frac{1}{2}\Big(\int_{\mathbb{R}^d} \delta(k\cdot p)\hat{R}_0(p)(p\cdot\nabla_k)^2 \frac{dp}{(2\pi)^d}\Big)\lambda(x,k), && \beta < 1, \\ &= \int_{\mathbb{R}^d} \delta\big(\frac{|k|^2 - |p|^2}{2}\big)\frac{\hat{R}_0(p-k)}{(2\pi)^d}\Big(\lambda(x,k) - \lambda(x,p)\Big)dp, && \beta = 1. \end{aligned} \tag{5.115}$$

Here $\hat{R}_0(p) = \int_{\mathbb{R}} \hat{R}(v,p)dv$. The proof is given in Lemma 5.18 for the case $\beta = 1$. Moreover, the same lemma shows that, denoting by $\mathcal{L}$ the above limit, $\varepsilon^{-\delta}(\mathcal{L}_\varepsilon - \mathcal{L})\lambda$ converges to 0 in the L^2 sense for sufficiently small values of δ. More precisely, the latter is of order $\varepsilon^{(\beta-\alpha)/2-\delta}$ so that $2\delta < \beta - \alpha$ is necessary. This allows us to pass to the diffusive limit. The case $\beta < 1$ is handled similarly.

Approximations of operators.

Let $\hat{R}(v,p)$ be a smooth even powerspectrum and let us define

$$\hat{R}_0(p) = \int_{\mathbb{R}} \hat{R}(v,p)dv. \tag{5.116}$$

We consider the following difference

$$I(k) = \int \Big(\eta\hat{R}(\eta\frac{k^2-p^2}{2}, k-p) - \delta\Big(\frac{k^2-p^2}{2}\Big)\hat{R}_0(p-k)\Big)(\lambda(k)-\lambda(p))dp, \tag{5.117}$$

for λ a smooth test function $\lambda(k) \in \mathcal{D}(\mathbb{R}^d)$, which moreover we assume is supported away from $k = 0$. Then we have the following result

Lemma 5.18. *Let $\lambda(k) \in \mathcal{D}(\mathbb{R}^d)$ supported away from $k = 0$. We assume that $\hat{R}$ is a smooth (non-negative) function such that $v \to |v|\hat{R}(v,\cdot)$ is uniformly integrable and such that the integral of $\hat{R}(v)$ over $|v| \geq V$ is of order V^{-1}.*

Then we verify that

$$\|I\|_{L^p(\mathbb{R}^d_k)} \leq C_\lambda \eta^{1/p}, \qquad 1 \leq p \leq \infty. \tag{5.118}$$

Proof. Let us first consider

$$I_1(k) = \int_{\mathbb{R}^d} \Big(\eta\hat{R}(\eta\frac{|k|^2-|p|^2}{2}, k-p) - \delta\Big(\frac{|k|^2-|p|^2}{2}\Big)\hat{R}_0(p-k)\Big)(\lambda(|k|\hat{k}) - \lambda(|k|\hat{p}))dp, \tag{5.119}$$

where $\hat{k} = k/|k|$. Using the change of variables $0 < |p| \to v(|p|) = \frac{1}{2}(|k|^2 - |p|^2)$, and $p = |p|^{d-1}dpd\hat{p}$, and the definition $\mu(v) = |p|^{d-1}(v)dp/dv$, we calculate that

$$\begin{aligned} &\int_{\mathbb{R}^d} \eta\hat{R}(\eta\frac{|k|^2-|p|^2}{2}, k-p)(\lambda(|k|\hat{k}) - \lambda(|k|\hat{p}))dp \\ &= \int \eta\hat{R}(\eta v, p - p(v)\hat{p})(\lambda(|k|\hat{k}) - \lambda(|k|\hat{p}))\mu(v)dvd\hat{p}. \end{aligned}$$

Now for k away from 0, we verify that

$$\int_{-\infty}^{|k|^2/2} \eta\hat R(\eta v, p-p(v)d\hat p)\mu(v)dv = \hat R_0(k-|k|\hat p)\mu(0)+O(\eta^{-1}).$$

Using the integrability and boundedness of λ in k and the change of variables $|k|\delta(\frac{1}{2}(|p|^2-|k|^2)) = \delta(|k|-|p|)$, we deduce that

$$\|I_1\|_{L^p(\mathbb{R}^d_k)} \le \frac{C}{\eta}, \qquad 1\le p\le\infty.$$

It remains to address the term

$$I_2(k) = \int_{\mathbb{R}^d} \eta\hat R(\eta\frac{|k|^2-|p|^2}{2}, k-p)(\lambda(|k|\hat p)-\lambda(|p|\hat p))dp. \tag{5.120}$$

Note that $I(k) = I_1(k)+I_2(k)$. We verify as above that $\|I_2\|_{L^\infty(\mathbb{R}^d_k)} \le C$ and calculate that

$$\int_{\mathbb{R}^d} |I_2(k)|dk \le \int_{\mathbb{R}^{2d}} \hat R(v, k-|p|(\frac{v}{\eta})\hat p)\Big|\lambda(k)-\lambda(|p|(\frac{v}{\eta})\hat p)\Big|\mu(\frac{v}{\eta})dvd\hat p dk. \tag{5.121}$$

Now for $|v|\le\eta$, the regularity of λ and $|p|\to v(|p|)$ imply the existence of a function $\varphi(k)\in\mathcal{D}(\mathbb{R}^d)$, which we can also choose supported away from $k=0$ though with a larger support than λ, such that

$$\Big|\lambda(kp)-\lambda(p(\frac{v}{\eta})\hat p)\Big| \le C\frac{|v|}{\eta}\varphi(k), \qquad |v|\le\eta.$$

This shows that the contribution in (5.121) for $|v|\le\eta$ is bounded by $O(\eta^{-1})$ since $|v|\hat R(v,\cdot)$ is integrable. The other contribution $|v|\ge\eta$ yields a contribution bounded by $\int_{|v|\ge\eta}\hat R(v,\cdot)dv \le C\eta^{-1}$ by hypothesis on $\hat R$. This shows that $\|I_2\|_{L^1(\mathbb{R}^d)}\le C\eta^{-1}$. The result follows by interpolation for $1\le p\le\infty$. ∎

6. **Kinetic models for correlations.** So far, we have used the radiative transfer, Fokker Planck, and diffusion equations to model the energy density of the waves. The energy density may be seen as the correlation function of a random field with itself, as is apparent from the definition of the Wigner transform. More generally, we may consider the correlation function of two vector fields, corresponding, for instance, to different initial conditions, and propagating possibly in two different media. We will see later in this section applications for such correlation functions. Such correlations also satisfy kinetic models, and in this section we describe the corresponding kinetic models.

6.1. **Radiative transport equations for correlations.** We first consider the weak-coupling regime for the Schrödinger equation

$$i\varepsilon\frac{\partial\psi_{j,\varepsilon}}{\partial t} + \frac{\varepsilon^2}{2}\Delta\psi_{j,\varepsilon} - \sqrt{\varepsilon}V_j(\frac{x}{\varepsilon})\psi_{j,\varepsilon} = 0, \tag{6.1}$$

for $j=1,2$ corresponding to two possibly different media.

6.1.1. *Formal derivation with iterated Duhamel expansion.* As is standard in all kinetic models derived so far, in the limiting equation, the influence of the random fluctuations is characterized by the covariance function of the random potentials defined by

$$R_{mn}(x) = \mathbb{E}\left\{V_m(y)V_n(y+x)\right\}, \qquad m,n=1,2. \tag{6.2}$$

The power spectra $\hat{R}_{mn}(k)$ are then their Fourier transforms. The correlation function of the two wave fields after Fourier transform is defined as the following Wigner transform

$$W_\varepsilon(t,x,k) = \int_{\mathbb{R}^3} e^{ik\cdot y}\psi_{1\varepsilon}(t,x-\frac{\varepsilon y}{2})\psi_{2\varepsilon}^*(t,x+\frac{\varepsilon y}{2})\frac{dy}{(2\pi)^d}. \tag{6.3}$$

We assume here that $d=3$ to simplify. First we obtain that W_ε solves the following equation:

$$\frac{\partial W_\varepsilon}{\partial t} + k\cdot\nabla W_\varepsilon = \int_{\mathbb{R}^3} K_\varepsilon(x,k-p)W_\varepsilon(t,x,p)dp, \tag{6.4}$$

where the convolution kernel K_ε is given by

$$K_\varepsilon(x,p) = \frac{1}{i\pi^3\sqrt{\varepsilon}}\Big(\hat{V}_1(2p)e^{i2p\cdot x/\varepsilon} - \hat{V}_2(-2p)e^{-i2p\cdot x/\varepsilon}\Big). \tag{6.5}$$

Inverting the free transport operator $\partial_t + k\cdot\nabla$ we obtain that

$$W_\varepsilon(t,x,k) = W_\varepsilon(0,x-tk,k) + \int_0^t\int K_\varepsilon(x-sk,k-p)W_\varepsilon(t-s,x-sk,p)dpds.$$

After one more iteration we have

$$\begin{aligned} &W_\varepsilon(t,x,k)\\ &= W_\varepsilon(0,x-tk,k) + \int_0^t\int K_\varepsilon(x-sk,k-p)W_\varepsilon(0,x-sk-(t-s)p,p)dpds\\ &+\int_0^t\int K_\varepsilon(x-sk,k-p)\int_0^{t-s}\int K_\varepsilon(x-sk-up,p-q)\\ &\times W_\varepsilon(t-s-u,x-sk-up,q)dqdudpds. \end{aligned} \tag{6.6}$$

We now average the above equation with respect to the realizations of the random medium. We assume that $\mathbb{E}\left\{K_\varepsilon\otimes K_\varepsilon W_\varepsilon\right\} = \mathbb{E}\left\{K_\varepsilon\otimes K_\varepsilon\right\}\mathbb{E}\left\{W_\varepsilon\right\}$ and that W_ε is sufficiently smooth. Of course, this assumption is only formal and cannot be justified but it is exactly equivalent to the assumptions about the leading order term in the asymptotic expansions for the Wigner transform that we have considered previously in our formal arguments. It is known to yield the correct result in the weak coupling regime where rigorous derivations are also available. In the following section, we will present a rigorous result obtained in the setting of Schrödinger equations with time dependent potentials.

Using that

$$\mathbb{E}\left\{\hat{V}_m(p)\hat{V}_n(q)\right\} = (2\pi)^3\hat{R}_{mn}(p)\delta(p+q), \qquad m,n=1,2,$$

we deduce that

$$\mathbb{E}\left\{K_\varepsilon(y,k-p)K_\varepsilon(y-up,p-q)\right\}$$
$$= -\frac{1}{\pi^3\varepsilon}\Big(\hat{R}_{11}(2(k-p))e^{2i(k-p)\cdot pu/\varepsilon}\delta(k-q) - \hat{R}_{12}(2(k-p))e^{2i(k-p)\cdot pu/\varepsilon}\delta(k+q-2p)$$
$$-\hat{R}_{21}(-2(k-p))e^{-2i(k-p)\cdot pu/\varepsilon}\delta(k+q-2p) + \hat{R}_{22}(2(k-p))e^{-2i(k-p)\cdot pu/\varepsilon}\delta(k-p)\Big).$$

The power spectrum $\hat{R}_{mn}$ is a 2×2 positive definite matrix such that $\hat{R}_{mn}(-p) = \hat{R}_{nm}(p)$, $m,n=1,2$. After the changes of variables $2p-k\to p$ and $u\to\varepsilon u$ and replacing $W_\varepsilon(t-s-\varepsilon u, x-sk-\varepsilon up, q)$ by $W_\varepsilon(t-s,x-sk,q)$ we deduce that the ensemble average of the last term in (6.6) is approximated by

$$\begin{aligned}&\int_0^t\int\int_0^{(t-s)/\varepsilon}\Big(-e^{iu\frac{|k|^2-|p|^2}{2}}R_{11}(p-k) - e^{-iu\frac{|k|^2-|p|^2}{2}}R_{22}(p-k)\Big)\\ &\quad\times\mathbb{E}\left\{W_\varepsilon\right\}(t-s,x-sk,k) + \Big(e^{iu\frac{|k|^2-|p|^2}{2}} + e^{-iu\frac{|k|^2-|p|^2}{2}}\Big)R_{21}(p-k)\\ &\quad\times\mathbb{E}\left\{W_\varepsilon\right\}(t-s,x-sk,p)\frac{dudpds}{(2\pi)^3}.\end{aligned}\tag{6.7}$$

We now pass to the limit $\varepsilon\to 0$ and replace $\mathbb{E}\left\{W_\varepsilon\right\}$ by its limit W. We first observe that

$$\int_0^\infty e^{\pm iu\omega}du = \pi\delta(\omega) \pm \frac{i}{\omega}.$$

Thus in the limit $\varepsilon\to 0$ we obtain the equation

$$\begin{aligned}W(t,x,k) &= W(0,x-tk,k) + \int_0^t\Big(\int\hat{R}_{21}(p-k)W(t-s,x-sk,p)\\ &\quad\times\delta(\frac{|k|^2-|p|^2}{2})\frac{dp}{(2\pi)^2} - (\Sigma(k)+i\Pi(k))W(t-s,x-sk,k)\Big)ds,\end{aligned}$$

where the total absorption and phase modulation terms are given by

$$\begin{aligned}\Sigma(k) &= \int\frac{\hat{R}_{11}(p-k)+\hat{R}_{22}(p-k)}{2}\delta(\frac{|k|^2-|p|^2}{2})\frac{dp}{(2\pi)^2}\\ \Pi(k) &= ⨍\Big(\hat{R}_{11}(p-k)-\hat{R}_{22}(p-k)\Big)\frac{2}{|k|^2-|p|^2}\frac{dp}{(2\pi)^3}.\end{aligned}\tag{6.8}$$

The latter integral defining Π has to be understood in the principal value sense. We assume that the power spectrum $\hat{R}_{mn}$ is such that the above integrals exist. This is the integral formulation of the following radiative transfer equation

$$\frac{\partial W}{\partial t} + k\cdot\nabla W + (\Sigma(k)+i\Pi(k))W = \int\hat{R}_{21}(p-k)W(t,x,p)\delta(\frac{|k|^2-|p|^2}{2})\frac{dp}{(2\pi)^2}.\tag{6.9}$$

The initial conditions for W are given by the Wigner transform of the two fields $\psi_{j,\varepsilon}$. These initial conditions need not be equal.

Such derivations may be generalized to other equations; for an explicit expression for acoustic wave equations, we refer the reader to [21].

6.1.2. *Rigorous results for time dependent potentials.* We consider in this section the general problem of the correlation of solutions of the linear paraxial Schrödinger equations in two different albeit correlated random media. We let $\psi_\varepsilon(t,x)$ and $\phi_\varepsilon(t,x)$ be the solutions of the family of Cauchy problems

$$
\begin{aligned}
&i\varepsilon\frac{\partial\psi_\varepsilon}{\partial t}+\frac{\varepsilon^2}{2}\Delta\psi_\varepsilon-\kappa^2\sqrt{\varepsilon}V_1\left(\frac{t}{\varepsilon},\frac{x}{\varepsilon}\right)\psi_\varepsilon=0 \\
&\psi_\varepsilon(0,x)=\psi_\varepsilon^0(x;\zeta)
\end{aligned}
\tag{6.10}
$$

and

$$
\begin{aligned}
&i\varepsilon\frac{\partial\phi_\varepsilon}{\partial t}+\frac{\varepsilon^2}{2}\Delta\phi_\varepsilon-\sqrt{\varepsilon}V_2\left(\frac{t}{\varepsilon},\frac{x}{\varepsilon}\right)\phi_\varepsilon=0, \\
&\phi_\varepsilon(0,x)=\phi_\varepsilon^0(x;\zeta)
\end{aligned}
\tag{6.11}
$$

with two different random potentials V_1 and V_2. The initial data depend on an additional random variable ζ defined over a state space S with a probability measure $d\mu(\zeta)$, as we are going, once again, to consider mixtures of states. The cross Wigner transform is defined by

$$W_\varepsilon(t,x,k)=\int_{\mathbb{R}^d\times S}e^{ik\cdot y}\psi_\varepsilon\left(t,x-\frac{\varepsilon y}{2};\zeta\right)\bar{\phi}_\varepsilon\left(t,x+\frac{\varepsilon y}{2};\zeta\right)\frac{dy}{(2\pi)^d}d\mu(\zeta).$$

The evolution equation for the Wigner transform is

$$
\begin{aligned}
&\frac{\partial W_\varepsilon}{\partial t}+k\cdot\nabla_x W_\varepsilon \\
&=\tfrac{1}{i\sqrt{\varepsilon}}\int_{\mathbb{R}^d}e^{ip\cdot x/\varepsilon}\left[\tilde{V}_1\left(\tfrac{t}{\varepsilon},p\right)W_\varepsilon\left(k-\tfrac{p}{2}\right)-\tilde{V}_2\left(\tfrac{t}{\varepsilon},p\right)W_\varepsilon\left(k+\tfrac{p}{2}\right)\right]\tfrac{dp}{(2\pi)^d}.
\end{aligned}
\tag{6.12}
$$

Here $\tilde{V}(t,p)$ is the partial Fourier transform of $V(t,x)$ in x only. We will make the same assumptions as in Section 5: we assume that the initial data $W_\varepsilon(0,x,k)$ converges strongly in $L^2(\mathbb{R}^d\times\mathbb{R}^d)$ to a limit $W_0(x,k)$, which is possible due to the introduction of the mixture of states . We also make all the assumptions on the random processes $V_{1,2}(z)$ that we assumed about $V(t,x)$ in Section 5, so we do not repeat them here.

The main result of this section is that under the above assumptions, the following theorem holds. Let us define the operator

$$
\begin{aligned}
\mathcal{L}f(x,k)=&\int_{\mathbb{R}^d}\left[\hat{R}_{12}(\tfrac{p^2-k^2}{2},p-k)W_0(p)-\tfrac{\hat{R}_{11}(\frac{p^2-k^2}{2},p-k)+\hat{R}_{22}(\frac{p^2-k^2}{2},p-k)}{2}W_0(k)\right] \\
&\times\tfrac{dp}{(2\pi)^d}-i\Pi(k)W_0(k)
\end{aligned}
$$

with

$$
\begin{aligned}
\Pi(k)&=\frac{1}{i}\int_{\mathbb{R}}dr\int_{\mathbb{R}^d}\tfrac{dp}{(2\pi)^d}\tfrac{\tilde{R}_{22}(r,p)-\tilde{R}_{11}(r,p)}{2}\exp\{ir(k-p/2)\cdot p\}\mathrm{sgn}(r) \\
&=\int_{\mathbb{R}^d}\text{p.v.}\int_{\mathbb{R}}\frac{\hat{R}_{22}(\omega,k-p)-\hat{R}_{11}(\omega,k-p)}{\omega-\frac{|p|^2-|k|^2}{2}}\frac{d\omega dp}{(2\pi)^{d+1}}.
\end{aligned}
$$

Here, $\tilde{R}(r,p)$ is the partial Fourier transform of R in x only. We denote the standard inner product on $L^2(\mathbb{R}^{2d})$ by $\langle f,g\rangle=\int_{\mathbb{R}^{2d}}f(x,k)\bar{g}(x,k)dxdk$. Then we have the following result.

Theorem 6.1. *Under the above assumptions, the Wigner distribution W_ε converges in probability and weakly in $L^2(\mathbb{R}^{2d})$ to the solution $\overline{W}$ of the transport equation*

$$\frac{\partial \overline{W}}{\partial t} + k \cdot \nabla_x \overline{W} = \mathcal{L}\overline{W}. \tag{6.13}$$

More precisely, for any test function $\lambda \in L^2(\mathbb{R}^{2d})$ the process $\langle W_\varepsilon(t), \lambda\rangle$ converges to $\langle \overline{W}(z), \lambda\rangle$ in probability as $\varepsilon \to 0$, uniformly on finite intervals $0 \le t \le T$.

The proof of the above theorem is very similar to that of Theorem 5.10. We refer the reader to [20] for the details.

6.1.3. *Correlations and Itô-Schrödinger models.* The above results may also be derived for the Itô-Schrödinger model (white noise limit) for wave propagation. Since we have access to exact equations for moments of the wave functions, the derivation is significantly simplified.

Let ψ_1 and ψ_2 satisfy

$$d\psi_m(t,x) = \frac{1}{2}\Big(i\varepsilon\Delta_x - K_m(\mathbf{0})\Big)\psi_m(t,x)dt + i\psi_m(t,x)B_m(dt,\frac{x}{\varepsilon}), \qquad m = 1,2. \tag{6.14}$$

The Wiener measures $B_{1,2}$ are described by different statistics $K_{1,2}$ for the forward propagation (index 1) and the backward propagation (index 2). The *cross-correlation* of the two media, is defined by

$$\mathbb{E}\{B_m(x,t)B_n(y,t')\} = K_{mn}(x-y)t \wedge t', \qquad 1 \le m,n \le 2. \tag{6.15}$$

We define the second moment $m_2(x,y)$ as

$$m_2(t,x,y,\kappa) = \mathbb{E}\{\psi_1(t,x+\frac{\varepsilon y}{2},\kappa)\psi_2^*(t,x-\frac{\varepsilon y}{2},\kappa)\}. \tag{6.16}$$

By an application of the Itô calculus [66] we obtain that

$$d(\psi_1(t,x)\psi_2^*(t,y)) = \psi_1(t,x)d\psi_2^*(t,y) + d\psi_1(t,x)\psi_2^*(t,y) + d\psi_1(t,x)d\psi_2^*(t,y).$$

We insert (6.14) into the above formula and taking mathematical expectation, obtain after some algebra [2] an equation for m_2:

$$\frac{\partial m_2}{\partial t} = \nabla_x \cdot \nabla_y m_2 - \Big(\frac{K_{11}(\mathbf{0}) + K_{22}(\mathbf{0})}{2} - K_{12}(y)\Big)m_2. \tag{6.17}$$

Now, defining the Wigner transform of the two fields as

$$W_{12}(t,x,k) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{ik\cdot x}\psi_1(t,x-\frac{\varepsilon y}{2})\psi_2^*(t,x+\frac{\varepsilon y}{2})dy, \tag{6.18}$$

we find that

$$m_2(t,x,y) = \int_{\mathbb{R}^d} e^{ik\cdot y}\mathbb{E}\{W_{12}\}(t,x,k,\kappa)dk. \tag{6.19}$$

Therefore, $\mathbb{E}\{W_{12}\}$ solves the following equation:

$$\frac{\partial W}{\partial t} + k\cdot\nabla_x W + \frac{K_{11}(\mathbf{0}) + K_{22}(\mathbf{0})}{2}W = \int_{\mathbb{R}^d} \hat{K}_{12}(p-k)W(p)dp. \tag{6.20}$$

The initial condition for W is simply given by the Wigner transform of the two wave fields (in the limit $\varepsilon \to 0$). This formal derivation may be turned into a rigorous derivation as in the case of energy densities: see [20].

6.2. **Fokker-Planck equation for correlations.** Equations for correlation functions may also be obtained when the correlation length is much larger than the wavelength. In such a regime, we recall that the energy density solves a Fokker-Planck equation in the limit of vanishing correlation length.

Let us assume that propagation occurs in media with some mismatch. The function ψ_ε satisfies

$$i\varepsilon\frac{\partial\psi_\varepsilon}{\partial t}+\frac{\varepsilon^2}{2}\Delta\psi_\varepsilon-V_\delta(x)\psi_\varepsilon=0 \tag{6.21}$$

and the function ϕ_ε satisfies

$$i\varepsilon\frac{\partial\phi_\varepsilon}{\partial t}+\frac{\varepsilon^2}{2}\Delta\phi_\varepsilon-[V_\delta(x)+\varepsilon S_\delta(x)]\phi_\varepsilon=0. \tag{6.22}$$

Both of the functions ϕ_ε and ψ_ε satisfy initially

$$\psi_\varepsilon(0,x)=\phi_\varepsilon(0,x)=\phi_0^\varepsilon(x) \tag{6.23}$$

to simplify, although different initial conditions can also be considered as in the preceding section. The family ϕ_ε^0 is ε-oscillatory and compact at infinity. The random potentials V_δ and S_δ vary on a scale δ that is much larger than the wave length ε of the initial data but is much smaller than the overall propagation distance that is of the order $O(1)$: $\varepsilon\ll\delta\ll 1$. To keep a non-trivial correlation of ψ_ε and ϕ_ε the mismatch of the potentials has to be weak – hence the coefficient ε in front of S_δ. We will see that in order to produce an order one contribution we will have eventually to take $S_\delta(x)=\delta^{-1/2}S(x/\delta)$ making the overall strength of the mismatch be of the order $O(\varepsilon/\sqrt{\delta})$.

In order to study the correlation of ψ_ε and ϕ_ε we introduce the cross Wigner transform as

$$W_\varepsilon(t,x,k)=\int e^{ik\cdot y}\psi_\varepsilon\left(t,x-\frac{\varepsilon y}{2}\right)\bar\phi_\varepsilon\left(t,x+\frac{\varepsilon y}{2}\right)\frac{dy}{(2\pi)^d}. \tag{6.24}$$

The distribution $W_\varepsilon(t,x,k)$ does not have to be real if $\phi_\varepsilon\neq\psi_\varepsilon$. Its phase measures the decoherence of the functions ϕ_ε and ψ_ε. In order to obtain an equation for W_ε we differentiate the Wigner transform with respect to time:

$$\frac{\partial W_\varepsilon}{\partial t}+k\cdot\nabla_x W_\varepsilon$$
$$=\frac{1}{i\varepsilon}\int e^{ik\cdot y}\left[V_\delta\left(x-\tfrac{\varepsilon y}{2}\right)-V_\delta\left(x+\tfrac{\varepsilon y}{2}\right)-\varepsilon S_\delta\left(x+\tfrac{\varepsilon y}{2}\right)\right]\psi_\varepsilon\left(x-\tfrac{\varepsilon y}{2}\right)\bar\phi_\varepsilon\left(x+\tfrac{\varepsilon y}{2}\right)\tfrac{dy}{(2\pi)^d}.$$

Passing to the limit $\varepsilon\to 0$ we obtain an equation for the distribution $W_\delta(t,x,k)$, the weak limit of W_ε as $\varepsilon\to 0$:

$$\frac{\partial W_\delta}{\partial t}+k\cdot\nabla_x W_\delta-\nabla V_\delta(x)\cdot\nabla_k W_\delta=iS_\delta(x)W_\delta. \tag{6.25}$$

In order to obtain a non-trivial limit of $W_\delta(t,x,k)$ as the correlation length $\delta\to 0$ we choose the random potential $V_\delta(x)$ and the mismatch $S_\delta(x)$ to be of the form

$$V_\delta(x)=\sqrt{\delta}V\left(\frac{x}{\delta}\right),\ S_\delta(x)=\frac{1}{\sqrt{\delta}}S\left(\frac{x}{\delta}\right).$$

Then (6.25) becomes

$$\frac{\partial W_\delta}{\partial t}+k\cdot\nabla_x W_\delta-\frac{1}{\sqrt{\delta}}\nabla V\left(\frac{x}{\delta}\right)\cdot\nabla_k W_\delta=\frac{i}{\sqrt{\delta}}S\left(\frac{x}{\delta}\right)W_\delta \tag{6.26}$$
$$W(0,x,k)=W_0(x,k).$$

The initial data $W_0(x,k)$ is simply the limit Wigner measure of the family ϕ_ε^0. Equation (6.26) is the starting point of our analysis.

A formal derivation of the Fokker-Planck limit may be obtained as follows. We introduce a multiple scales expansion

$$W_\delta = W(t,x,k) + \sqrt{\delta} W_1(t,x,y,k) + \delta W_2(t,x,y,k) + \ldots, \ y = x/\delta$$

and insert it into (6.26). As usual we make an additional assumption that the leading order term $W(t,x,k)$ is deterministic and does not depend on the fast scale variable y. In the leading order we obtain

$$k \cdot \nabla_y W_1 + \theta W_1 = \nabla V(y) \cdot \nabla_k W + iS(y) W.$$

Here $\theta > 0$ is an auxiliary regularizing parameter that we will send to zero at the end. Define the correctors χ_j and η as mean-zero solutions of

$$k \cdot \nabla_y \chi_j + \theta \chi_j = \frac{\partial V}{\partial y_j}$$
$$k \cdot \nabla_y \eta + \theta \eta = S(y).$$

They are given explicitly by

$$\chi_j(y,k) = \int_0^\infty \frac{\partial V(y-sk)}{\partial y_j} e^{-\theta s} ds \tag{6.27}$$

and

$$\eta(y,k) = \int_0^\infty e^{-\theta s} S(y-sk) ds. \tag{6.28}$$

The function W_1 is given in terms of the correctors as

$$W_1(t,x,y,k) = \sum_{j=1}^d \chi_j(y,k) \frac{\partial W(t,x,k)}{\partial k_j} + i\eta(y,k) W(t,x,k).$$

The equation for W_2 is

$$\frac{\partial W}{\partial t} + k \cdot \nabla_x W + k \cdot \nabla_y W_2 = \nabla V(y) \cdot \nabla_k W_1 + iS(y) W_1.$$

Averaging under the assumption that $\mathbb{E}\{k \cdot \nabla_y W_2\} = 0$ we obtain the following closed equation for the leading order term W:

$$\frac{\partial W}{\partial t} + k \cdot \nabla_x W = \mathbb{E}\left\{\nabla V(y) \cdot \nabla_k W_1 + iS(y) W_1\right\} = J_I + J_{II}. \tag{6.29}$$

The two terms on the right side are computed using the explicit expressions (6.27) and (6.28) for the correctors. The first term may be split as

$$J_I = \mathbb{E}\left\{\nabla V(y) \cdot \nabla_k W_1\right\} = J_I^1 + J_I^2$$

with

$$\begin{aligned} J_I^1 &= \mathbb{E}\left\{ \frac{\partial V}{\partial y_j}(y) \frac{\partial}{\partial k_j} \left[\chi_m(y,k) \frac{\partial W(t,x,k)}{\partial k_m} \right] \right\} \\ &= \frac{\partial}{\partial k_j} \left[\mathbb{E}\left\{ \frac{\partial V}{\partial y_j}(y) \int_0^\infty \frac{\partial V(y-sk)}{\partial y_m} e^{-\theta s} ds \right\} \frac{\partial W(t,x,k)}{\partial k_m} \right] \\ &= \frac{\partial}{\partial k_j} \left(D_{jm}(k) \frac{\partial W(t,x,k)}{\partial k_m} \right) \end{aligned}$$

where the diffusion matrix D_{jm} is given by

$$D_{mn}(k) = -\frac{1}{2}\int_{-\infty}^{\infty}\frac{\partial^2 R^{VV}(ks)}{\partial x_n \partial x_m}ds \\ = -\frac{1}{2|k|}\int_{-\infty}^{\infty}\frac{\partial^2 R^{VV}(s\hat{k})}{\partial x_n \partial x_m}ds, \quad m,n = 1,\dots,d,\ \hat{k} = k/|k|. \tag{6.30}$$

The term J_I^2 is

$$J_I^2 = \mathbb{E}\left\{\frac{\partial V}{\partial y_j}(y)\frac{\partial}{\partial k_j}\left[i\eta(y,k)W(t,x,k)\right]\right\} \\ = i\frac{\partial}{\partial k_j}\left[\mathbb{E}\left\{\frac{\partial V}{\partial y_j}(y)\int_0^\infty S(y-sk)e^{-\theta s}ds\right\}W(t,x,k)\right] \\ = i\frac{\partial}{\partial k_j}\left(E_j'(k)W(t,x,k)\right)$$

with the drift

$$E_j'(k) = \int_0^\infty \frac{\partial R^{SV}(sk)}{\partial x_j}ds.$$

Now we look at the second term in the right side of (6.29)

$$J_{II} = \mathbb{E}\left\{iS(y)W_1\right\} = J_{II}^1 + J_{II}^2 \tag{6.31}$$

with

$$J_{II}^1 = \mathbb{E}\left\{iS(y)\chi_m(y,k)\frac{\partial W(t,x,k)}{\partial k_m}\right\} \\ = i\mathbb{E}\left\{S(y)\int_0^\infty \frac{\partial V(y-sk)}{\partial y_m}e^{-\theta s}ds\right\}\frac{\partial W(t,x,k)}{\partial k_m} \\ = iE_m''(k)\frac{\partial W(t,x,k)}{\partial k_m}$$

with

$$E_m'' = -\int_0^\infty \frac{\partial R^{VS}(sk)}{\partial x_j}ds = \int_0^\infty \frac{\partial R^{SV}(-sk)}{\partial x_j}ds = \int_{-\infty}^0 \frac{\partial R^{SV}(sk)}{\partial x_j}ds.$$

Note that

$$J_I^2 + J_{II}^1 = i\frac{\partial}{\partial k_j}\left(E_j'(k)W(t,x,k)\right) + iF_m(k)\frac{\partial W(t,x,k)}{\partial k_m} \\ = i(E_j' + E_j'')\frac{\partial W(t,x,k)}{\partial k_j} + i(\nabla_k \cdot E')W(t,x,k) \\ = E_j\frac{\partial W(t,x,k)}{\partial k_j} + FW(t,x,k)$$

with

$$E_j = E_j' + E_j'' = \int_{-\infty}^\infty \frac{\partial R^{SV}(sk)}{\partial x_j}ds \tag{6.32}$$

and

$$F = \nabla_k \cdot E' = \int_0^\infty s\Delta R^{SV}(sk)ds. \tag{6.33}$$

The last term in (6.31) is

$$J_{II}^2 = \mathbb{E}\left\{iS(z)\eta(z,k)W(t,x,k)\right\} = -\kappa(k)W(t,x,k)$$

with the absorption coefficient

$$\kappa(k) = \int_0^\infty R^{SS}(sk)ds. \tag{6.34}$$

Putting together all the terms above we get the equation for W:

$$\frac{\partial W}{\partial t} + k\cdot\nabla_x W = iE(k)\cdot\nabla_k W + iF(k)W + \frac{\partial}{\partial k_m}\left(D_{mn}(k)\frac{\partial W}{\partial k_n}\right) - \kappa(k)W. \tag{6.35}$$

If S and V are independent then $F = E = 0$ and this simplifies to

$$\frac{\partial W}{\partial t} + k\cdot\nabla_x W = \frac{\partial}{\partial k_m}\left(D_{mn}(k)\frac{\partial W}{\partial k_n}\right) - \kappa(k)W. \tag{6.36}$$

In its simplest version, we therefore see that a mismatch in the random fluctuations generates absorption $\kappa(k)$ for the limiting correlation function W. This is consistent with the picture obtained for the radiative transfer equation in the preceding section.

6.3. **More general models for correlations.** The preceding kinetic equations for correlations were derived for a Schödinger wave model. Kinetic equations may be obtained for more general dispersive and non-dispersive wave models. Such models were derived for scalar wave field models in e.g. [3]. We will present a brief summary of the kinetic models derived in that reference, to which we refer the reader for the details.

Consider an equation of the form

$$R(\varepsilon D_t)p_\varepsilon + \mathcal{H}_\varepsilon p_\varepsilon = 0, \qquad \mathcal{H}_\varepsilon = b_\varepsilon(x)\beta(\varepsilon\mathbf{D}_x)d_\varepsilon(x)\gamma(\varepsilon\mathbf{D}_x). \tag{6.37}$$

The operators $R(\varepsilon D_t)$, $\beta(\varepsilon\mathbf{D}_x)$, and $\gamma(\varepsilon\mathbf{D}_x)$ are pseudo-differential operators with constant coefficients and with symbols defined by $R(i\varepsilon\omega)$, $\beta(i\varepsilon k)$, and $\gamma(\varepsilon ik)$, respectively. This means that $R(\varepsilon D_t) = \mathcal{F}^{-1}R(\varepsilon i\omega)\mathcal{F}$, where $\mathcal{F}$ is the Fourier transform, with similar expressions for $\beta(\varepsilon\mathbf{D}_x)$ and $\gamma(\varepsilon\mathbf{D}_x)$. We assume that $R(i\omega)$ is real-valued and that $\beta(\varepsilon\mathbf{D}_x)$ is the formal adjoint operator to $\gamma(\varepsilon\mathbf{D}_x)$.

For instance, the scalar wave equation

$$\frac{1}{c^2(x)}\partial_t^2 p = \Delta p,$$

corresponds to the case $R(i\omega) = -\omega^2$ and $\gamma(ik) = ik$ with $b(x) = c^2$ and $d(x) = 1$.

The advantage of the formulation (6.37) is that other (dispersive) equations such as Klein Gordon with $R(i\omega) = -\omega^2 + \alpha^2$ or Schrödinger with $R(i\omega) = -\omega$. Discretization effects may also be accounted for, for instance by choosing $R(i\omega) = i\frac{\sin\omega\Delta}{\Delta}$, where Δ is a time discretization step. The derivation of kinetic models is thus possible for a large class of wave models. Here, as in [3], we restrict ourselves to scalar models although the generalization to vectorial wave models is possible. The above model (6.37) is not sufficiently general to account for spatial operators of the form $-\Delta + V_\varepsilon$ as for the Schrödinger equation. We thus consider scalar equations of the form

$$\begin{aligned} &R(\varepsilon D_t)p_\varepsilon^\varphi + \mathcal{H}_\varepsilon^\varphi p_\varepsilon^\varphi = 0, \quad \mathcal{H}_\varepsilon^\varphi = \sum_{\mu=1}^M \mathcal{H}_{\mu\varepsilon}^\varphi, \quad 1\le\varphi\le 2, \\ &\mathcal{H}_{\mu\varepsilon}^\varphi = b_{\mu\varepsilon}^\varphi(x)\beta_\mu(\varepsilon\mathbf{D}_x)d_{\mu\varepsilon}^\varphi(x)\gamma_\mu(\varepsilon\mathbf{D}_x), \quad 1\le\varphi\le 2,\ 1\le\mu\le M. \end{aligned} \tag{6.38}$$

Here $M \geq 1$ is a fixed positive integer. We define the operators

$$H(x,k) = \sum_{\mu=1}^{N} H_\mu(x,k),$$

$$H_\mu(x,k) = b_{\mu 0}(x) d_{\mu 0}(x) \beta_\mu(ik)\gamma_\mu(ik),$$

for $1 \leq \mu \leq M$. We assume that the random fluctuations and the power spectra are given by

$$\begin{aligned}
&b^\varphi_{\mu\varepsilon}(x) = b_{\mu 0}(x) + \sqrt{\varepsilon} b^\varphi_{\mu 1}(x, \frac{x}{\varepsilon}), \qquad d^\varphi_{\mu\varepsilon}(x) = d_{\mu 0}(x) + \sqrt{\varepsilon} d^\varphi_{\mu 1}(x, \frac{x}{\varepsilon}),\\
&(2\pi)^d x_{\mu 0} y_{\mu 0} \hat{R}^{\varphi\psi}_{\mu\nu xy}(x,p)\delta(p+q) = \mathbb{E}\left\{\hat{x}^\varphi_{\mu 1}(x,p)\hat{y}^\psi_{\nu 1}(x,q)\right\}\\
&(x,y) \in \{(b,b),(b,d),(d,d)\}, \quad 1 \leq \varphi,\psi \leq 2, \quad 1 \leq \mu,\nu \leq M.
\end{aligned} \tag{6.39}$$

Let us now define the Wigner transform of the two wave fields:

$$W_\varepsilon(t,\omega,x,k) = W[p^1_\varepsilon, p^2_\varepsilon](t,\omega,x,k).$$

We want to find a kinetic model to represent the evolution of $W_\varepsilon(t,\omega,x,k)$ in the limit $\varepsilon \to 0$. Since the random fluctuations have a negligible effect on the dispersion relation, we find that

$$\Big(R(i\omega) + H(x,k)\Big) W(t,\omega,x,k) = 0. \tag{6.40}$$

This implies that $W(t,\omega,x,k)$ is a distribution supported on the manifold given by

$$R(i\omega) + H(x,k) = 0, \tag{6.41}$$

which we assume admits $\omega_n(x,k)$ as distinct solutions. We assume that all solutions $\omega_n(x,k)$ are real-valued. When $R(i\omega) = S(\omega^2)$, they come in pairs $\omega_{-n} = \omega_n$. The (even) number of modes indexed by n may be finite or infinite but is assumed to be independent of (x,k). This generalizes the case of the wave equation where $n = \pm 1$. This allows us to decompose W_0 as

$$W_0(t,\omega,x,k) = b_0(x) \sum_n a_n(t,x,k)\delta(\omega - \omega_n(x,k)). \tag{6.42}$$

In other words, the correlation function peaks at those values of ω given by the dispersion relation. It remains to find equations for the various amplitudes $a_n(t,x,k)$. Generalizations of the techniques developed earlier in this paper allow us to obtain that the modes $a_m(t,x,k)$ then satisfy the following equation

$$\begin{aligned}
&\frac{\partial a_m}{\partial t} + \{\omega_m(x,k), a_m\} + \big(\Sigma_m(x,k) + i\Pi_m(x,k)\big) a_m\\
&= \int_{\mathbb{R}^d} \sigma_m(x,k,q) \sum_n a_n(q) \delta\big(\omega_n(x,q) - \omega_m(x,k)\big) dq.
\end{aligned} \tag{6.43}$$

Here, the Poisson bracket is defined as

$$\{P, W\}(x,k) = (\nabla_x P \cdot \nabla_k W - \nabla_x W \cdot \nabla_k P)(x,k). \tag{6.44}$$

The coefficients appearing in the above kinetic equation have the form

$$\begin{aligned}
\Sigma_m(x,k) &= \int_{\mathbb{R}^d} \sum_{\mu,\nu=1}^{M} \frac{B^{11}_{\mu\nu}(x,k,q)+B^{22}_{\mu\nu}(x,k,q)}{2|R'(i\omega_m(x,k))|^2} \sum_n \\
&\quad \times \delta\big(\omega_n(x,q)-\omega_m(x,k)\big)\frac{dq}{(2\pi)^{d-1}},\\
\Pi_m(x,k) &= \frac{1}{R'(i\omega_m(x,k))}\text{p.v.}\int_{\mathbb{R}^d}\sum_{\mu,\nu=1}^{M}\frac{B^{11}_{\mu\nu}(x,k,q)-B^{22}_{\mu\nu}(x,k,q)}{H(x,k)-H(x,q)}\frac{dq}{(2\pi)^d},\\
\sigma_m(x,k,q) &= \sum_{\mu,\nu=1}^{M}\frac{B^{12}_{\mu\nu}(x,k,q)}{|R'(i\omega_m(x,k))|^2}\frac{1}{(2\pi)^{d-1}},\\
\alpha^{\varphi}_{\mu}(x,p,k,q) &= d_{\mu 0}(x)\hat{b}^{\varphi}_{\mu 1}(x,p)\beta_\mu(iq)\gamma_\mu(iq)+b_{\mu 0}(x)\hat{d}^{\varphi}_{\mu 1}(x,p)\beta_\mu(ik)\gamma_\mu(iq),\\
B^{\varphi\psi}_{\mu\nu}(x,k,q)\delta(0) &= (2\pi)^{-d}\mathbb{E}\left\{\alpha^{\varphi}_{\mu}(x,k-q,k,q)\alpha^{\psi}_{\nu}(x,q-k,q,k)\right\}.
\end{aligned} \tag{6.45}$$

Such models may be applied for scalar wave equations, Klein-Gordon equations, Schrödinger equations as we mentioned earlier, but also for some systems of Maxwell's equations; see [3] for the details.

The above kinetic equations were obtained for scalar wave models. There is no fundamental difficulty to extend the algebra to the systems analyzed in [71] although such calculations have not been carried out to date, except for the system of acoustic waves that will be presented below.

7. Application to time reversal.

7.1. Time reversal modeling.

Propagation of acoustic waves is described by the following linear hyperbolic system

$$A(x)\frac{\partial \mathbf{u}}{\partial t}+D^j\frac{\partial \mathbf{u}}{\partial x^j}=0,\ \ x\in\mathbb{R}^3 \tag{7.1}$$

with the vector $\mathbf{u}=(v,p)\in\mathbb{C}^4$. The matrix $A=\text{Diag}(\rho,\rho,\rho,\kappa)$ is positive definite. We recall that the 4×4 matrices D^j, $j=1,2,3$, are symmetric and given by $D^j_{mn}=\delta_{m4}\delta_{nj}+\delta_{n4}\delta_{mj}$. We use the Einstein convention of summation over repeated indices. We restrict ourselves to the case of dimension $d=3$ for concreteness although spatial dimension $d\geq 2$ may be treated similarly.

The time reversal experiment considered here consists of two steps. First, the direct problem

$$\begin{aligned}
A(x)\frac{\partial \mathbf{u}}{\partial t}+D^j\frac{\partial \mathbf{u}}{\partial x^j}&=0,\ \ 0\leq t\leq T\\
\mathbf{u}(0,x)&=\mathbf{S}(x)
\end{aligned} \tag{7.2}$$

with a localized source $\mathbf{S}$ centered at a point x_0 is solved until time $t=T$ to yield $u(T^-,x)$. 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 $\Gamma(x)$, which in the ideal case is given by $\Gamma=\Gamma_0=\text{Diag}(-1,-1,-1,1)$. After processing, we have

$$u(T^+,x)=\Gamma(f*(\chi u))(T^-,x)\chi(x). \tag{7.3}$$

The processed signal then propagates for the same time duration T:

$$\begin{aligned} A(x)\frac{\partial \mathbf{u}}{\partial t} + D^j\frac{\partial \mathbf{u}}{\partial x^j} &= 0,\ T \le t \le 2T \\ u(T^+, x) &= \Gamma(f * (\chi u))(T^-, x)\chi(x). \end{aligned} \tag{7.4}$$

To compare the signal $u(2T, x)$ with the initial pulse $\mathbf{S}$, we need to reverse the acoustic velocity once again and define the *back-propagated signal*

$$u^B(x) = \Gamma_0 u(2T, x).$$

The main question is whether the back-propagated signal $u^B(x) = \Gamma_0 u(2T, x)$ 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 $u^B(x) = \mathbf{S}(x)$, which corresponds to exact refocusing.

When only partial measurements are available we shall see that $\mathbf{u}(2T, x)$ is closer to $\Gamma\mathbf{S}(x)$ when propagation occurs in a heterogeneous medium than in a homogeneous medium. This is one of the main striking results of the theory of time reversal. Kinetic models offer a very precise quantitative description of this behavior.

7.2. **Kinetic model for the refocusing signal.** Let $G(T, x; z)$ be the Green matrix solution of

$$\begin{aligned} A(x)\frac{\partial G(t, x; y)}{\partial t} + D^j\frac{\partial G(t, x; y)}{\partial x^j} &= 0,\ 0 \le t \le T \\ G(0, x; y) &= I\delta(x - y). \end{aligned} \tag{7.5}$$

The back-propagated signal may thus be written as

$$u^B(x) = \Gamma_0 u(2T, x) = \int_{\mathbb{R}^9} \Gamma_0 G(T, x; y)\Gamma(y')G(T, y'; z)\chi(y)\chi(y')f(y-y')\mathbf{S}(z)dydy'dz. \tag{7.6}$$

We observe that the signal involves the product of two Green's functions, which has the units of an energy density. It is therefore tempting to believe that in some regimes such a product will be the solution of a kinetic equation and that the kinetic model will offer the right transition kernel from the source $\mathbf{S}$ to the back-propagated signal u^B.

We consider an asymptotic solution of the time reversal problem 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 consequently consider the initial pulse to be of the form

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

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 $u^B(x)$ in the vicinity of x_0. We therefore introduce the scaling $x = x_0 + \varepsilon\xi$. With these changes of variables, expression (7.6) is recast as

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

where

$$\chi(y,y') = \chi(y)\chi(y')f(\frac{y-y'}{\varepsilon}). \tag{7.7}$$

We are interested in the limit of $u^B(\xi;x_0)$ as $\varepsilon \to 0$.

Let us introduce some notation. We define the *adjoint* Green's matrix, solution of

$$\begin{aligned} &\frac{\partial G_*(t,x;y)}{\partial t}A(x) + \frac{\partial G_*(t,x;y)}{\partial x^j}D^j = 0 \\ &G_*(0,x;y) = \Gamma(x)\Gamma_0 A^{-1}(x)\delta(x-y). \end{aligned} \tag{7.8}$$

This solution is useful because $G_*(t,x;y) = \Gamma(y)G(t,y;x)A^{-1}(x)\Gamma_0$. Define now

$$\begin{aligned} Q(T,x;q) &= \int_{\mathbb{R}^d} G(T,x;y)\chi(y)e^{iq\cdot y/\varepsilon}dy, \\ Q_*(T,x;q) &= \int_{\mathbb{R}^d} G_*(T,x;y)\chi(y)e^{-iq\cdot y/\varepsilon}dy, \end{aligned} \tag{7.9}$$

and the Wigner measure

$$W_\varepsilon(t,x,k) = \int_{\mathbb{R}^d} \hat{f}(q)U_\varepsilon(t,x,k;q)dq, \tag{7.10}$$

where

$$U_\varepsilon(t,x,k;q) = \int_{\mathbb{R}^d} e^{ik\cdot y}Q(t,x-\frac{\varepsilon y}{2};q)Q_*(t,x+\frac{\varepsilon y}{2};q)\frac{dy}{(2\pi)^d}. \tag{7.11}$$

Then some algebra [18] shows that

$$u^B(\xi;x_0) = \int_{\mathbb{R}^6} e^{ik\cdot(\xi-z)}\Gamma_0 W_\varepsilon(T,x_0+\varepsilon\frac{z+\xi}{2},k)\Gamma_0 A(x_0+\varepsilon z)\mathbf{S}(z)\frac{dzdk}{(2\pi)^d}. \tag{7.12}$$

We have thus reduced the analysis of $u(\xi;x_0)$ as $\varepsilon \to 0$ to that of the asymptotic properties of the Wigner transform W_ε. But we already know that the Wigner transform solves a kinetic equation in appropriate regimes of wave propagation. In the weak-coupling regime for instance, we obtain for sources of the form

$$\mathbf{S}(x) = \begin{pmatrix} \nabla\phi(x) \\ p(x) \end{pmatrix} \tag{7.13}$$

that in the limit $\varepsilon \to 0$, the Fourier transform $\xi \to k$ of the back-propagated signal is given by

$$\hat{u}^B(k;x_0) = a_-(T,x_0,k)\hat{S}_+(k)b_+(x_0,k) + a_+(T,x_0,k)\hat{S}_-(k)b_-(x_0,k). \tag{7.14}$$

The eigenvectors $b_\pm$ associated to the eigenvalues $\omega_\pm(x,k) = \pm c(x)|k|$ of the dispersion relation are as in (4.71).

The propagating amplitudes $a_\pm$ solve the following kinetic model

$$\frac{\partial a_\pm}{\partial t} \pm c_0\hat{k}\cdot\nabla_x a_\pm = \int_{\mathbb{R}^d} \sigma(k,p)(a_\pm(t,x,p) - a_\pm(t,x,k))\delta(c_0(|k|-|p|))dp \tag{7.15}$$

The scattering coefficient $\sigma(k,p)$ is the same as before. The initial conditions for the amplitudes $a_\pm$ are given by

$$a_\pm(0,x,k) = |\chi(x)|^2\hat{f}(k)(A_0(x)\Gamma(x)\mathbf{b}_\mp(x,k)\cdot\mathbf{b}_\pm(x,k)). \tag{7.16}$$

Here $A_0 = \mathrm{Diag}(\rho_0,\rho_0,\rho_0,\kappa_0)$. Note that when $\Gamma(x) = \Gamma_0$, then $a_\pm(0,x,k) = |\chi(x)|^2\hat{f}(k)$.

We thus see that the refocusing properties of the time reversal experiment are captured by the behavior of the amplitudes $a_\pm$.

7.3. **The random medium in time reversal: a filtering process.** Equation (7.14) thus shows that the underlying medium acts as a *filter* for the time reversal experiment. A refocused signal will be "good" if it looks like the emitted signal. This means that the filter in (7.14) should be as close as possible to a constant value.

In a homogeneous medium with $\Gamma(x) = \Gamma_0$, we observe that

$$a_\pm(t, x_0, k) = |\chi(x_0 \mp c_0 \hat{k} t)|^2 \hat{f}(k). \tag{7.17}$$

Unless χ is uniformly equal to a constant (when there are detectors everywhere) and $\hat{f}(k) = 1$ (when measurements are perfect), we observe that these amplitudes become more and more singular in k as time grows since their gradient in k grows linearly with time. In a random medium however, multiple scattering renders the solution smoother and smoother as time grows (see [18] for a concrete definition of this). In the limit of vanishing mean free path (highly scattering media), the amplitudes

$$a_+(t, x, k) \approx a_-(t, x, k) \approx a(t, x, |k|),$$

asymptotically solve a diffusion equation

$$\begin{aligned} &\frac{\partial a(t, x, |k|)}{\partial t} - D(|k|)\Delta_x a(t, x, |k|) = 0, \\ &a(0, x, |k|) = |\chi(x)|^2 \frac{1}{4\pi |k|^2} \int_{\mathbb{R}^d} \hat{f}(q)\delta(|q| - |k|) dq. \end{aligned} \tag{7.18}$$

In this setting, we obtain that

$$\hat{u}^B(k; x_0) = a(T, x_0, |k|)\hat{\mathbf{S}}(k), \tag{7.19}$$

for a solution $a(T, x_0, |k|)$ that is quite smooth (in the $|k|$ variable).

Assuming that $\rho_1 \equiv 0$ and that $\mathbb{E}\{\hat{\kappa}_1(p)\hat{\kappa}_1(q)\} = \kappa_0^2 \hat{R}_0 \delta(p+q)$, then we find that

$$D(|k|) = \frac{c_0^2}{3\Sigma(|k|)} = \frac{c_0}{6\pi^2 |k|^4 \hat{R}_0}. \tag{7.20}$$

When $\hat{f}(k) = \hat{f}(|k|)$, the solution of (7.18) takes the form

$$a(T, x_0, |k|) = \hat{f}(|k|)\Big(\frac{3\pi |k|^4 \hat{R}_0}{2c_0 T}\Big)^{3/2} \int_{\mathbb{R}^d} \exp\Big(-\frac{3\pi^2 |k|^4 \hat{R}_0 |x_0 - y|^2}{2c_0 T}\Big)|\chi(y)|^2 dy. \tag{7.21}$$

When $f(x) = \delta(x)$, and $\Omega = \mathbb{R}^d$, 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 single-time time reversal process. Notice that both the low and high frequencies are damped when $\chi(x_0) = 0$, i.e., there are no detectors at the source location.

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: $x_0 \in \Omega$.

Expressions such as (7.19) and (7.21) give both a quantitative and physically simple explanation for the good refocusing properties of time reversed waves in highly heterogeneous media. Note that the statistical stability we obtained for kinetic models extends to the analysis of time reversed waves, at least in those regimes of wave propagation where rigorous results may be obtained.

7.4. **Time reversal and changing media.** Let us now assume that the medium in which the back-propagating signal propagates in the second stage of the time reversal experiment is *different* from the medium in which the first signal propagated. From the point of view of the modeling, this is straightforward. All we need to do is to replace (7.6) by

$$\begin{aligned} u^B(x) &= \Gamma_0 u(2T,x) \\ &= \int_{\mathbb{R}^9} \Gamma_0 G_2(T,x;y)\Gamma(y')G_1(T,y';z)\chi(y)\chi(y')f(y-y')\mathbf{S}(z)dydy'dz, \end{aligned} \tag{7.22}$$

where G_j is the Green's function corresponding to propagation in medium $j=1$ for the forward propagation and $j=2$ for the backward propagation. The Wigner transform in (7.10)-(7.11) thus needs to be modified accordingly. But we have already obtained kinetic models in the weak coupling regime for the correlation of wave fields propagating in different media.

Consider fluctuations of the form

$$\begin{aligned} (c_\varepsilon^\varphi)^2(x) &= c_0^2 - \sqrt{\varepsilon}V^\varphi(\frac{x}{\varepsilon}), \quad \varphi=1,2, \\ c_0^2 &= \frac{1}{\kappa_0\rho_0}, \qquad V^\varphi(x) = \frac{c_0^2}{\kappa_0}\kappa_1^\varphi(x), \end{aligned}$$

where c_0 is the average background speed and κ_1^φ and V^φ are random fluctuations in the compressibility and sound speed, respectively.We assume that $V^\varphi(x)$, $\varphi = 1,2$, are **statistically homogeneous** mean-zero random fields with correlation functions and power spectra given by:

$$\begin{aligned} c_0^4 R^{\varphi\psi}(x) &= \mathbb{E}\left\{V^\varphi(y)V^\psi(y+x)\right\}, \qquad 1\le\varphi,\psi\le 2, \\ (2\pi)^d c_0^4 \hat{R}^{\varphi\psi}(p)\delta(p+q) &= \mathbb{E}\left\{\hat{V}^\varphi(p)\hat{V}^\psi(q)\right\}. \end{aligned}$$

Then we find [21, 71] that $a_+(t,x,k)$ solves the following equation

$$\begin{aligned} &\frac{\partial a_+}{\partial t} + c_0\hat{k}\cdot\nabla a_+ + (\Sigma(k)+i\Pi(k))a_+ \\ &\quad = \frac{\pi\omega_+^2(k)}{2(2\pi)^d}\int_{\mathbb{R}^d}\hat{R}^{12}(k-q)a_+(q)\delta\big(\omega_+(q)-\omega_+(k)\big)dq, \\ &\Sigma(k) = \frac{\pi\omega_+^2(k)}{2(2\pi)^d}\int_{\mathbb{R}^d}\frac{\hat{R}^{11}+\hat{R}^{22}}{2}(k-q)\delta\big(\omega_+(q)-\omega_+(k)\big)dq \\ &i\Pi(k) = \frac{i\pi\sum_{j=\pm}}{4(2\pi)^d}\,\text{p.v.}\int_{\mathbb{R}^d}\Big(\hat{R}^{11}-\hat{R}^{22}\Big)(k-q)\frac{\omega_j(k)\omega_+(q)}{\omega_j(q)-\omega_+(k)}dq. \end{aligned} \tag{7.23}$$

Here, we have $\omega_\pm(k) = \pm c|k|$. The other mode is given by $a_-(t,x,k) = a_+^*(t,x,-k)$, a property which is always satisfied by uniqueness of the transport solutions and which ensures that the back-propagated signal is real-valued. The bouncary conditions, as in (7.16), are still given by $a_\pm(0,x,k) = |\chi(x)|^2\hat{f}(k)$ assuming $\Gamma(x)=\Gamma_0$.

Let us consider a very simple example where the randomness is spatially shifted by a factor $\varepsilon\boldsymbol{\tau}$ before back-propagation. This corresponds to assuming that $\hat{V}^2(p) = e^{ip\cdot\boldsymbol{\tau}}\hat{V}^1(p)$. Then we observe that $\hat{R}^{11} = \hat{R}^{22} = \hat{R}$ and that $\hat{R}^{12}(p) = e^{-ip\cdot\boldsymbol{\tau}}\hat{R}(p)$. As a consequence, $\alpha_+(t,x,k) = e^{ip\cdot\boldsymbol{\tau}}a(t,x,k)$ solves the transport equation (7.23) with $\boldsymbol{\tau}\equiv 0$ (i.e., the transport equation when both propagating media are the same).

In the diffusion approximation, with perfect detectors $\hat{f}(q) = 1$, we thus obtain that

$$\begin{aligned}&\frac{\partial \alpha(t,x,|k|)}{\partial t} - D(|k|)\Delta_x \alpha(t,x,|k|) = 0,\\ &\alpha(0,x,|k|) = |\chi(x)|^2 \frac{1}{4\pi |k|^2}\int_{\mathbb{R}^d} e^{iq\cdot\boldsymbol{\tau}} \delta(|q|-|k|)dq = |\chi(x)|^2 \frac{\sin|\boldsymbol{\tau}||k|}{|\boldsymbol{\tau}||k|}.\end{aligned} \tag{7.24}$$

As a consequence, we find that

$$\hat{u}^B(k;x_0) = e^{-i\boldsymbol{\tau}\cdot k}\frac{\sin|\boldsymbol{\tau}||k|}{|\boldsymbol{\tau}||k|} a_0(T,x_0,|k|)\hat{\mathbf{S}}(k), \tag{7.25}$$

where $a_0(T,x_0,|k|)$ is the solution of (7.18) obtained for $\boldsymbol{\tau} = 0$.

In the simplified diffusive regime of propagation, we thus obtain the following striking result: if the random medium is shifted by $\frac{\lambda}{2}$, then the the filter for the corresponding wavenumber $\lambda|k| = 2\pi$ *vanishes*. There is simply no refocusing for such a frequency. This behavior as a sinc function, which should be replaced by a Bessel function is two dimensions of space, was very well reproduced in numerical simulations [21] as well as experimental data [61].

Even though time reversal is relatively robust with respect to errors in the measurements (modeled by blurring $f(x)$ or by very inaccurate time reversion $\Gamma(x)$), it is seen to be quite unstable with respect to changes in the random medium between the two propagation stages of the experiment. This also explains the degrading refocusing properties observed in time reversal oceanic experiments [36] as the heterogeneities in the ocean change as a function of time.

7.5. **Time reversal and imaging.** Time reversal enhanced refocusing properties at first look very promising in imaging of buried inclusions in random media. If the inclusion may be made active, then the back-propagated signal will indeed refocus at the location of the inclusion. If back-propagation is performed numerically on a computer rather than in the physical environment, then time reversal seems to provide a very powerful method to localize such inclusions.

In most applications, however, the underlying heterogeneous medium (the medium we have modeled as random for want of a better description) is not known very accurately. Back-propagation on the computed will therefore inevitably occur using an approximation of the underlying medium. How accurate should this approximation be for the refocused signal to be tightly focused at the location of the inclusion? The answer, unfortunately, is that the approximation should be very accurate. The theory of time reversal in changing media seen above was precisely devised to quantify how accurate the approximation should be.

We have already mentioned that refocusing was "good" when the transport solution was smooth. And what generates the smoothness of the transport equation is strong scattering. Yet, the scattering operator in the above equation is seen to be proportional to the quantity $\hat{R}^{12}$, in other words to the cross-correlation of the two random media. When such a correlation is large, then refocusing will be very strong. When the correlation is weak, then refocusing will be weak as well. Assuming that we know the statistics of the random medium, it is relatively easy to construct a realization of the random medium. If that realization is chosen uncorrelated with the "true" heterogeneous medium in which the physical signals propagate, then we may conclude that the two random media V^1 and V^2 are uncorrelated, in which case $\hat{R}^{12} = 0$. This is often the best we can achieve on a computer unless detailed knowledge of the highly oscillatory random medium is available. Back-propagation will thus occur with $a_\pm$ solutions of transport equations with no scattering. We know that such solutions are not smooth and that the refocusing properties of the reconstructed signal will be poor. For this reason, time reversal is difficult to use as an imaging methodology.

REFERENCES

[1] F. Bailly, J. F. Clouet and J.-P. Fouque, *Parabolic and gaussian white noise approximation for wave propagation in random media*, SIAM J. Appl. Math, **56** (1996), 1445–1470.

[2] G. Bal, *On the self-averaging of wave energy in random media*, Multiscale Model. Simul., **2** (2004), 398–420.

[3] G. Bal, *Kinetics of scalar wave fields in random media*, Wave Motion, **43** (2005), 132–157.

[4] G. Bal, *Inverse problems in random media: A kinetic approach*, J. Phys. Conf. Series, **124** (2008), 012001.

[5] G. Bal, *Inverse transport theory and applications*, Inverse Problems, **25** (2009), 053001.

[6] G. Bal, L. Carin, D. Liu and K. Ren, *Experimental validation of a transport-based imaging method in highly scattering environments*, Inverse Problems, **26** (2007), 2527–2539.

[7] G. Bal, D. Liu, S. Vasudevan, J. Krolik and L. Carin, *Electromagnetic time-reversal imaging in changing media: Experiment and analysis*, IEEE Trans. Anten. and Prop., **55** (2007), 344–354.

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

[9] G. Bal, T. Komorowski and L. Ryzhik, *Asymptotics of the solutions of the random Schródinger equation*, to appear in Arch. Rat. Mech., (2010).

[10] G. Bal, G. Papanicolaou and L. Ryzhik, *Radiative transport limit for the random Schrödinger equations*, Nonlinearity, **15** (2002), 513–529.

[11] G. Bal, G. Papanicolaou and L. Ryzhik, *Self-averaging in time reversal for the parabolic wave equation*, Stochastics and Dynamics, **4** (2002), 507–531.

[12] G. Bal and O. Pinaud, *Time reversal-based imaging in random media*, Inverse Problems, **21** (2005), 1593–1620.

[13] G. Bal and O. Pinaud, *Accuracy of transport models for waves in random media*, Wave Motion, **43** (2006), 561–578.

[14] G. Bal and O. Pinaud, *Kinetic models for imaging in random media*, Multiscale Model. Simul., **6** (2007), 792–819.

[15] G. Bal and O. Pinaud, *Self-averaging of kinetic models for waves in random media*, Kinetic Related Models, **1** (2008), 85–100.

[16] G. Bal and O. Pinaud, *Imaging using transport models for wave-wave correlations*, to appear in M3AS, (2011).

[17] G. Bal and K. Ren, *Transport-based imaging in random media*, SIAM J. Applied Math., **68** (2008), 1738–1762.

[18] G. Bal and L. Ryzhik, *Time reversal and refocusing in random media*, SIAM J. Appl. Math., **63** (2003), 1475–1498.

[19] G. Bal and L. Ryzhik, *Time splitting for wave equations in random media*, preprint, (2004).

[20] G. Bal and L. Ryzhik, *Stability of time reversed waves in changing media*, Discrete Contin. Dyn. Syst., **12** (2005), 793–815.

[21] G. Bal and R. Verástegui, *Time reversal in changing environment*, Multiscale Model. Simul., **2** (2004), 639–661.

[22] W. Bao, S. Jin and P. A. Markowich, *On Time-Splitting spectral approximations for the Schrödinger equation in the semiclassical regime*, J. Comp. Phys., **175** (2002), 487–524.

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

[24] P. Blankenship and G. C. Papanicolaou, *Stability and control of stochastic systems with wide-band noise disturbances. i*, SIAM J. Appl. Math., **34** (1978), 437–476.

[25] P. Blomgren, G. C. Papanicolaou and H. Zhao, *Super-Resolution in Time-Reversal Acoustics*, J. Acoust. Soc. Am., **111** (2002), 230–248.

[26] B. Borcea, G. C. Papanicolaou and C. Tsogka, *Theory and applications of time reversal and interferometric imaging*, Inverse Problems, **19** (2003), S139–S164.

[27] B. Borcea, G. C. Papanicolaou and C. Tsogka, *Interferometric array imaging in clutter*, Inverse Problems, **21** (2005), 1419–1460.

[28] B. Borcea, G. C. Papanicolaou and C. Tsogka, *Adaptive interferometric imaging in clutter and optimal illumination*, Inverse Problems, **22** (2006), 1405–1436.

[29] A. N. Borodin, *A limit theorem for solutions of differential equations with random right hand side*, Teor. Veroyatn. Ee Primen, **22** (1977), 498–512.

[30] R. A. Carmona and S. A. Molchanov, *Parabolic Anderson problem and intermittency*, Mem. Amer. Math. Soc, **108** (1994), viii+125.

[31] S. Chandrasekhar, "Radiative Transfer," Dover Publications, New York, 1960.

[32] G. C. Cohen, "Higher-Order Numerical Methods for Transient Wave Equations," Scientific Computation, Springer Verlag, Berlin, 2002.

[33] D. A. Dawson and G. C. Papanicolaou, *A random wave process*, Appl. Math. Optim., **12** (1984), 97–114.

[34] 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.

[35] D. R. Durran, "Numerical Methods for Wave equations in Geophysical Fluid Dynamics," Springer, New York, 1999.

[36] G. F. Edelmann, T. Akal, W. S. Hodgkiss, S. Kim, W. A. Kuperman and H. C. Song, *An initial demonstration of underwater acoustic communication using time reversal*, IEEE J. Oceanic Eng., **27** (2002), 602–609.

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

[38] L. Evans and M. Zworski, Lectures on semiclassical analysis (Berkeley).

[39] A. Fannjiang, *Self-averaging in scaling limits for random high-frequency parabolic waves*, Archives of Rational Mechanics and Analysis, **175** (2005), 343–387.

[40] J. P. Fouque, *La convergence en loi pour les processus à valeur dans un espace nucléaire*, Ann. Inst. H. Poincaré Prob. Stat, **20** (1984), 225–245.

[41] J.-P. Fouque, J. Garnier, G. Papanicolaou and K. Sølna, "Wave Propagation and Time Reversal in Randomly Layered Media," Springer Verlag, New York, 2007.

[42] P. Gérard, *Microlocal defect measures*, Comm. PDEs, **16** (1991), 1761–1794.

[43] P. Gérard, P. A. Markowich, N. J. Mauser and F. Poupaud, *Homogenization limits and Wigner transforms*, Comm. Pure Appl. Math., **50** (1997), 323–380.

[44] F. Golse, S. Jin and C. D. Levermore, *The convergence of numerical transfer schemes in diffusive regimes. I. Discrete-ordinate method*, SIAM J. Numer. Anal., **36** (1999), 1333–1369.

[45] T. G. Ho, L. J. Landau and A. J. Wilkins, *On the weak coupling limit for a Fermi gas in a random potential*, Rev. Math. Phys., **5** (1992), 209–298.

[46] H. Hochstadt, "The Functions of Mathematical Physics," Dover Publications, New York, 1986.

[47] T. Y. Hou, X. Wu and Z. Cai, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., **227** (1999), 913–943.

[48] , I. A. Ibragimov and Yu. V. Linnik, "Independent and Stationary Sequences of Random Variables," Wolters-Noordhoff Publishing, Groningen, 1971.

[49] A. Ishimaru, "Wave Propagation and Scattering in Random Media," New York, Academics, 1978.

[50] J. Jacod and A. N. Shiryaev, "Limit Theorems for Stochastic Processes," Grundlehren der mathematischen Wissenschaft **288**, Springer-Verlag, 2003.

[51] A. Jakubowski, *A non-Skorohod topology on the Skorohod space*, Electron. J. Probability, **2** (1997), 1–21.

[52] H. Kesten and G. Papanicolaou, *A limit theorem for turbulent diffusion*, Comm. Math. Phys., **65** (1979), 97–128.

[53] H. Kesten and G. C. Papanicolaou, *A limit theorem for stochastic acceleration*, Comm. Math. Phys., **78** (1980), 19–63.

[54] R. Khasminskii, *A limit theorem for solutions of differential equations with a random right hand side*, Theory Probab. Appl., **11** (1966), 390–406.

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

[56] T. Komorowski, Sz. Peszat and L. Ryzhik, *Limit of fluctuations of solutions of Wigner equation*, Comm. Math. Phys., **292** (2009), 479–510.

[57] T. Komorowski and L. Ryzhik, *Diffusion in a weakly random Hamiltonian flow*, Comm. Math. Phys., **263** (2006), 277–323.

[58] T. Komorowski and L. Ryzhik, *The stochastic acceleration problem in two dimensions*, Israel Jour.Math., **155** (2006), 157–204.
[59] T. Komorowski and L. Ryzhik, *Asymptotics of the phase of the solutions of the random Schrödinger equation*, preprint, (2010).
[60] P.-L. Lions and T. Paul, *Sur les mesures de Wigner*, Rev. Mat. Iberoamericana, **9** (1993), 553–618.
[61] D. Liu, S. Vasudevan, J. Krolik, G. Bal and L. Carin, *Electromagnetic time-reversal imaging in changing media: Experiment and analysis*, IEEE Trans. Antennas and Prop., **55** (2007), 344–354.
[62] J. Lukkarinen and H. Spohn, *Kinetic limit for wave propagation in a random medium*, Arch. Ration. Mech. Anal., **183** (2007), 93–162.
[63] P. Markowich, P. Pietra and C. Pohl, *Numerical approximation of quadratic observables of Schrödinger-type equations in the semi-classical limit*, Numer. Math., **81** (1999), 595–630.
[64] P. Markowich, P. Pietra, C. Pohl and H. P. Stimming, *A Wigner-measure analysis of the Dufort-Frankel scheme for the Schrödinger equation*, SIAM J. Numer. Anal., **40** (2002), 1281–1310.
[65] I. Mitoma, *On the sample continuity of $\mathcal{S}'$ processes*, J. Math. Soc. Japan, **35** (1983), 629–636.
[66] B. Øksendal, "Stochastic Differential Equations," Springer-Verlag, Berlin, 2000.
[67] G. Papanicolaou and W. Kohler, *Asymptotic theory of mixing stochastic ordinary differential equations*, Comm. Pure Appl. Math., **27** (1974), 641–668.
[68] G. Papanicolaou, L. Ryzhik and K. Sølna, *The parabolic wave approximation and time reversal*, Matematica Contemporanea, **23** (2002), 139–159.
[69] G. C. Papanicolaou, L. Ryzhik and K. Sølna, *Self-averaging from lateral diversity in the Ito-Schroedinger equation*, Multiscale Model. Simul., **6** (2007), 468–492.
[70] F. Poupaud and A. Vasseur, *Classical and quantum transport in random media*, J. Math. Pures Appl., **6** (2003), 711–748.
[71] L. Ryzhik, G. Papanicolaou and J. B. Keller, *Transport equations for elastic and other waves in random media*, Wave Motion, **24** (1996), 327–370.
[72] H. Sato and M. C. Fehler, "Seismic Wave Propagation and Scattering in the Heterogeneous Earth," AIP series in modern acoustics and signal processing, AIP Press, Springer, New York, 1998.
[73] P. Sheng, "Introduction to Wave Scattering, Localization and Mesoscopic Phenomena," Academic Press, New York, 1995.
[74] H. Spohn, *Derivation of the transport equation for electrons moving through random impurities*, Jour. Stat. Phys., **17** (1977), 385–412.
[75] G. Strang, *On the construction and comparison of difference schemes*, SIAM J. Numer. Anal., **5** (1968), 507–517.
[76] D. W. Stroock and S. R. S. Varadhan, "Multidimensional Diffusion Processes," Grundlehren der mathematischen Wissenschaften **233**, Berlin, Heidelberg, New York, 1979, 1997.
[77] C. R. Vogel, "Computational Methods for Inverse Problems," Frontiers Appl. Math., SIAM, Philadelphia, 2002.
[78] E. Wigner, *On the quantum correction for thermodynamic equilibrium*, Physical Rev., **40** (1932), 749–759.
[79] B. White, *The stochastic caustic*, SIMA Jour. Appl. Math., **44** (1984), 127–149.

Received September 2010; revised October 2010.

E-mail address: gb2030@columbia.edu
E-mail address: komorow@hektor.umcs.lublin.pl
E-mail address: ryzhik@math.stanford.edu