Dynamical large deviations for an inhomogeneous wave kinetic theory: linear wave scattering by a random medium

Yohei Onuki, Jules Guioth, Freddy Bouchet

Dynamical large deviations for an inhomogeneous wave kinetic theory: linear wave scattering by a random medium

Yohei Onuki[1,2*], Jules Guioth[2] and Freddy Bouchet[2,3]

[1]Research Institute for Applied Mechanics, Kyushu University, Kasuga, Fukuoka, Japan.
[2]Univ. Lyon, ENS de Lyon, CNRS, Laboratoire de Physique, Lyon, France.
[3]Laboratoire de Météorologie Dynamique, CNRS, Ecole Normale Supérieure, Institut Pierre-Simon Laplace, Paris Sciences Lettres Université, Paris, France.

*Corresponding author(s). E-mail(s): onuki@riam.kyushu-u.ac.jp;

Abstract

The wave kinetic equation predicts the averaged temporal evolution of a continuous spectral density of waves either randomly interacting or scattered by the fine structure of a medium. In a wide range of systems, the wave kinetic equation is derived from a fundamental equation of wave motion, which is symmetric through time-reversal. By contrast, the corresponding wave kinetic equations is time-irreversible: its solutions monotonically increase an entropy-like quantity. A similar paradox appears whenever one make a mesoscopic description of the evolution of a very large number of microscopic degrees of freedom, the paradigmatic example being the kinetic theory of dilute gas molecules leading to the Boltzmann equation. Since Boltzmann, it has been understood that a probabilistic understanding solves the apparent paradox. More recently, it has been understood that the kinetic description itself, at a mesoscopic level, should not break time reversal symmetry [1]. The time reversal symmetry remains a fundamental property of the mesoscopic stochastic process: without external forcing the path probabilities obey a detailed balance relation with respect to an equilibrium quasipotential. The proper theoretical or mathematical tool to derive fully this mesoscopic time reversal stochastic process is large deviation theory: a

large deviation principle uncovers a time reversible field theory, characterized by a large deviation Hamiltonian, for which the deterministic wave kinetic equation appears as the most probable evolution. Its irreversibility appears as a consequence of an incomplete description, rather than as a consequence of the kinetic limit itself, or some related chaotic hypothesis. This paper follows [1] and a series of other works that derive the large deviation Hamiltonians of the main classical kinetic theories, for instance [2] for homogeneous wave kinetics. We propose here a derivation of the large deviation principle in an inhomogeneous situation, for the linear scattering of waves by a weak random potential. This problem involves microscopic scales corresponding to the typical wavelengths and periods of the waves and mesoscopic ones which are the scales of spatial inhomogeneities in the spectral density of both the random scatterers and the wave spectrum, and the time needed for the random scatterers to alter the wave spectrum. The main assumption of the kinetic regime is a large separation of these microscopic and mesoscopic scales. For the sake of simplicity, we consider a generic model of wave scattering by weak disorder: the Schrödinger equation with a random potential. We derive the path large deviation principle for the local spectral density and discuss its main properties. We show that the mesoscopic process obeys a time-reversal symmetry at the level of large deviations.

This publication is part of a special issue in homage of the memory of Krzysztof Gawędzki. The subject of this work is large deviation theory applied to wave turbulence. Large deviation theory applied to complex dynamics and turbulent flows was one of the subjects for which Krzysztof Gawędzki made a number of important contributions during the last few years, see for instance [3–8]. He taught many of us, including Freddy Bouchet, many aspects of large deviation theory. We wrote a common paper on the subject of large deviation theory and non-equilibrium quasipotentials for stochastic particles with mean field interactions [8]. Given his scientific qualities, and his deep sense of friendship, it is great pleasure for us to pay homage to Krzysztof Gawędzki through this modest contribution.

1 Introduction

The aim of this paper is to extend the existing kinetic theory to describe probabilistically mesoscopic evolutions of wave fields interacting with random potentials. We will derive for the first time a large deviation principle that describes completely typical and rare fluctuations of the wave local spectral density. This is also the first large deviation principle for wave kinetic theory in an inhomogeneous setup. This work lies at the intersection of three different active fields in theoretical and mathematical physics: the description of waves interacting with random media and their applications to ocean and atmosphere

dynamics, recent mathematical and theoretical advances in the kinetic theory of wave turbulence, and large deviation theory for kinetic theories.

For the first field, we note that wave propagation in disordered media is a ubiquitous phenomenon appearing in various areas of physics. Typical examples include light radiation through the atmosphere, acoustic or internal gravity waves in turbulent flows, and elastic waves in solid Earth. In most cases, one is not interested in the individual wave interference or scattering processes but in the statistical description of the overall wave field at a mesoscopic spatial scale much greater than the extent of disorder or wavelength. For this purpose, it is customary to define the spectral density of the wave signal at each location and investigate its statistical properties. The wave kinetic equation, sometimes referred to as a radiative transport equation or simply as a transport equation, is known as the universal model to describe the evolution of the local spectral density. It commonly derives from elementary wave equations and has a broad range of applications [9–14]. Recently, the evolution of wave spectra under scattering interactions with a turbulent flow were studied in a two-dimensional model [15].

Wave kinetic theory is of special interest in some specific areas of ocean and atmosphere research. Since the celebrated work by Klaus Hasselmann [16], the kinetic description of nonlinear 4-wave interactions among water waves has been used for estimating energy transfer rates in a wind wave spectrum and forecasting the sea surface states. The linear counterpart of the wave kinetic equation is relevant to surface or internal wave energy dispersion in a slowly evolving turbulent flow [12, 13, 17–19]. In these actual problems, the scale-separation assumption at the heart of the kinetic theory might be valid, but is not necessarily always valid. For example, internal wave activity in the ocean is highly heterogeneous, which is imprinted on the variability of energy dissipation rates on scales of order 10 to 100 km, in the mid-depth layer [20]. For tide or wind generated waves with 10-100 km horizontal wavelengths, deviations of the spectral evolutions from that predicted by the kinetic equation may not be negligible, and a first principle theory of fluctuation is missing. This motivates us to revisit the theoretical basis of wave kinetic theory. Our work can be considered as the first building block for stochastic parameterization of the local spectral density from first principles, for the specific case of wave interacting with random potentials.

In relation with geophysical applications, several experiments with fundamental scopes in wave kinetic theory have been recently devised. For instance this led to the very first observation of the regime of inertial wave turbulence in a rotating flow [21], the identification of regimes of weakly and strongly nonlinear internal wave turbulence in an experiment of stratified turbulence [22], experiments on statistical properties of water waves in a large basin [23], the validation of the inverse cascade phenomenon [24], and extension of the range of scales for observing pure gravity wave turbulence in the laboratory [25] using reduced gravity experiments.

The second field, fundamental theoretical developments in wave turbulence theory, has seen many new advances recently. For the first time, using numerical simulations of the non-linear Schrödinger wave kinetic equation, predictions by the wave kinetic equation were tested for several kinetic times [26, 27]. Novel finite-size effects in wave turbulence were systematically studied in a one-dimensional model using a combination of theory and numerics [28]. Significant recent progress has been made to give a mathematical foundation for wave turbulence theory: theorem about approximations of the dynamics for times much shorter than the kinetic time [29–34], the understanding of propagation of chaos [35], and a remarkable first full rigorous derivation of the wave kinetic theory at the kinetic timescale, for the non-linear Schrödinger equation [30] and for water waves equation [36]. From the point of view of these fundamental perspectives, our work gives for the first time a description of all the cumulants of the local spectral density, through a large deviation principle, in an inhomogeneous setting.

The third field is the development of large deviation principles in relation with kinetic theory. Many classical equations of mathematical physics arise from a law of large numbers, when faster and smaller scale degrees of freedom are averaged out. This is the case for all classical kinetic theories. It is natural to extend all these theories to look for the statistics of fluctuations. Generically, one expects to derive a large deviation principle that describes a statistical field theory quantifying the probabilities of any fluctuations, either typical or extremely rare, in a way analogous to macroscopic fluctuation theory [37] for stochastic diffusive systems, or large deviation theory for stochastic dynamics with mean field interaction [8]. Deriving such large deviation principles from deterministic microscopic dynamics is a fundamental endeavor in theoretical and mathematical physics. Recently, the large deviation principles for a number of classical kinetic theories, starting from first principles, have been uncovered: for discrete models that mimic dilute gases and with Boltzmann like behavior [38, 39], for dilute gases related to the Boltzmann equation [1, 40], for the Kac model [41, 42], for plasma at length scales much smaller than the Debye length related to the Landau equation [43], for homogeneous systems with long range interactions related to the Balescu–Guernsey–Lenard equation [44], for weakly interacting waves in a homogeneous setup [2] related to the wave kinetic equation. The large deviation principles describe fluctuations but also uncover gradient structure for the deterministic kinetic equation, see [45] and a simple explanation in [1]. Several mathematical results, usually valid for a fraction of the kinetic time in the spirit of Lanford results for the deterministic equation, have been obtained for the large deviation principles, for instance for the Boltzmann equation [40, 46], or for the Kac model with unexpected corrections to the expected large deviation principle [41, 42] associated to giant concentrations and solutions of the Boltzmann equation without energy conservation.

One aim of large deviation theory is to study rare events. In the context of wave dynamics, large deviation theory has been used to study rare events

for the evolution of the empirical spectrum [2] on the kinetic time scale, but also for studying the appearance of very large amplitude waves [47], for the non-linear Schrödinger dynamics for shorter time scales. Instantons structures have been predicted and compared with experimental data taken from a 300 m long wave [47]. Large deviation principles for the wave amplitude due to short time phase mixing has also been studied [48].

The result described in this paper is a new example of a large deviation principle for a kinetic theory, derived from microscopic dynamics. It is the first extension of large deviation theory for the local spectral density in an inhomogeneous setup. It opens the way for other inhomogeneous large deviation principle for wave turbulence, and for the study of new wave turbulence phenomena where rare events play an important role.

As a generic model of waves interacting with random medium, we consider the linear Schrödinger equation in a weak random potential

$$i\frac{\partial \psi}{\partial t} = -\frac{D}{2}\nabla_x^2 \psi + V\psi,$$

where $\psi(\boldsymbol{x}, t)$ is a wave function defined on $\mathbb{R}^{d+1}$ and $V(\boldsymbol{x})$ is a homogeneous random potential. For this model, we assume a regime with a wave spectrum which is dominated by waves of typical wavelengths λ, and with modulations of the statistical properties of the wave spectrum on scales of order λ/μ. The second assumption is that the typical correlation length of the potential is of order λ and that interactions between the waves and the potential is weak, more precise definitions are given in section 2. Then, for small value of μ, we have a separation of scales and of the associated times, where a huge amount of waves experience multiple scattering in domains of typical size λ/μ. It is natural to focus on variations of the field on the mesoscopic scales of order λ/μ. This defines a kinetic regime where such mesoscopic variations are captured by the Wigner distribution n, that somehow measures the wave energy density in both position and wave-vector space. After time and length rescaling $t \to \mu t$ and $\boldsymbol{x} \to \mu \boldsymbol{x}$, in the small μ limit, the wave kinetic equation is classically derived (see for instance [9, 49]):

$$\frac{\partial n(\boldsymbol{x}, \boldsymbol{p}, t)}{\partial t} + \boldsymbol{p} \cdot \nabla_x n(\boldsymbol{x}, \boldsymbol{p}, t) = c \int d\boldsymbol{\eta} \sigma(\boldsymbol{p}, \boldsymbol{\eta}) \left(n(\boldsymbol{x}, \boldsymbol{\eta}, t) - n(\boldsymbol{x}, \boldsymbol{p}, t) \right),$$

where $n(\boldsymbol{x}, \boldsymbol{p}, t)$ is the Wigner distribution at position $\boldsymbol{x}$, wave vector $\boldsymbol{p}$ and time t, and $c\sigma(\boldsymbol{p}_1, \boldsymbol{p}_2)$ is the scattering cross section.

Interestingly, the evolution of the Wigner distribution predicted from the wave kinetic equation is an irreversible relaxation process. For this problem, a Lyapunov function, $S = \int d\boldsymbol{x} d\boldsymbol{p} \log n(\boldsymbol{x}, \boldsymbol{p})$, monotonically increases with time, even though the fundamental equation of motion possesses a time-reversal symmetry. This old irreversibility paradox has been recently revisited for the kinetic theory of particles [1], using dynamical large deviation principles, in the case of the Boltzmann equation. It turns out that the dynamical large deviation

principle that quantifies the probability for the evolution of any trajectory has a time reversal symmetry. The kinetic equation corresponds to the most probable path of the system, while the probability of a path and its time-reversed path is related through detailed balance, a manifestation of time reversal symmetry for the mesoscopic stochastic process. This gives an extremely simple and enlightening explication of the irreversibility paradox. The main result of the present paper is a large deviation principle for the Schrödinger equation in a weak random potential, which has also a time reversal symmetry. This gives a new very clear explanation of the time reversal paradox.

The purpose of this paper is to formulate a path large deviation principle for wave scattering by random disorder in spatially inhomogeneous problems. In particular, to make the discussion as concise as possible, we restrict our attention to the simplest Schrödinger equation model. Our fundamental results are as follows. First, (i) for a small but finite μ we show that the probability that a path of local spectral density $\{n^\mu(t)\}$ evolves at a vicinity of a prescribed specific path $\{n(t)\}$ satisfies a large deviation principle:

$$\mathbb{P}\left[\{n^\mu(t) = n(t)\}_{0\leqslant t\leqslant t_f}\right]$$
$$\underset{\mu\to 0}{\asymp} \exp\left(-\frac{1}{(2\pi\mu)^d}\int_0^{t_f} dt \sup_\lambda\left\{\int \lambda\dot{n} - \mathcal{H}[n,\lambda]\right\}\right)\mathbb{P}_0[n(0)],$$

where $\mathcal{H}$ is the large deviation Hamiltonian that generally governs the stochastic fluctuations of macroscopic variables and $\mathbb{P}_0[n(0)]$ is the probability of the initial condition $n^\mu(t=0)$. Obtaining the explicit expression for $\mathcal{H}$ from the microscopic equation is one of the main results of this paper. Next, (ii) we verify that the ordinary wave kinetic equation describes the path that minimizes the exponent of the probability functional. Then, (iii) we establish a large deviation principle for the microcanonical measure that defines the quasipotential of the mesoscopic stochastic process of the local spectral density. We analyze (iv) the property of the large deviation Hamiltonian, check its symmetries related to conservation laws and the time-reversal symmetry, and derive an expression of the detailed balance that connects the probabilities of a path and its time-reversed path. Finally (v) we study the diffusive limit when the scales of variation of the random potential are much larger than the typical wavelength of the waves. For this case we obtain a diffusive large deviation Hamiltonian, for which we check all the desired symmetries.

The paper is organized in the following order. In section 2, we first set up the basic problem, introduce scaling and statistical assumptions, and derive the ordinary form of the wave kinetic equation. In section 3, we derive the path large deviation principle for the temporal variations of the local spectral density. The approach has some analogy with that of [2], with a slightly different scaling assumption, and working with the Wigner distribution to describe the wave local spectral density. We then show that this Hamiltonian satisfies the expected properties. A remarkable point is that the quasipotential entering into the detailed balance relation is consistent with the one obtained from

a direct computation of microcanonical ensemble, formulated in Appendix B. Section 4 proposes several possible extensions in future studies.

2 Wave kinetic equation for a linear Schrödinger equation

2.1 Problem setup

We consider the Schrödinger equation for a wave function $\psi^*(\boldsymbol{x}^*, t^*) : \mathbb{R}^{d+1} \to \mathbb{C}$, where t^* is time and $\boldsymbol{x}^*$ is the position vector, and with potential $V^*(\boldsymbol{x}^*) : \mathbb{R}^d \to \mathbb{R}$:

$$i\frac{\partial \psi^*}{\partial t^*} = -\frac{D}{2}\nabla^2_{x^*}\psi^* + V^*\psi^*. \tag{1}$$

We use an upper script $*$ to represent variables with physical dimensions. The physical parameter $D > 0$ has dimension L^2T^{-1}. In absence of interaction with the potential, the free Schrödinger equation is the dynamics of linear waves with a dispersion relation $\omega(\boldsymbol{k}^*) = D|\boldsymbol{k}^*|^2/2$, where $\boldsymbol{k}^*$ is a wave vector. A localised wave packet propagates at group velocity $\nabla_{k^*}\omega(\boldsymbol{k}^*) = D\boldsymbol{k}^*$.

The potential $V^*(\boldsymbol{x}^*)$ is assumed to be a spatially homogeneous random field with zero-average, $\mathbb{E}[V^*] = 0$, with its spectral density given by

$$\Pi^*(\boldsymbol{k}^*) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} d\boldsymbol{y}^* e^{-i\boldsymbol{k}^*\cdot\boldsymbol{y}^*}\mathbb{E}\left[V^*\left(\boldsymbol{x}^* + \frac{\boldsymbol{y}^*}{2}\right)V^*\left(\boldsymbol{x}^* - \frac{\boldsymbol{y}^*}{2}\right)\right]. \tag{2}$$

For homogeneous fields, the two-point correlation function $\mathbb{E}\left[V^*(\boldsymbol{x}_1^*)V^*(\boldsymbol{x}_2^*)\right]$ depends only on the point separation $\boldsymbol{x}_1^* - \boldsymbol{x}_2^*$. The spectral density of the potential is the Fourier transform of the two-point correlation function $\mathbb{E}\left[V^*(\boldsymbol{x}_1^*)V^*(\boldsymbol{x}_2^*)\right]$ with respect to $\boldsymbol{x}_1^* - \boldsymbol{x}_2^*$ and thus contains the same information. Note that a prescription of higher order cumulants would be needed to fully characterize the potential distribution. As we will see in the following, the higher order statistics of the potential will not affect the dynamics of the spectral density of the waves in the kinetic regime. Hence, although we do not specify all the cumulants, the potential needs not to be Gaussian.

In this article, we assume that the spectral density Π^* is concentrated around wave vectors $|\boldsymbol{k}^*| \sim 2\pi/\lambda$ where λ is the typical wavelength. In real space, the wavelength λ is interpreted as the typical correlation length of the potential. For such a potential, the Schrödinger dynamics Eq. (1) may display many different regimes, depending on the order of magnitude of λ compared to typical wavelengths in the initial condition of ψ^*. For instance, if the initial condition is made of waves with wavelengths much smaller than λ, the wave-potential interaction corresponds to random but smooth refraction that leads to diffusion in the macroscopic limit [17–19, 50]. We will see in section 3.4 that this diffusive limit can be recovered from the wave kinetic regime. On the other hand, if the initial waves have wavelength much greater that λ, one faces

a homogenization problem that is not described by the wave kinetic equation [51, 52]. In the present paper, we will focus on an intermediate regime, when the initial condition is made of waves with typical wavelengths which are of order λ. We also make an assumption that the potential term is very small compared to the Laplacian term, by setting

$$\epsilon \equiv \frac{\lambda^2 V_0}{D} \ll 1, \tag{3}$$

where V_0 is a constant scaling the potential, i.e., the potential spectrum is typically $\Pi^* \sim V_0^2 \lambda^d$.

Wave energy or wave action measures the local amplitude of the signal. We denote ℓ the typical scale for spatial variation of wave action, and call it the mesoscopic spatial scale. We introduce the second natural dimensionless parameter

$$\mu \equiv \frac{\lambda}{\ell}. \tag{4}$$

The kinetic limit is the limit $\mu \ll 1$. In the context of wave kinetics, we are interested in the statistical behavior of the system at the mesoscopic scale, avoiding chasing rapid phase oscillations at scale λ. Since the group velocity of a wave packet is $D|\boldsymbol{k}^*| \propto D\lambda^{-1}$, the migration time of a wave packet over a mesoscopic distance ℓ is $\lambda^2/D\mu$. We call this time the mesoscopic time. Choosing such mesoscopic units naturally yields the following dimensionless coordinates

$$\boldsymbol{x} \equiv \mu \frac{\boldsymbol{x}^*}{\lambda}, \quad \boldsymbol{p} \equiv \lambda \boldsymbol{k}^*, \quad t \equiv \mu \frac{Dt^*}{\lambda^2}, \tag{5}$$

where the scaled wave vector is now represented by $\boldsymbol{p}$ in a customary way of quantum mechanics with μ corresponding to the Dirac constant. Physically, the square of the absolute value of the wave function, $|\psi^*|^2$, represents the wave action density that is proportional to energy, momentum, or number of particles contained in a unit volume. Therefore, on the dimensional ground, the wave function should be dependent on the scaling parameters as

$$\psi^\mu(\boldsymbol{x}) = \frac{\lambda^{d/2} \psi^*(\boldsymbol{x}^*)}{\mu^{d/2}}.$$

The potential and its spectrum are scaled as

$$V^\mu(\boldsymbol{x}) = \frac{V^*(\boldsymbol{x}^*)}{V_0}, \quad \Pi(\boldsymbol{p}) = \frac{\Pi^*(\boldsymbol{k}^*)}{V_0^2 \lambda^d},$$

such that

$$\Pi(\boldsymbol{p}) = \frac{1}{(2\pi\mu)^d} \int_{\mathbb{R}^d} d\boldsymbol{y} e^{-i\boldsymbol{p}\cdot\boldsymbol{y}/\mu} \mathbb{E}\left[V^\mu\left(\boldsymbol{x} + \frac{\boldsymbol{y}}{2}\right) V^\mu\left(\boldsymbol{x} - \frac{\boldsymbol{y}}{2}\right)\right]. \tag{6}$$

In the end, the governing equation (1) is rewritten in the non-dimensional form

$$i\mu \frac{\partial \psi^\mu}{\partial t} = -\frac{\mu^2}{2}\nabla_x^2 \psi^\mu + \epsilon V^\mu \psi^\mu, \tag{7}$$

which is the fundamental model of the present work.

For $\epsilon = 0$, waves $e^{i(\boldsymbol{p}\cdot\boldsymbol{x}-\omega t)/\mu}$, with $\omega = |\boldsymbol{p}|^2/2$, are exact solutions of the equations, illustrating that the microscopic time scale and spatial scales are $t_m \sim \mathcal{O}(\mu)$ and $\boldsymbol{x}_m \sim \mathcal{O}(\mu)$ respectively. A wave packet $a(\boldsymbol{x}, t)e^{i(\boldsymbol{p}\cdot\boldsymbol{x}-\omega t)/\mu}$ with modulation of its amplitude a on spatial scales of order one (slow modulation compared to the microscopic scale), will actually see an evolution of a on time scales of order one, according to the group velocity $\boldsymbol{p}$. In the limit of small ϵ, the effect of the inhomogeneous potential term is very small on the microscopic time scale. Therefore, a wave packet propagates almost freely in a microscopic time scale. Since $\mathbb{E}[V^\mu] = 0$ has been assumed, effects of terms proportional to ϵ will vanish on average. Accumulation of the random potential effect will then give non zero contribution of order ϵ^2. In order for this to be on the same order as effects of free propagation on the wave action requires

$$\epsilon = \sqrt{c\mu} \tag{8}$$

with $c > 0$ a finite constant which accounts for the relative importance of the scattering interactions with respect to propagation. The constant c is strictly speaking not needed, and it could be absorbed in the definition of Π, but it is useful for the physical discussion. The kinetic regime, or kinetic limit, is the joint limit $\mu \to 0$ with $\epsilon = \sqrt{c\mu}$, where c is a fixed constant. Consequently, the pertinent equation for the kinetic scaling will be

$$i\mu \frac{\partial \psi^\mu}{\partial t} = -\frac{\mu^2}{2}\nabla_x^2 \psi^\mu + \sqrt{c\mu} V^\mu \psi^\mu. \tag{9}$$

In some parts of the following sections, we will perform asymptotic expansions of the effect of the random potential by expanding (7) in power of ϵ, while integrating out explicitly the wave propagation effects. For this reason, we often consider (7) instead of (9) for those technical parts.

2.2 Local spectral density

In the regime of wave kinetics, one is interested in the amount of wave action existing at each position and wave vector. This is provided by the (rescaled) Wigner distribution of the signal that is defined, following previous work on

inhomogeneous wave kinetics [9, 10], as

$$n^{\mu}(\boldsymbol{x},\boldsymbol{p},t)=\frac{1}{(2\pi\mu)^{d}}\int_{\mathbb{R}^{d}}d\boldsymbol{y}\,e^{-i\boldsymbol{p}\cdot\boldsymbol{y}/\mu}\psi^{\mu}\left(\boldsymbol{x}+\frac{\boldsymbol{y}}{2}\right)\psi^{\mu\dagger}\left(\boldsymbol{x}-\frac{\boldsymbol{y}}{2}\right), \tag{10}$$

where we have denoted by $\psi^{\mu\dagger}$ the complex conjugate of ψ^{μ}. The Wigner distribution function is the local spectral density of the wave action defined in both space of position and wave vector. Indeed, the integral of n^{μ} over wave-vector space coincides with the action density in position space, such that $\int_{\mathbb{R}^{d}}d\boldsymbol{p}\,n^{\mu}(\boldsymbol{x},\boldsymbol{p})=|\psi^{\mu}(\boldsymbol{x})|^{2}$. A caution for the physical interpretation of n^{μ} is that it allows the existence of negative values, in contrast to wave action or energy that should be strictly non-negative. The negativity of n^{μ} can be eliminated by averaging it over a scale comparable to μ [53].

In this study, we will take the asymptotically small limit of both μ and ϵ. Statistically, the limit of $\mu\to 0$ is regarded as a kind of thermodynamic limit; since μ represents the typical correlation length of the wave signal, the total number of degrees of freedom increase as μ^{-d}. In general, the thermodynamic limit makes sense when we specify the macroscopic or mesoscopic variables, e.g., temperature or pressure for gas molecules. In the present model, the potential spectrum $\Pi(\boldsymbol{p})$ is a mesoscopic control parameter that should not depend on μ, and the local spectral density n^{μ} is a mesoscopic variable to be determined. In our formulation, a superscript μ is put on a variable that depends on the scaling parameter. It is worth keeping in mind that, although the spectral density function $\Pi(\boldsymbol{p})$ is fixed, the corresponding potential function $V^{\mu}(\boldsymbol{x})$ depends on μ because its structure becomes finer and finer for $\mu\to 0$.

2.3 Conservation properties

We consider equation (7) on a spatial domain Γ, which is either $\mathbb{R}^{d}$ or a periodic domain $\mathbb{V}$. The norm $\int_{\Gamma}d\boldsymbol{x}|\psi^{\mu}(\boldsymbol{x},t)|^{2}$ is conserved by the dynamics. If Γ is $\mathbb{R}^{d}$, the norm can be either finite for localized solution, or infinite. A local conservation law always exists for $|\psi^{\mu}(\boldsymbol{x},t)|^{2}$.

We now consider the other invariants, for equation (7), or the associated local conservation laws. If not specified, integrations for positions and wave vectors are always performed over Γ and $\mathbb{R}^{d}$, respectively. For a polynomial function $f:\mathbb{R}\to\mathbb{R}$, we define an operator $\mathcal{K}_{f}^{\epsilon,\mu}=f(-(\mu^{2}/2)\nabla_{x}^{2}+\epsilon V^{\mu})$. Since the Schrödinger operator is self-adjoint, a straightforward computation shows that

$$\langle f\rangle\equiv\int d\boldsymbol{x}\psi^{\mu\dagger}(\boldsymbol{x},t)\mathcal{K}_{f}^{\epsilon,\mu}\psi^{\mu}(\boldsymbol{x},t) \tag{11}$$

is independent of the time t when ψ^{μ} is a solution of Eq. (7). Equation (11) is rewritten in terms of the local spectral density as

$$\langle f\rangle=\int d\boldsymbol{x}d\boldsymbol{p}K_{f}^{\epsilon,\mu}(\boldsymbol{x},\boldsymbol{p})n^{\mu}(\boldsymbol{x},\boldsymbol{p},t), \tag{12}$$

where

$$K_f^{\epsilon,\mu}(\boldsymbol{x},\boldsymbol{p}) = \int_{\mathbb{R}^d} d\boldsymbol{y} e^{-i\boldsymbol{p}\cdot\boldsymbol{y}/\mu} \hat{K}_f^{\epsilon,\mu}\left(\boldsymbol{x}+\frac{\boldsymbol{y}}{2}, \boldsymbol{x}-\frac{\boldsymbol{y}}{2}\right)$$
$$\text{and} \quad \hat{K}_f^{\epsilon,\mu}(\boldsymbol{x}_1,\boldsymbol{x}_2) \equiv \mathcal{K}_f^{\epsilon,\mu}\delta(\boldsymbol{x}_1-\boldsymbol{x}_2)$$

is the Weyl symbol of the operator $\mathcal{K}_f^{\epsilon,\mu}$. In general, $K_f^{\epsilon,\mu}$ can be expanded as a power series of ϵ and μ. The leading-order term is immediately obtained as $K_f^{0,0}(\boldsymbol{x},\boldsymbol{p}) = f\left(|\boldsymbol{p}|^2/2\right)$. Therefore, for the small limit of ϵ and μ, we may write the general invariant as

$$\langle f \rangle = \int d\boldsymbol{x} d\boldsymbol{p}\, f\left(\frac{|\boldsymbol{p}|^2}{2}\right) n^\mu(\boldsymbol{x},\boldsymbol{p}). \tag{13}$$

Since the choice of f is arbitrary, the present system possesses an infinite number of invariant. This property is related to wave frequency conservation. For a scattering of wave action in spectral space by a time-independent potential, wave frequency does not change. Therefore, the amount of wave action with frequency less than ω remains constant. Setting $f(\sigma) = h(\omega-\sigma)$ in (13), where h is the Heaviside function and $\omega \in \mathbb{R}^+$, we define

$$\mathcal{A}_\omega[n^\mu] \equiv \int d\boldsymbol{x} d\boldsymbol{p}\, h\left(\omega - \frac{|\boldsymbol{p}|^2}{2}\right) n^\mu(\boldsymbol{x},\boldsymbol{p}). \tag{14}$$

Conservation of $\langle f \rangle$ for an arbitrary f is equivalent to the conservation of $\mathcal{A}_\omega$ for an arbitrary $\omega \in \mathbb{R}^+$, as $\langle f \rangle = \int_{\mathbb{R}^+} d\omega \frac{d\mathcal{A}_\omega}{d\omega} f(\omega)$.

2.4 Wave kinetic equation

In this subsection, we shall derive the classical form of the wave kinetic equation. A common derivation of the wave kinetic equation starts from a closed equation on the Wigner distribution and performs a multiple time-scale expansion, see e.g. [9]. We rather adopt here a perturbative approach in ϵ from the Schrödinger equation (7). This approach is also classical in the wave turbulence literature [2, 54, 55], will appear helpful in the derivation of the dynamical large deviation theory in section 3, and has the advantage to generalize easily to the case of the kinetic theory of non-linear waves with 3-wave interactions.

The first step is to express the solution of the Schrödinger equation (7) using an expansion in power of ϵ, such that $\psi^\mu = \psi_0^\mu + \epsilon\psi_1^\mu + \epsilon^2\psi_2^\mu + \mathcal{O}(\epsilon^3)$. Inserting this expansion to (7), we derive the series of equations:

$$i\mu\frac{\partial\psi_0^\mu}{\partial t} = -\frac{\mu^2}{2}\nabla_x^2\psi_0^\mu$$
$$i\mu\frac{\partial\psi_1^\mu}{\partial t} = -\frac{\mu^2}{2}\nabla_x^2\psi_1^\mu + V^\mu\psi_0^\mu$$

$$i\mu\frac{\partial\psi_2^\mu}{\partial t} = -\frac{\mu^2}{2}\nabla_x^2\psi_2^\mu + V^\mu\psi_1^\mu$$
$$\vdots$$

We consider this expansion for any given initial condition $\psi^\mu(\boldsymbol{x},0) = \psi^{\mu,0}(\boldsymbol{x})$. We assume that the initial condition does not depend on ϵ, i.e., $\psi_0^\mu(\boldsymbol{x},0) = \psi^{\mu,0}(\boldsymbol{x})$ and $\psi_j^\mu(\boldsymbol{x},0) = 0$ for $j \geq 1$. To write down the solution, we introduce the propagator $G^\mu(\boldsymbol{x},t)$ such that

$$\left(i\mu\frac{\partial}{\partial t} + \frac{\mu^2}{2}\nabla_x^2\right)G^\mu = i\mu\delta(\boldsymbol{x})\delta(t), \tag{15}$$

and $G^\mu = 0$ for $t < 0$. We then obtain

$$\psi_0^\mu(\boldsymbol{x},t) = \int_{\mathbb{R}^d} d\boldsymbol{\xi} G^\mu(\boldsymbol{x}-\boldsymbol{\xi},t)\psi^\mu(\boldsymbol{\xi},0) \tag{16a}$$

$$\begin{aligned}\psi_1^\mu(\boldsymbol{x},t) &= \frac{1}{i\mu}\int_0^t d\tau\int_{\mathbb{R}^d} d\boldsymbol{\xi} G^\mu(\boldsymbol{x}-\boldsymbol{\xi},t-\tau)V^\mu(\boldsymbol{\xi})\psi_0^\mu(\boldsymbol{\xi},\tau)\\ &= \frac{1}{i\mu}\int_0^t d\tau\int_{\mathbb{R}^{2d}} d\boldsymbol{\xi}_{12} G^\mu(\boldsymbol{x}-\boldsymbol{\xi}_1,t-\tau)V^\mu(\boldsymbol{\xi}_1)G^\mu(\boldsymbol{\xi}_1-\boldsymbol{\xi}_2,\tau)\psi^\mu(\boldsymbol{\xi}_2,0)\end{aligned} \tag{16b}$$

$$\begin{aligned}\psi_2^\mu(\boldsymbol{x},t) &= \frac{1}{i\mu}\int_0^t d\tau\int_{\mathbb{R}^d} d\boldsymbol{\xi} G^\mu(\boldsymbol{x}-\boldsymbol{\xi},t-\tau)V^\mu(\boldsymbol{\xi})\psi_1^\mu(\boldsymbol{\xi},\tau)\\ &= \frac{1}{-\mu^2}\int_0^t d\tau_1\int_0^{\tau_1} d\tau_2\int_{\mathbb{R}^{3d}} d\boldsymbol{\xi}_{123} G^\mu(\boldsymbol{x}-\boldsymbol{\xi}_1,t-\tau_1)V^\mu(\boldsymbol{\xi}_1)\\ &\times G^\mu(\boldsymbol{\xi}_1-\boldsymbol{\xi}_2,\tau_1-\tau_2)V^\mu(\boldsymbol{\xi}_2)G^\mu(\boldsymbol{\xi}_2-\boldsymbol{\xi}_3,\tau_2)\psi^\mu(\boldsymbol{\xi}_3,0).\end{aligned} \tag{16c}$$

The equation for the propagator (15) is analytically solved as

$$G^\mu(\boldsymbol{x},t) = \frac{h(t)}{(2\pi\mu)^d}\int d\boldsymbol{p} e^{-i|\boldsymbol{p}|^2 t/2\mu}e^{i\boldsymbol{p}\cdot\boldsymbol{x}/\mu}, \tag{17}$$

where $h(t)$ is again the Heaviside function. Although the wave vector integration of this expression can be carried out, we keep this form because it is convenient for later computations.

Importantly, the perturbation solution will be valid for not too large t. For longer times, the higher order terms will not be small compared to ψ_0^μ even though ϵ is small. As will be clear in the following discussion, we will need the perturbative solution to be valid up to $\mu \ll t \ll 1$; an intermediate range between the microscopic and mesoscopic time scales.

Based on the perturbation solution of ψ^μ derived above, we shall consider the evolution of the local spectral density n^μ. To simplify the computation, we

introduce the Wigner transform of two functions, $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$, as

$$w^{\mu}(f,g)(\boldsymbol{x},\boldsymbol{p}) \equiv \frac{1}{(2\pi\mu)^d}\int_{\mathbb{R}^d} d\boldsymbol{y}\, f\left(\boldsymbol{x}+\frac{\boldsymbol{y}}{2}\right) g^{\dagger}\left(\boldsymbol{x}-\frac{\boldsymbol{y}}{2}\right) e^{-i\boldsymbol{p}\cdot\boldsymbol{y}/\mu}. \tag{18}$$

Its inverse is

$$f(\boldsymbol{x}_1)g^{\dagger}(\boldsymbol{x}_2) = \int d\boldsymbol{p}\, w^{\mu}\left(\frac{\boldsymbol{x}_1+\boldsymbol{x}_2}{2},\boldsymbol{p}\right) e^{i\boldsymbol{p}\cdot(\boldsymbol{x}_1-\boldsymbol{x}_2)/\mu}. \tag{19}$$

Basically, the local spectral density is the Wigner transform of identical wave functions,

$$n^{\mu}(\boldsymbol{x},\boldsymbol{p},t) = w^{\mu}(\psi^{\mu},\psi^{\mu}). \tag{20}$$

Inserting $\psi^{\mu} = \psi_0^{\mu} + \epsilon\psi_1^{\mu} + \epsilon^2\psi_2^{\mu} + \dots$ to (20), and using (16), we obtain an expansion of the local spectral density n^{μ} in terms of ϵ as

$$\begin{aligned} n^{\mu}(\boldsymbol{x},\boldsymbol{p},t) = {} & w^{\mu}(\psi_0^{\mu},\psi_0^{\mu}) + \epsilon w^{\mu}(\psi_1^{\mu},\psi_0^{\mu}) + \epsilon w^{\mu}(\psi_0^{\mu},\psi_1^{\mu}) \\ & + \epsilon^2 w^{\mu}(\psi_1^{\mu},\psi_1^{\mu}) + \epsilon^2 w^{\mu}(\psi_2^{\mu},\psi_0^{\mu}) + \epsilon^2 w^{\mu}(\psi_0^{\mu},\psi_2^{\mu}) + \mathcal{O}(\epsilon^3). \end{aligned} \tag{21}$$

The first term on the right-hand side is easily computed as

$$w^{\mu}(\psi_0^{\mu},\psi_0^{\mu}) = n^{\mu}(\boldsymbol{x}-\boldsymbol{p}t,\boldsymbol{p},0). \tag{22}$$

This expression means that, in absence of the potential, the propagation of free waves transports the spatial distribution of wave action density at the group velocity $\boldsymbol{p}$. It is notable that Eq. (22) is valid without taking an ensemble average or an asymptotic limit.

We shall consider the expectation of (21) with respect to the realization of the random potential V^{μ}. Since $\mathbb{E}[V^{\mu}] = 0$ has been assumed, terms proportional to ϵ vanish. The dominant contribution from the random potential to the variations in local spectral density comes from terms of order ϵ^2. Direct computations, performed in Appendix C, yield the expectation values of perturbation terms at the leading-order, (114), (115) and (116). Consequently, we obtain

$$\begin{aligned} & \mathbb{E}\left[\epsilon^2 w^{\mu}(\psi_1^{\mu},\psi_1^{\mu}) + \epsilon^2 w^{\mu}(\psi_2^{\mu},\psi_0^{\mu}) + \epsilon^2 w^{\mu}(\psi_0^{\mu},\psi_2^{\mu})\right] \\ = & \frac{\epsilon^2 t}{\mu}\int d\boldsymbol{\eta}\sigma(\boldsymbol{p},\boldsymbol{\eta})\left(n^{\mu}(\boldsymbol{x},\boldsymbol{\eta},0) - n^{\mu}(\boldsymbol{x},\boldsymbol{p},0)\right) + \text{h.o.t.} \end{aligned}$$

with

$$\sigma(\boldsymbol{p}_1,\boldsymbol{p}_2) \equiv 2\pi\Pi(\boldsymbol{p}_1-\boldsymbol{p}_2)\delta\left(\frac{|\boldsymbol{p}_1|^2}{2} - \frac{|\boldsymbol{p}_2|^2}{2}\right). \tag{23}$$

Here, h.o.t. represents the higher-order terms in the expansion that are negligible in the asymptotic limit. Setting $\epsilon = \sqrt{c\mu}$, we have

$$\lim_{\mu\to 0} \frac{\mathbb{E}[n^\mu(\boldsymbol{x},\boldsymbol{p},t)] - n^\mu(\boldsymbol{x}-\boldsymbol{p}t,\boldsymbol{p},0)}{t} = \lim_{\mu\to 0} c\int d\boldsymbol{\eta}\sigma(\boldsymbol{p},\boldsymbol{\eta})\left(n^\mu(\boldsymbol{x},\boldsymbol{\eta},0) - n^\mu(\boldsymbol{x},\boldsymbol{p},0)\right).$$

Taking the limit $t \to 0$, we obtain a differential equation for the ensemble average of the local spectral density, $\lim_{\mu\to 0}\mathbb{E}[n^\mu] = n$:

$$\frac{\partial n(\boldsymbol{x},\boldsymbol{p},0)}{\partial t} + \boldsymbol{p}\cdot\nabla_x n(\boldsymbol{x},\boldsymbol{p},0) = c\int d\boldsymbol{\eta}\sigma(\boldsymbol{p},\boldsymbol{\eta})\left(n(\boldsymbol{x},\boldsymbol{\eta},0) - n(\boldsymbol{x},\boldsymbol{p},0)\right).$$

At this stage, this expression is only valid at the initial time, $t = 0$. One cannot extend it to $t > 0$ because the wave function ψ^μ is *a priori* correlated with the potential field V^μ. However, we shall argue that the correlation between ψ^μ and V^μ remains always weak at any time. Indeed, in the present problem, it is assumed that the significant modification of the wave field ψ^μ by scattering on the potential occurs at a time scale of wave propagation over a mesoscopic distance. The typical correlation length of the random potential V^μ is much shorter than this mesoscopic scale by a factor of μ. Therefore, even though interferences of the field with the potential produce slight correlations, the free propagation of the field makes the correlation vanishes. This situation resembles the loss of memory for particle collision in dilute gas—the molecular chaos hypothesis. The present weak-correlation assumption allows us to regard the temporal evolution of n^μ as a Markovian process such that the wave kinetic equation would be valid any time as far as μ is sufficiently small. We note that this explanation applies to $d \geq 2$ cases. For one-dimensional problems, an interesting phenomenon named localization is known to occur [56, 57]. This localization phenomenon which invalidates the present kinetic regime [50, Sec. 5.2] is thus discarded in this paper.

Once this Markovian hypothesis is accepted, we may iterate the reasoning expounded above for $t = 0$ in order to reach any time $t > 0$. We eventually obtain the equation

$$\frac{\partial n(\boldsymbol{x},\boldsymbol{p},t)}{\partial t} + \boldsymbol{p}\cdot\nabla_x n(\boldsymbol{x},\boldsymbol{p},t) = c\int d\boldsymbol{\eta}\sigma(\boldsymbol{p},\boldsymbol{\eta})\left(n(\boldsymbol{x},\boldsymbol{\eta},t) - n(\boldsymbol{x},\boldsymbol{p},t)\right), \quad (24)$$

that is the ordinary form of the wave kinetic equation. Since the group velocity of a free wave is now $\boldsymbol{p}$, the second term on the left-hand side is understood as the motion of the Wigner distribution at the group velocity due to the free dynamics. The right-hand side represents wave scattering in wave-vector space that occurs at microscopic scale. The function $c\sigma(\boldsymbol{p}_1,\boldsymbol{p}_2)$ is the scattering cross section determining the rate of wave action converted from wave vector $\boldsymbol{p}_1$ to $\boldsymbol{p}_2$ per unit time.

The wave kinetic equation inherits the conservation property of the original Schrödinger equation, namely that $d\mathcal{A}_\omega[n]/dt = 0$. However, the wave kinetic equation differs from the microscopic Schrödinger dynamics (7) since the former appears time-irreversible whereas the latter is time reversible. Indeed, for a prescribed potential field $V^\mu(\boldsymbol{x})$, let $\psi^\mu(\boldsymbol{x}, t)$ be a solution of (7). Its time-reverse counterpart is defined as $\psi^\mu_R(\boldsymbol{x}, t) = \psi^\dagger(\boldsymbol{x}, -t)$. As a consequence, the time-reversed local spectral density reads as $n^\mu_R(\boldsymbol{x}, \boldsymbol{p}, t) = n^\mu(\boldsymbol{x}, -\boldsymbol{p}, -t)$. Change in sign of $\boldsymbol{p}$ is a natural outcome because the wave group velocities of the forward and the reverse paths should be opposite. If $\psi^\mu(\boldsymbol{x}, t)$ is a solution of the Schrödinger equation, $\psi^\mu_R(\boldsymbol{x}, t)$ is also a solution. This is the time-reversal symmetry. By contrast, the wave kinetic equation violates this symmetry: if n is a solution, the time-reversed n_R is not a solution of the wave kinetic equation, unless both n and n_R are an identical stationary state (see Eq. (26) defined below). This irreversibility paradox is reminiscent to the one raised by Boltzmann for the case of dilute gases. As explained in [1] for the case of the Boltzmann equation, we will see in the next section that one can recover time-reversibility for the kinetic theory at the large deviation level, as a time-reversibility for the stochastic process of the local spectral density. Time-irreversibility arises because the wave kinetic equation describes the evolution of the *average* $n = \mathbb{E}[n^\mu]$ only, or equivalently in this case the most probable evolution only. Other paths, including time-reversed paths, are possible: they are just extremely unlikely.

Before moving toward the large deviation theory of wave kinetics, we can remark that time-irreversibility can be also quantified by introducing a Lyapunov function

$$S \equiv \int d\boldsymbol{x} d\boldsymbol{p} \log n(\boldsymbol{x}, \boldsymbol{p}). \tag{25}$$

Following solutions of the wave kinetic equation, S increases monotonically with time, $dS/dt \geq 0$. If the spatial domain Γ is finite, S achieves the maximum when the spectral density $n(\boldsymbol{x}, \boldsymbol{p})$ is homogeneous in $\boldsymbol{x}$ and isotropic in $\boldsymbol{p}$. We write this homogeneous distribution of the spectral density under the constraint of $\mathcal{A}_\omega[n] = A(\omega)$ as

$$n^A_h(\boldsymbol{p}) = \frac{A'(|\boldsymbol{p}|^2/2)}{\int d\boldsymbol{x} d\boldsymbol{\eta} \delta(|\boldsymbol{p}|^2/2 - |\boldsymbol{\eta}|^2/2)}, \tag{26}$$

where the denominator is introduced for normalization purpose, and $A' = dA/d\omega$. It is obvious that $n(\boldsymbol{x}, \boldsymbol{p}) = n^A_h(\boldsymbol{p})$ is a stationary solution of the wave kinetic equation. We note that the Lyapunov function S is actually the microcanonical entropy of the macrostate specified by $(n^\mu, \mathcal{A}_\omega) = (n, A)$ for the Schrödinger equation as we discuss in Appendix B, and is related to the quasipotential that appears in the discussion of the large deviation theory, as we discuss in section 3.2.3.

3 Large deviation formulation for random wave scattering

In the previous section, we derived the wave kinetic equation as an equation for the ensemble average $\mathbb{E}[n^\mu]$ of the empirical local spectral density, with respect to the probability measure of the random potential, in the limit $\mu \to 0$. This kinetic limit can be understood as a law of large number: $\lim_{\mu\to 0} n^\mu = \mathbb{E}[n^\mu]$, where the limit has to be understood in a weak sense (for instance, the limit holds when both n^μ and $\mathbb{E}[n^\mu]$ are integrated over any subset $\mathbb{U} \subset \Gamma \times \mathbb{R}^d$).

In statistical mechanics, it is quite common that an empirical macroscopic quantity converges to its ensemble average for the large limit of the number of elements. The deviation from the average is often exponentially small and evaluated asymptotically by a large deviation principle. In this section, our aim is to generalize the law of large number for the kinetic theory and to compute the probability to observe any possible fluctuations for n^μ, as a large deviation principle, in the limit $\mu \to 0$. Such fluctuations are expected to be characterized by a large deviation parameter proportional to μ^d, where d is the space dimension, because the number of statistically independent degrees of freedom is of order μ^{-d}.

3.1 Large deviation Hamiltonian

We define the Newton ratio as the time increment for the local spectral density: $\Delta n^\mu/\Delta t = (n^\mu(\Delta t) - n^\mu(0))/\Delta t$. We regard the temporal variations in n^μ as a stochastic process, and look for the probability to observe a value of the Newton ratio, conditioned on the value of the local spectral density at time 0: $n^\mu(0) = n$. Our aim is to justify that it satisfies

$$-\lim_{\Delta t\to 0}\lim_{\substack{\mu\to 0\\ \epsilon=\sqrt{c\mu}}} \frac{(2\pi\mu)^d}{\Delta t}\log \mathbb{P}\left[\tfrac{\Delta n^\mu}{\Delta t} = \dot{n}|n^\mu = n\right] = \mathcal{L}[n,\dot{n}], \tag{27}$$

and to derive an explicit formula for the Lagrangian $\mathcal{L}$. The limit $\lim_{\substack{\mu\to 0\\ \epsilon=\sqrt{c\mu}}}$ corresponds to the kinetic limit $\mu \to 0$ where one has fixed $\epsilon = \sqrt{c\mu}$. Here, $\mathcal{L}[n,\dot{n}]$ is the rate function of the probability of the Newton ratio $\mathbb{P}[\dot{n}|n]$.

Through the fast microscopic dynamics, the memory of the initial condition of the phases of ψ^μ are expected to be lost after some times, meaning that two-times, or multi-times, correlation functions are expected to decay with the time differences of two or several phase observables. Such a mixing property is expected to be due to the conjunction of phase mixing (oscillating integrals and the Riemann–Lebesgue lemma), spatial transport and dispersion, and the effect of the random potential. Because the natural timescale for phase dynamics is the microscopic timescale, one might expect a typical mixing time to be much smaller than the kinetic time scale and to decay to zero in the kinetic limit. This mixing property is required to justify a Markovian behavior and to propagate local in time results, like the Lagrangian (27), in order to describe

the dynamics at finite times. In none of the existing classical kinetic theories, neither physicists nor mathematicians have been so far able to justify or prove the requested mixing properties, in order to justify the long time validity of kinetic theories or their probabilistic large deviation generalizations. This is the main reason why mathematical results, for instance the celebrated Lanford's result for the Boltzmann equation [58], or its generalizations [59], are usually valid only for a fraction of the kinetic time. As it is customary in theoretical physics, we will assume the validity of such a mixing property in the following.

We now assume the natural mixing hypothesis and the related Markov behavior of the stochastic process. As a consequence, the evolution of n^μ does not depend on its previous state. Then the path probability for the stochastic process is Markovian, and the probability of a path of n^μ for a finite time interval, $0 \leqslant t \leqslant t_f$, can be derived from the local in time Lagrangian (27). The path probability conditioned on the local spectral density at the initial time $n^\mu(0) = n(0)$ is then

$$\mathbb{P}_{n(0)}\left[\{n^\mu(t) = n(t)\}_{0\leqslant t\leqslant t_f}\right] \underset{\mu\to 0}{\asymp} \exp\left(-\frac{1}{(2\pi\mu)^d}\int_0^{t_f} dt\mathcal{L}[n,\dot{n}]\right). \tag{28}$$

This expression is analogous to the path-integral formulation for the probability density in quantum theory. Note that the initial conditions $n^\mu(0) = n(0)$ is fixed here, but one can easily consider a set of initial condition complemented by an initial probability density $\mathbb{P}_0$. In such a case, the path probability of a trajectory $\{n(t)\}_{0\leqslant t\leqslant t_f}$ reads as

$$\mathbb{P}\left[\{n^\mu(t) = n(t)\}_{0\leqslant t\leqslant t_f}\right] \underset{\mu\to 0}{\asymp} \exp\left(-\frac{1}{(2\pi\mu)^d}\int_0^{t_f} dt\mathcal{L}[n,\dot{n}]\right)\mathbb{P}_0[n(0)]. \tag{29}$$

To compute the Lagrangian (27), we will use the Gärtner-Ellis theorem that connects the rate function $\mathcal{L}$ to the cumulant-generating function, or the Hamiltonian $\mathcal{H}$ defined by

$$\begin{aligned}\mathcal{H}[n,\lambda] = &\lim_{\Delta t\to 0}\lim_{\substack{\mu\to 0\\ \epsilon=\sqrt{c\mu}}}\frac{(2\pi\mu)^d}{\Delta t}\\ &\times\log\mathbb{E}\left[\exp\left(\frac{\int d\boldsymbol{x}d\boldsymbol{p}\lambda(\boldsymbol{x},\boldsymbol{p})(n^\mu(\boldsymbol{x},\boldsymbol{p},\Delta t)-n(\boldsymbol{x},\boldsymbol{p}))}{(2\pi\mu)^d}\right)\right],\end{aligned} \tag{30}$$

through the Legendre-Fenchel transform,

$$\mathcal{L}[n,\dot{n}] \equiv \sup_\lambda\left\{\int d\boldsymbol{x}d\boldsymbol{p}\lambda(\boldsymbol{x},\boldsymbol{p})\dot{n}(\boldsymbol{x},\boldsymbol{p}) - \mathcal{H}[n,\lambda]\right\}. \tag{31}$$

Our aim here is to derive the specific form of $\mathcal{H}$ directly from the perturbation solutions of the original Schrödinger equation (7).

Following [2], we first compute the moment-generating function for the increment of n^μ,

$$\mathcal{Z}^\mu[n,\lambda;\ \Delta t] \equiv \mathbb{E}\left[\exp\left(\frac{\int d\boldsymbol{x}d\boldsymbol{p}\lambda(\boldsymbol{x},\boldsymbol{p})(n^\mu(\boldsymbol{x},\boldsymbol{p},\Delta t)-n(\boldsymbol{x},\boldsymbol{p}))}{(2\pi\mu)^d}\right)\right]. \tag{32}$$

We insert the expansion of $n^\mu(\boldsymbol{x},\boldsymbol{p},t)$ in terms of ϵ into (32). Since at the leading order $w^\mu(\psi_0^\mu,\psi_0^\mu)$ is statistically independent of V^μ by the weak correlation hypothesis described before, it is possible to decompose $\mathcal{Z}^\mu$ into two parts as

$$\mathcal{Z}^\mu = \mathcal{Z}_0^\mu \mathcal{Z}_\epsilon^\mu \tag{33a}$$

$$\mathcal{Z}_0^\mu \equiv \exp\left(\frac{1}{(2\pi\mu)^d}\int d\boldsymbol{x}d\boldsymbol{p}\lambda(\boldsymbol{x},\boldsymbol{p})\left(w^\mu(\psi_0^\mu,\psi_0^\mu)-n(\boldsymbol{x},\boldsymbol{p})\right)\right) \tag{33b}$$

$$\mathcal{Z}_\epsilon^\mu \equiv \mathbb{E}\left[\exp\left(\frac{1}{(2\pi\mu)^d}\int d\boldsymbol{x}d\boldsymbol{p}\lambda(\boldsymbol{x},\boldsymbol{p})\left(n^\mu(\boldsymbol{x},\boldsymbol{p},\Delta t)-w^\mu(\psi_0^\mu,\psi_0^\mu)\right)\right)\right]. \tag{33c}$$

Using (22), the first part, $\mathcal{Z}_0^\mu$, is immediately rewritten as

$$\mathcal{Z}_0^\mu = \exp\left(\frac{1}{(2\pi\mu)^d}\int d\boldsymbol{x}d\boldsymbol{p}\lambda(\boldsymbol{x},\boldsymbol{p})\left(n(\boldsymbol{x}-\boldsymbol{p}\Delta t,\boldsymbol{p},0))-n(\boldsymbol{x},\boldsymbol{p})\right)\right). \tag{34}$$

The second part is expanded in terms of ϵ to obtain

$$\mathcal{Z}_\epsilon^\mu = 1+\epsilon^2\mathcal{Z}_2^\mu + o(\epsilon^2), \tag{35}$$

where we have used $\mathbb{E}[V^\mu]=0$. The Landau notation $o(\epsilon^2)$ gathers all the terms that are negligible compared to $\mathcal{O}(\epsilon^2)$ terms in the expansion. In the following, we shall discard the higher order terms of $o(\epsilon^2)$ and concentrate on computing $\mathcal{Z}_2^\mu$. Because we are considering the simultaneous limit of μ, $\epsilon=\sqrt{c\mu}$, neglecting $o(\epsilon^2)$ terms cannot be justified *a priori*. This kind of problem commonly arises in wave kinetic theory [2] but we expect $o(\epsilon^2)$ to be negligible in the kinetic limit. The direct computation of $\mathcal{Z}_2^\mu$ yields

$$\begin{aligned}\mathcal{Z}_2^\mu = &\frac{1}{(2\pi\mu)^d}\int d\boldsymbol{x}d\boldsymbol{p}\lambda(\boldsymbol{x},\boldsymbol{p})\mathbb{E}\left[w^\mu(\psi_2^\mu,\psi_0^\mu)+w^\mu(\psi_0^\mu,\psi_2^\mu)+w^\mu(\psi_1^\mu,\psi_1^\mu)\right]\\ &+\frac{1}{2(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\\ &\times\mathbb{E}\big[(w^\mu(\psi_1^\mu,\psi_0^\mu)(\boldsymbol{x}_1,\boldsymbol{p}_1)+w^\mu(\psi_0^\mu,\psi_1^\mu)(\boldsymbol{x}_1,\boldsymbol{p}_1))\\ &\times(w^\mu(\psi_1^\mu,\psi_0^\mu)(\boldsymbol{x}_2,\boldsymbol{p}_2)+w^\mu(\psi_0^\mu,\psi_1^\mu)(\boldsymbol{x}_2,\boldsymbol{p}_2))\big].\end{aligned} \tag{36}$$

From the computations in Appendix C, specifically (114)-(116) and (118)-(121), we obtain

$$\mathcal{Z}_2^\mu = \frac{\Delta t}{\mu(2\pi\mu)^d}\int d\boldsymbol{x}d\boldsymbol{p}_{12}\Big[(\lambda(\boldsymbol{x},\boldsymbol{p}_1)-\lambda(\boldsymbol{x},\boldsymbol{p}_2))\sigma(\boldsymbol{p}_1,\boldsymbol{p}_2)n(\boldsymbol{x},\boldsymbol{p}_2) + \frac{1}{2}(\lambda(\boldsymbol{x},\boldsymbol{p}_1)-\lambda(\boldsymbol{x},\boldsymbol{p}_2))^2\sigma(\boldsymbol{p}_1,\boldsymbol{p}_2)n(\boldsymbol{x},\boldsymbol{p}_1)n(\boldsymbol{x},\boldsymbol{p}_2)\Big] + \text{h.o.t..} \tag{37}$$

Inserting $\mathcal{Z}^\mu = \mathcal{Z}_0^\mu\left(1+\epsilon^2\mathcal{Z}_2^\mu\right)$ into $\mathcal{H} = \lim_{\Delta t\to 0}\lim_{\substack{\mu\to 0\\ \epsilon=\sqrt{c\mu}}}((2\pi\mu)^d/\Delta t)\log\mathcal{Z}^\mu$, we obtain the Hamiltonian as

$$\mathcal{H}[n,\lambda] = \mathcal{H}_F + \mathcal{H}_S \tag{38a}$$

$$\mathcal{H}_F = -\int d\boldsymbol{x}d\boldsymbol{p}\lambda(\boldsymbol{x},\boldsymbol{p})\boldsymbol{p}\cdot\nabla_x n(\boldsymbol{x},\boldsymbol{p}) \tag{38b}$$

$$\mathcal{H}_S = c\int d\boldsymbol{x}d\boldsymbol{p}_{12}\Big[(\lambda(\boldsymbol{x},\boldsymbol{p}_1)-\lambda(\boldsymbol{x},\boldsymbol{p}_2))\sigma(\boldsymbol{p}_1,\boldsymbol{p}_2)n(\boldsymbol{x},\boldsymbol{p}_2) + \frac{1}{2}(\lambda(\boldsymbol{x},\boldsymbol{p}_1)-\lambda(\boldsymbol{x},\boldsymbol{p}_2))^2\sigma(\boldsymbol{p}_1,\boldsymbol{p}_2)n(\boldsymbol{x},\boldsymbol{p}_1)n(\boldsymbol{x},\boldsymbol{p}_2)\Big]. \tag{38c}$$

We have separated the Hamiltonian into two parts, $\mathcal{H}_F$ and $\mathcal{H}_S$. The first part, $\mathcal{H}_F$, represents the free wave propagation in position space and the second one, $\mathcal{H}_S$, the wave scattering in wave-vector space by the random potential.

3.2 Properties of the large deviation Hamiltonian

Once the specific form of the Hamiltonian is obtained, we can discuss the properties of the stochastic process governed by the path-integral formula (28) and (31). Formulations in the paper [1] are simple and informative, and we summarize several important properties of dynamical large deviation theory in Appendix A. We now check the classical expected properties of $\mathcal{H}$.

3.2.1 Weak noise Langevin dynamics and wave kinetic equation

The first important property of the large deviation Hamiltonian $\mathcal{H}$ is that it is quadratic and convex with respect to the conjugated field λ. This means that the fluctuations of the infinitesimal current $\dot{n}dt$ are, locally in time, Gaussian. Reading the quadratic part of the Hamiltonian (38c), we see that the local covariance of the local in time Gaussian fluctuations are given by the diffusion kernel

$$\Sigma[n](\boldsymbol{x};\ \boldsymbol{p}_1,\boldsymbol{p}_2) = -c\sigma(\boldsymbol{p}_1,\boldsymbol{p}_2)n(\boldsymbol{x},\boldsymbol{p}_1)n(\boldsymbol{x},\boldsymbol{p}_2) + c\int_{\mathbb{R}^d} d\boldsymbol{\eta}\sigma(\boldsymbol{p}_1,\boldsymbol{\eta})n(\boldsymbol{x},\boldsymbol{p}_1)n(\boldsymbol{x},\boldsymbol{\eta})\delta(\boldsymbol{p}_1-\boldsymbol{p}_2). \tag{39}$$

As a consequence, in the kinetic limit $\mu \ll 1$, the dynamics of the local spectral density in the kinetic regime corresponds to a weak noise Langevin dynamics [60, 61]

$$\dot{n}(\boldsymbol{x},\boldsymbol{p},t) = -\boldsymbol{p}\cdot\nabla_x n + c\int d\boldsymbol{\eta}\,\sigma(\boldsymbol{p},\boldsymbol{\eta})\,(n(\boldsymbol{x},\boldsymbol{\eta},t) - n(\boldsymbol{x},\boldsymbol{p},t)) + \sqrt{2}(2\pi\mu)^{d/2}\int d\boldsymbol{\eta}\,\Sigma^{1/2}[n](\boldsymbol{x};\ \boldsymbol{p},\boldsymbol{\eta})\xi(\boldsymbol{\eta},t) \tag{40}$$

with $\xi(\boldsymbol{p},t)$ a white noise such that $\mathbb{E}\left[\xi(\boldsymbol{p},t)\xi(\boldsymbol{\eta},s)\right] = \delta(t-s)\delta(\boldsymbol{p}-\boldsymbol{\eta})$, and where the kernel $\Sigma^{1/2}[n]$ is defined as a square root of diffusion kernel, meaning that $\int d\boldsymbol{\eta}\Sigma^{1/2}(\boldsymbol{x},\boldsymbol{p}_1,\boldsymbol{\eta})\Sigma^{1/2}(\boldsymbol{x},\boldsymbol{\eta},\boldsymbol{p}_2) = \Sigma(\boldsymbol{x},\boldsymbol{p}_1,\boldsymbol{p}_2)$.

The second expected property is that the most probable path is the solution of the wave kinetic equation. This is easily checked by noticing that the most probable path, for which the action $\int_0^{t_f}\mathcal{L}[n,\dot{n}]dt = 0$ vanishes, satisfies

$$\frac{\partial n}{\partial t} = \left.\frac{\delta\mathcal{H}}{\delta\lambda}\right|_{\lambda=0} \tag{41}$$

(see Appendix A.1.1), and is also the linear term in λ of the Hamiltonian $\mathcal{H}$ and the drift term of the Langevin dynamics (40). Then, the path large deviation analysis confirms that the wave kinetic equation can be understood as a law of large number at the level of trajectories.

It is enlightening to rewrite the scattering part of the Hamiltonian $\mathcal{H}_S$ in terms of the diffusion kernel (39) as

$$\mathcal{H}_S = \int d\boldsymbol{x}d\boldsymbol{p}_1 d\boldsymbol{p}_2\,\lambda(\boldsymbol{x},\boldsymbol{p}_1)\Sigma[n](\boldsymbol{x};\ \boldsymbol{p}_1,\boldsymbol{p}_2)\left(\frac{\delta S}{\delta n(\boldsymbol{x},\boldsymbol{p}_2)} + \lambda(\boldsymbol{x},\boldsymbol{p}_2)\right), \tag{42}$$

where S is the entropy (25), and the Langevin equation (40) as

$$\dot{n}(\boldsymbol{x},\boldsymbol{p},t) + \boldsymbol{p}\cdot\nabla_x n = \int d\boldsymbol{\eta}\,\Sigma[n](\boldsymbol{x};\ \boldsymbol{p},\boldsymbol{\eta})\frac{\delta S}{\delta n(\boldsymbol{x},\boldsymbol{\eta})} + \sqrt{2}(2\pi\mu)^{d/2}\int d\boldsymbol{\eta}\,\Sigma^{1/2}[n](\boldsymbol{x};\ \boldsymbol{p},\boldsymbol{\eta})\xi(\boldsymbol{\eta},t). \tag{43}$$

This suggestive forms immediately emphasize that S is a Lyapunov function, and that the dynamics has a detailed balance structure, as further explained in section 3.2.4.

3.2.2 Conservation of the wave action distribution

As we discussed in section 2.3, in the original Schrödinger equation, the quantity $\mathcal{A}_\omega[n^\mu] = \int d\boldsymbol{x}d\boldsymbol{p}h(\omega - |\boldsymbol{p}|^2/2)n^\mu(\boldsymbol{x},\boldsymbol{p})$ for any $\omega \in \mathbb{R}^+$ is conserved. This conservation property has to be also verified at the large deviation level, meaning that any trajectory $\{n(t)\}_{0\leqslant t\leqslant t_f}$ has to lie on the manifold

$\mathcal{A}_\omega[n] = A(\omega)$. As explained in Appendix A.2.1, this is equivalent to the Hamiltonian symmetry

$$\mathcal{H}\left[n, \lambda + \alpha \frac{\delta \mathcal{A}_\omega}{\delta n}\right] = \mathcal{H}\left[n, \lambda\right], \tag{44}$$

for an arbitrary $\alpha \in \mathbb{R}$, n and λ. For the specific form of the Hamiltonian $\mathcal{H} = \mathcal{H}_F + \mathcal{H}_S$ (38a) with $\mathcal{H}_S$ written in the symmetric form (42), one can directly check the above Hamiltonian symmetry boils down to the following property on the diffusive kernel (39)

$$\int d\boldsymbol{\eta}\, \Sigma[n](\boldsymbol{x};\, \boldsymbol{p}, \boldsymbol{\eta}) \frac{\delta \mathcal{A}_\omega}{\delta n(\boldsymbol{x}, \boldsymbol{\eta})} = 0, \tag{45}$$

for any n, $\boldsymbol{x}$, $\boldsymbol{p}$, and ω.

3.2.3 Stationary quasipotential

It was shown in the previous section that the wave kinetic equation possesses an attractor which is a homogeneous distribution, n_h, with the prescribed constraints $\mathcal{A}_\omega[n] = A(\omega)$, for any ω. We now consider the fluctuations of n from n_h at the large deviation level. More precisely, we define the equilibrium distribution of the stochastic process at a large deviation level:

$$\mathbb{P}^\mu_{A,S}[\{n^\mu = n\}] \underset{\mu \to 0}{\asymp} \exp\left(-\frac{\mathcal{U}_A[n]}{(2\pi\mu)^d}\right), \tag{46}$$

where $\mathbb{P}^\mu_{A,S}$ is the stationary probability measure of the microcanonical ensemble which is parameterized by a small constant μ as well as a function $A(\omega)$ specifying the action conservation constraint. The rate function $\mathcal{U}_A$ is named the quasipotential (see Appendix A.1).

In principle, the quasipotential can be computed from the dynamics, starting from the large deviation Hamiltonian, see for instance Appendix A.1.3. Formula (69) gives an expression for the quasipotential, in the cases when the wave kinetic equation has a single attractor. However, when one knows explicitly the microscopic stationary distribution, for instance in the case of equilibrium statistical mechanics, one can compute directly the quasipotential. It is then related to the entropy. Those different expressions have to give consistent results. Indeed, one may see a simple example of the relation between microcanonical entropy and the quasipotential for the dilute gas dynamics in [1]. The case of spatially homogeneous weakly nonlinear wave dynamics has been discussed in [2]. For the present problem, we compute the quasipotential from a microcanonical ensemble for the original Schrödinger equation model

in Appendix B. One obtains

$$\mathcal{U}_A[n] = \begin{cases} -\int d\boldsymbol{x}d\boldsymbol{p} \log\left(\dfrac{n(\boldsymbol{x},\boldsymbol{p})}{n_h^A(\boldsymbol{p})}\right) & \text{if} \quad \mathcal{A}_\omega[n] = A(\omega) \\ +\infty & \text{otherwise} \end{cases} . \tag{47}$$

This result is technically not obvious. As far as we know, it had never been derived before.

As shown in Appendix A.1.4, it is a generic property of the quasipotential to play the role of a Lyapunov function for the deterministic relaxation dynamics, in this case the wave kinetic equation. Such a property can be derived generically from the existence of a large deviation principle, independently on the specific form of the large deviation Hamiltonian. One can note that the quasipotential (47) is the opposite of the entropy S (25), up to an additive constant. The quasipotential now depends on $A(\omega)$ and satisfies the normalization condition, $\min_n \mathcal{U}_A[n] = 0$, where the minimum is achieved when $n(\boldsymbol{x},\boldsymbol{p}) = n_h^A(\boldsymbol{p})$ (26).

From Eq. (42) and the fact that $\mathcal{U}_A$ is equal to $-S$ up to a constant, it is immediately checked that $\mathcal{U}_A$ solves the stationary Hamiltonian-Jacobi equation,

$$\mathcal{H}\left[n, \frac{\delta \mathcal{U}_A}{\delta n}\right] = 0. \tag{48}$$

3.2.4 Time-reversal symmetry and detailed balance

Finally, we consider the time-reversal symmetry of the dynamics (see Appendix A.2.2). For the present wave kinetic theory, we showed at the end of section 2.4 that the time-reversed local spectral density is defined as $n_R(\boldsymbol{x},\boldsymbol{p},t) = n(\boldsymbol{x},-\boldsymbol{p},-t)$, namely that wave vector needs to change sign in addition to time-reversal $t \to -t$. Therefore, it is useful to introduce the involution operator I such that $I[n(\boldsymbol{x},\boldsymbol{p})] = n(\boldsymbol{x},-\boldsymbol{p})$.

Since the wave kinetic dynamics is an equilibrium dynamics whose stationary state is characterised by the microcanonical quasipotential (47) at the large deviation level, we expect the fluctuating dynamics to be time-reversible. Following Appendix A.2.2, we indeed check that the Hamiltonian satisfies the detailed-balance condition

$$\mathcal{H}\left[n, \lambda + \frac{\delta \mathcal{U}_A}{\delta n}\right] = \mathcal{H}\left[I[n], -I[\lambda]\right], \tag{49}$$

thus proving the time-reversal symmetry of the dynamics. For a trajectory that follows the wave kinetic equation, $\mathcal{U}_A$ monotonically decreases towards 0, and therefore the local spectral density irreversibly approaches the homogeneous distribution n_h^A. However, when μ is small but finite, there remains a possibility that $\mathcal{U}_A$ increases, namely that the spectrum moves away from the attractor

n_h^A. This property is quantified by the fluctuation relation that is equivalent to the detailed balance relation (49)

$$\frac{\mathbb{P}_{n(0)}[\{n^\mu(t)=n(t)\}_{0\leqslant t\leqslant t_f}]}{\mathbb{P}_{n(t_f)}[\{n^\mu(t)=n_R(t-t_f)\}_{0\leqslant t\leqslant t_f}]} \underset{\mu\to 0}{\asymp} \exp\left(\frac{\mathcal{U}_A[n(0)]-\mathcal{U}_A[n(t_f)]}{(2\pi\mu)^d}\right). \quad (50)$$

The fluctuation relation (50) is the fundamental answer to the irreversibility paradox that arises in any classical kinetic theory, and in particular for the classical wave kinetic theory. Indeed, let us consider a path starting from $n(0)$ that follows the wave kinetic equation until it reaches a state $n(t_f)$ at some time $t_f > 0$. Since the quasipotential $\mathcal{U}_A$ is a Lyapunov function, one has $\mathcal{U}_A[n(t_f)] < \mathcal{U}_A[n(0)]$. From the fluctuation relation Eq. (50), the exact time-reversed path starting from the state $n(t_f)$ has a probablity $\sim \exp\left(-(2\pi\mu)^{-d}\Delta\mathcal{U}\right)$ (with $\Delta\mathcal{U}=\mathcal{U}_A[n(0)]-\mathcal{U}_A[n(t_f)]>0$) to occur in the small μ limit. The irreversibility turns into an improbability.

3.3 Decomposition of the Hamiltonian

Since a free wave packet does not change its wave vector during a free propagation, and since also wave frequency is conserved during scattering by a time-independent potential, waves with different pulsation $\omega(\boldsymbol{p}) = |\boldsymbol{p}|^2/2$ (*i.e.* located on distinct spherical shells) do not interfere with each other. Therefore, the dynamics can be separated into an infinite number of subsystems in which the degrees of freedom are mutually independent. Indeed, this expectation can be verified by showing that the Hamiltonian is decomposed as an integration over frequency, as we do now.

To do so, let us rewrite the wave vector as $\boldsymbol{p} = p\boldsymbol{e}_\theta$ and define the corresponding frequency, $\omega = p^2/2$. The vector $\boldsymbol{e}_\theta$ is a unit vector whose angle is specified by θ. Generally, the number of degrees of freedom for the angle θ is $d-1$. For example, for a $d=3$ case, elevation and azimuthal angles would be selected as a set of representative coordinate variables. The following consideration is not dependent on the choice of these coordinates. We just need to assume that a pair of opposite angles, θ and $-\theta$, are defined such that $\boldsymbol{e}_\theta = -\boldsymbol{e}_{-\theta}$. We decompose the wave vector integration element as $d\boldsymbol{p} = p^{d-1}dpd\theta = (2\omega)^{(d-2)/2}d\omega d\theta$ with $d\theta$ a surface element on a $(d-1)$-dimensional unit sphere, $\mathbb{S}^{d-1}$. We define new variables labeled by frequency ω as

$$n^\omega(\boldsymbol{x},\theta) = p^{d-2}n(\boldsymbol{x},p\boldsymbol{e}_\theta) \quad (51a)$$

$$\lambda^\omega(\boldsymbol{x},\theta) = \lambda(\boldsymbol{x},p\boldsymbol{e}_\theta) \quad (51b)$$

$$\sigma^\omega(\theta_1,\theta_2) = 2\pi\Pi(p(\boldsymbol{e}_{\theta_1}-\boldsymbol{e}_{\theta_2})) \quad (51c)$$

Please do not confuse n^ω with n^μ, in the context of this section. The Hamiltonian is decomposed into independent subdynamics, each of which involves the

new variables labeled by ω,

$$\mathcal{H}[n,\lambda] = \int \mathcal{H}^\omega[n^\omega,\lambda^\omega]d\omega = \int (\mathcal{H}_F^\omega + \mathcal{H}_S^\omega)d\omega \tag{52a}$$

$$\mathcal{H}_F^\omega[n^\omega,\lambda^\omega] = -\int d\boldsymbol{x}d\theta(2\omega)^{1/2}\lambda^\omega(\boldsymbol{x},\theta)\boldsymbol{e}_\theta\cdot\nabla_x n^\omega(\boldsymbol{x},\theta) \tag{52b}$$

$$\begin{aligned}\mathcal{H}_S^\omega[n^\omega,\lambda^\omega] = c\int d\boldsymbol{x}d\theta_1 d\theta_2 \\ \times\Bigg[(2\omega)^{(d-2)/2}(\lambda^\omega(\boldsymbol{x},\theta_1)-\lambda^\omega(\boldsymbol{x},\theta_2))\sigma^\omega(\theta_1,\theta_2)n^\omega(\boldsymbol{x},\theta_1) \\ +\frac{1}{2}(\lambda^\omega(\boldsymbol{x},\theta_1)-\lambda^\omega(\boldsymbol{x},\theta_2))^2\sigma^\omega(\theta_1,\theta_2)n^\omega(\boldsymbol{x},\theta_1)n^\omega(\boldsymbol{x},\theta_2)\Bigg],\end{aligned} \tag{52c}$$

where the integration for $(\boldsymbol{x},\theta)$ is carried out over $\Gamma\times\mathbb{S}^{d-1}$. For this system, the diffusion kernel and the Lyapunov function are

$$\begin{aligned}\Sigma^\omega[n^\omega](\boldsymbol{x};\ \theta_1,\theta_2) = -c\sigma^\omega(\theta_1,\theta_2)n^\omega(\boldsymbol{x},\theta_1)n^\omega(\boldsymbol{x},\theta_2) \\ + c\int d\theta'\sigma^\omega(\theta_1,\theta')n^\omega(\boldsymbol{x},\theta_1)n^\omega(\boldsymbol{x},\theta')\delta(\theta_1-\theta_2)\end{aligned} \tag{53a}$$

$$\text{and}\quad S^\omega[n^\omega] = \int d\boldsymbol{x}d\theta(2\omega)^{(d-2)/2}\log n^\omega(\boldsymbol{x},\theta). \tag{53b}$$

The scattering part of the Hamiltonian is accordingly represented as

$$\mathcal{H}_S^\omega = \int d\boldsymbol{x}d\theta_1 d\theta_2\lambda^\omega(\boldsymbol{x},\theta_1)\Sigma[n^\omega](\boldsymbol{x};\ \theta_1,\theta_2)\left(\frac{\delta S^\omega}{\delta n^\omega(\boldsymbol{x},\theta_2)}+\lambda^\omega(\boldsymbol{x},\theta_2)\right). \tag{54}$$

We may compute the probability of a path of local spectral density separately for each frequency band. For each subsystem, the amount of wave action, $\propto\int d\boldsymbol{x}d\theta n^\omega(\boldsymbol{x},\theta)\equiv\mathcal{N}[n^\omega]$, is invariant. For the attractor of the wave kinetic equation, namely the homogeneous distribution, n^ω is everywhere constant in $\Gamma\times\mathbb{S}^{d-1}$, which we write n_h^ω. Finally, we find that the quasipotential,

$$\mathcal{U}^\omega[n^\omega] = \begin{cases}-\int d\boldsymbol{x}d\theta(2\omega)^{(d-2)/2}\log\left(\dfrac{n^\omega(\boldsymbol{x},\theta)}{n_h^\omega}\right) & \text{if}\quad \mathcal{N}[n^\omega]=\mathcal{N}[n_h^\omega] \\ +\infty & \text{otherwise}\end{cases} \tag{55}$$

satisfies the detailed balance condition, $\mathcal{H}^\omega[n^\omega,\lambda^\omega+\delta\mathcal{U}^\omega/\delta n^\omega] = \mathcal{H}^\omega[\mathcal{I}[n^\omega],-\mathcal{I}[\lambda^\omega]]$ with the involution $\mathcal{I}$ defined as $\mathcal{I}[n^\omega(\boldsymbol{x},\theta)]=n^\omega(\boldsymbol{x},-\theta)$. Consequently, the fluctuation relation for a path and its reverse applies to each subsystem separately.

3.4 Diffusive limit

In this paper, we have focused on the wave kinetic regime when the random potential presents high oscillations at a scale comparable to a typical wavelength of the wave field. Another relevant limit, referred to as the diffusive approximation [17] or the Fokker-Planck limit [50], corresponds to the regime when the spatial oscillations of the potential are at a scale larger than those of the wave field. In this regime, a wave signal propagates along rays which are randomly *refracted* by the potential, leading to an asymptotic diffusion equation for the local spectral density [17, 50].

Technically, the diffusion equation on the local spectral density is often derived from a multiscale expansion from the microscopic dynamics (7) [17, 50]. But interestingly, the diffusive limit can also be obtained from the scattering kinetic regime that we have considered in this paper. Our goal in this subsection is to show how one can derive a path large deviation theory for wave kinetics in the diffusive regime from the large deviation Hamiltonian (42). In a recent paper, a similar weak scattering limit has been considered to derive the path large deviation principle for plasma below the Debye length, related to the Landau equation, from the path large deviation principle for dilute gases, related to the Boltzmann equation [43].

In the diffusive regime, the random potential has typical variations over large lengthscales compared to the typical wavelength of the signal. As a consequence, the spectrum of the potential (6) is localised around $\boldsymbol{p} = 0$, which implies from the definition of the cross section (23) that incoming wave vectors are randomly refracted by an infinitesimal amount at each time step by the potential. One thus expects to obtain the path large deviation theory of wave transport in the diffusive regime from the scattering regime by assuming that the potential spectrum Π is supported in the vicinity of $\boldsymbol{p} \approx 0$.

In order to derive the diffusive limit from the large deviation Hamiltonian, we use Eq. (42). We show in Appendix D that the diffusion kernel Σ transforms into a differential operator such that for any test function f and g:

$$\int d\boldsymbol{p}_1 d\boldsymbol{p}_2 f(\boldsymbol{p}_1)\Sigma(\boldsymbol{x};\, \boldsymbol{p}_1, \boldsymbol{p}_2)g(\boldsymbol{p}_2) \approx \int d\boldsymbol{p}\, \nabla_p f(\boldsymbol{p}) \cdot \left[n(\boldsymbol{x}, \boldsymbol{p})^2 \boldsymbol{D}(\boldsymbol{p})\right] \cdot \nabla_p g(\boldsymbol{p}). \tag{56}$$

Here, $\boldsymbol{D}$ is the diffusion matrix computed from the potential spectrum Π as

$$\boldsymbol{D}(\boldsymbol{p}) = -\frac{c}{2} \int_{\mathbb{R}} \nabla_y \otimes \nabla_y R(\boldsymbol{p}s) ds \quad \text{and} \quad R(\boldsymbol{y}) = \int d\boldsymbol{\eta}\, \Pi(\boldsymbol{\eta}) \mathrm{e}^{i\boldsymbol{\eta}\cdot\boldsymbol{y}}. \tag{57}$$

From its definition, $R(\boldsymbol{y})$ corresponds to the rescaled two-point correlation function of the random potential.

Using Eq. (56), the large deviation Hamiltonian in the diffusive limit reads as

$$H[n, \lambda] = \int d\boldsymbol{x} d\boldsymbol{p}\, \lambda(\boldsymbol{x}, \boldsymbol{p}) \boldsymbol{p} \cdot \nabla_x n(\boldsymbol{x}, \boldsymbol{p})$$

$$+\int d\boldsymbol{x}d\boldsymbol{p}\,\nabla_p\lambda(\boldsymbol{x},\boldsymbol{p})\cdot\left[n(\boldsymbol{x},\boldsymbol{p})^2\boldsymbol{D}(\boldsymbol{p})\right]\cdot\nabla_p\left(\frac{\delta S}{\delta n(\boldsymbol{x},\boldsymbol{p})}+\lambda(\boldsymbol{x},\boldsymbol{p})\right). \tag{58}$$

The diffusive Hamiltonian (58) conserves the spectral density at a given pulsation $\omega(\boldsymbol{p}) = |\boldsymbol{p}|^2/2$ because of the fundamental property

$$\boldsymbol{D}(\boldsymbol{p})\cdot\boldsymbol{p} = 0, \tag{59}$$

which results from $\boldsymbol{p}\cdot\nabla R(\boldsymbol{p}s) = \partial R(\boldsymbol{p}s)/\partial s$. This property means that the diffusion of the spectrum is always orthogonal to the group velocity.

Moreover, the expression of the Hamiltonian (58) makes it clear that the quasipotential of the diffusive dynamics is the same as the quasipotential of the scattering limit Eq. (47). Furthermore, one can easily check that the detailed balance relation (49) is still satisfied by the Hamiltonian (58).

4 Conclusions and perspective

The linear wave kinetic equation is a statistical model that governs wave action density spreading in position and wave-vector space through propagation and scattering in random media. Motivated by recent works on dynamical large deviation principles for kinetic theory, this study has derived the large deviation principle describing the probability of a finite-time evolution of the local spectral density of wave action in an asymptotic limit of scale separation. Importantly, the large deviation principle that is derived in this paper satisfies a time-reversal symmetry with respect to the microcanoical quasipotential, that is directly (and independently) computed from the microcanonical measure.

In this paper, we restrict our considerations to the simplest Schrödinger model with a homogeneous random potential. The next step is to extend the present formulation to a wider range of situations. Possible technical difficulties encountered during such future works involve (i) generalization of the dispersion relation to a function of position and wave vector, $\omega(\boldsymbol{x},\boldsymbol{p})$, (ii) coping with spatial inhomogeneity for the randomness, that leads to a space-dependent scattering cross section, $\sigma(\boldsymbol{x},\boldsymbol{p}_1,\boldsymbol{p}_2)$, (iii) consideration of a vector field that involves polarized waves or multiple waves, e.g., elastic media holding compressional and shear waves. Combining the present approach and a previous work on wave turbulence [2], we have also derived a formula of path-large deviations for inhomogeneous spectral density in nonlinear 4-wave interacting systems [62].

In the context of geophysics, the wave kinetic equation is used to discuss energy cascades of internal waves in the oceans and atmosphere where rotation and density stratification play key roles [13]. Because the dispersion relation of internal waves depends not on wave vector but on its angle against the gravity direction, even linear theory predicts interscale energy transfer. In this process, balanced geostrophic turbulent flow acts as a random potential. If

assuming a stationary flow state, i.e., fixing the potential field in time, the formulation will be analogous to the present case. On the other hand, once we allow the temporal variations in the geostrophic flow field, the situation essentially changes. Wave frequency is no longer conserved during propagation and scattering. Spreading of action density in frequency space associates gain or loss of wave energy. Quantification of the energy exchange rate between evolving turbulent eddies and waves remains an open problem. Recently, Dong et al. [19] discussed wave frequency diffusion in geostrophic turbulence based on a kind of kinetic equation model. As pointed out in the present study, the ordinary kinetic equation predicts an irreversible change in the spectral density. In the actual environmental situation, the scale separation parameter, μ, is not necessarily small and there should be non-negligible fluctuations in spectral density. The large deviation formulation has a possible application for the estimation of a fluctuating energy transfer rate, in such regime where the kinetic theory is marginally valid.

Acknowledgments. This work was supported by JSPS Overseas Research Fellowship as well as KAKENHI Grant Number JP20K14556 (Y.O.), and by the Simons Foundation through the Collaboration Grant 651463 "Wave Turbulence" (F.B. and J.G.) and the Targeted Grant in MPS 663054 "Revisiting the Turbulence Problem Using Statistical Mechanics" (F.B.). We thank Gregory Eyink, Laure Saint-Raymond, Jacques Vanneste, and Antoine Venaille for fruitful discussions.

A Properties of the stochastic system specified by a large deviation Hamiltonian

A.1 Path large deviation

This appendix presents some general properties of a stochastic process $X^\epsilon(t)$ whose probability conditioned on an initial value, $X^\epsilon(t_i) = X(t_i)$, is specified at the large deviation level via a formula,

$$\mathbb{P}\left[\{X^\epsilon(t) = X(t)\}_{t_i \leqslant t \leqslant t_f}\right] \underset{\epsilon \to 0}{\asymp} \exp\left(-\frac{\mathcal{S}[X]}{\epsilon}\right) \tag{60a}$$

$$\mathcal{S}[X] = \int_{t_i}^{t_f} dt \mathcal{L}(X, \dot{X}) \equiv \int_{t_i}^{t_f} dt \sup_P \left\{P \cdot \dot{X} - \mathcal{H}(X, P)\right\}. \tag{60b}$$

Here, $X^\epsilon(t)$ can be a scalar, vector, or continuous function defined on some space. Basic requirements are that an inner product is properly defined, and the dynamical property of the system is controlled by a single non-negative parameter ϵ. For the simplicity, we regard X as a scalar but the following consideration can be immediately extended to general cases. The large deviation Hamiltonian $\mathcal{H}(X, P)$ is a convex function of P and satisfies $\mathcal{H}(X, 0) = 0$ for any X. From the definition, the Lagrangian $\mathcal{L}$ satisfies

$\mathcal{L}(X, \dot{X}) \geq P \cdot \dot{X} - \mathcal{H}(X, P)$ for any X, $\dot{X}$ and P. Therefore, inserting $P = 0$, we know $\mathcal{L} \geq 0$.

A.1.1 Relaxation path

Clearly, in the limit of $\epsilon \to 0$, the system becomes deterministic with a single path that minimizes the action $\mathcal{S}[X]$ for a prescribed initial condition $X(t_i)$—named the relaxation path. Since $\mathcal{L} \geq 0$, if there exists a function $R(X)$ that satisfies $\mathcal{L}(X, R(X)) = 0$, a path solving $\dot{X} = R(X)$ minimizes the action and yields $\min_X \mathcal{S}[X] = 0$. From the facts that $\mathcal{L}(X, \dot{X}) = P \cdot \dot{X} - \mathcal{H}(X, P)$ with P solving $\dot{X} - \partial\mathcal{H}/\partial P = 0$ and $\mathcal{H}(X, 0) = 0$ for any X, we understand that a function $R(X) = \partial\mathcal{H}/\partial P|_{P=0}$ fulfills $\mathcal{L}(X, R(X)) = 0$. We thus assert that an equation

$$\dot{X}^r = R(X^r) \equiv \frac{\partial \mathcal{H}}{\partial P}(X^r, 0) \tag{61}$$

determines the relaxation path $X^r(t)$.

A.1.2 Optimal path

Slightly changing the situation, if we fix both the initial and final states, $X(t_i) = x_i$ and $X(t_f) = x_f$, respectively, the most probable path from x_i to x_f, namely the optimal path, or the instanton, is obtained by again minimizing $\mathcal{S}[X]$. This problem is equivalent to the principle of least action in analytical mechanics. In this context, P is called the generalized momenta and represented by $P = \partial\mathcal{L}/\partial\dot{X}$ which is no longer 0. The optimal path in phase space is governed by a set of canonical equations,

$$\dot{X} = \frac{\partial \mathcal{H}}{\partial P} \tag{62a}$$

$$\dot{P} = -\frac{\partial \mathcal{H}}{\partial X}, \tag{62b}$$

and we shall write their solutions as $X^o[x_f, t_f;\ x_i, t_i]$ and $P^o[x_f, t_f;\ x_i, t_i]$. For the simplicity, we fix the initial conditions, t_i and x_i, and rewrite the final state as x and t. We then introduce the Hamilton's principal function $\mathcal{Q}$ as an integration of the action following the optimal path as

$$\mathcal{Q}(x, t) = \int_{t_i}^{t} d\tau \mathcal{L}(X^o(\tau), \dot{X}^o(\tau)). \tag{63}$$

It is known in analytical mechanics that in this case the generalized momenta at the final time is represented as $P^o(t) = \partial\mathcal{L}/\partial\dot{X}\Big|_t = \partial\mathcal{Q}/\partial x$, and $\mathcal{Q}$ solves

the Hamilton-Jacobi equation,

$$\frac{\partial \mathcal{Q}}{\partial t} + \mathcal{H}\left(x, \frac{\partial \mathcal{Q}}{\partial x}\right) = 0. \tag{64}$$

In the definition of $\mathcal{Q}(x,t)$, a set of arguments, (x,t), is arbitrary chosen. When we pick up an optimal path $\{X^o(\tau), P^o(\tau)\}$, at any point on this path, the generalized momenta P and the Hamilton's principle function $\mathcal{Q}$ is related via

$$P^o(\tau) = \frac{\partial \mathcal{Q}}{\partial x}(X^o(\tau), \tau). \tag{65}$$

Therefore, combining (65) and the first line of (62), we formally obtain a single equation determining the optimal path,

$$\frac{dX^o}{d\tau} = \frac{\partial \mathcal{H}}{\partial P}\left(X^o, \frac{\partial \mathcal{Q}}{\partial x}(X^o, \tau)\right). \tag{66}$$

This equation is, however, not generally useful because $\mathcal{Q}(x,t)$ is inaccessible in most cases.

A.1.3 Quasipotential and fluctuation path

Going back to the original stochastic model, the meaning of $\mathcal{Q}$ is understood as the rate function for the probability that X^ϵ reaches x_f at $t = t_f$. Indeed, we may derive from (60) an expression

$$\mathbb{P}\left[X^\epsilon(t) = x | X^\epsilon(t_i) = x_i\right] \underset{\epsilon \to 0}{\asymp} \exp\left(-\frac{\mathcal{Q}(x,t)}{\epsilon}\right) \tag{67}$$

based on the contraction principle. Now, we shall consider the stationary distribution of the probability density of X^ϵ. This can be done simply setting $t_i = -\infty$ in (67). To make the discussion more specific, let us assume that the relaxation dynamics (61) has a unique global attractor x_0, where $R(x_0) = 0$. Then, we set $x_i = x_0$ and also $t = 0$, to write a large deviation formula for the stationary distribution

$$\mathbb{P}_s(x) \underset{\epsilon \to 0}{\asymp} \exp\left(-\frac{\mathcal{U}(x)}{\epsilon}\right) \tag{68}$$

with

$$\mathcal{U}(x) = \inf_{X(t)|X(-\infty)=x_0 \text{ and } X(0)=x} \int_{-\infty}^{0} dt \mathcal{L}(X(t), \dot{X}(t)). \tag{69}$$

The rate function $\mathcal{U}$ is called the quasipotential. Since $\mathcal{U}$ is the special case of $\mathcal{Q}$ but independent of t, it solves the stationary version of the Hamilton-Jacobi

equation (64),

$$\mathcal{H}\left(x, \frac{\partial \mathcal{U}}{\partial x}\right) = 0. \tag{70}$$

For the present case, the optical path $X^o(\tau)$ represents the most probable route from an attractor x_0 to a specific point x. This route is called the fluctuation path and is denoted by $X^f(t)$. Once we obtain the quasipotential $\mathcal{U}(x)$, (66) provides an equation determining the fluctuation path as

$$\dot{X}^f = F(X^f) \equiv \frac{\partial \mathcal{H}}{\partial P}\left(X^f, \frac{\partial \mathcal{U}}{\partial x}(X^f)\right). \tag{71}$$

Since the vector field $F(x)$ does not depend on t, this equation is more useful than the original one (66). On the fluctuation path, the generalized momenta is computed based on (65) as $P^f = \partial \mathcal{U}/\partial x(X^f)$. Combining (70) with the fact that $\mathcal{H}$ is constant along the optical path, we understand that $\mathcal{H}(X^f, P^f) = 0$ always holds.

A.1.4 Quasipotential as a Lyapunov function

A relaxation path and a fluctuation path have distinct properties for the variations in $\mathcal{U}$. For a relaxation path, we have

$$\begin{aligned}\frac{d\mathcal{U}}{dt}(X^r) &= \dot{X}^r \frac{\partial \mathcal{U}}{\partial x}(X^r) = \frac{\partial \mathcal{H}}{\partial P}(X^r, 0)\frac{\partial \mathcal{U}}{\partial x}(X^r) \\ &= \underbrace{\mathcal{H}(X^r, 0)}_{=0} - \underbrace{\mathcal{H}\left(X^r, \frac{\partial \mathcal{U}}{\partial x}(X^r))\right)}_{=0} + \frac{\partial \mathcal{H}}{\partial P}(X^r, 0)\frac{\partial \mathcal{U}}{\partial x}(X^r) \le 0,\end{aligned}$$

where we have used the general expressions, $\mathcal{H}(X, 0) = \mathcal{H}(X, \partial \mathcal{U}/\partial x(X)) = 0$, and the convexity of $\mathcal{H}(X, P)$ for P. For a fluctuation path, we have

$$\begin{aligned}\frac{d\mathcal{U}}{dt}(X^f) &= \dot{X}^f \frac{\partial \mathcal{U}}{\partial x}(X^f) = \frac{\partial \mathcal{H}}{\partial P}\left(X^f, \frac{\partial \mathcal{U}}{\partial x}(X^f)\right)\frac{\partial \mathcal{U}}{\partial x}(X^f) \\ &= \underbrace{\mathcal{H}(X^f, 0)}_{=0} - \underbrace{\mathcal{H}\left(X^f, \frac{\partial \mathcal{U}}{\partial x}(X^f))\right)}_{=0} + \frac{\partial \mathcal{H}}{\partial P}\left(X^f, \frac{\partial \mathcal{U}}{\partial x}(X^f)\right)\frac{\partial \mathcal{U}}{\partial x}(X^f) \ge 0,\end{aligned}$$

again from the convexity of $\mathcal{H}$. We have thus learned that the quasipotential is a Lyapunov function because it monotonically decreases in a relaxation path while increases in a fluctuation path. These results are natural consequence from a basic property that $\mathcal{U}(x)$ is minimum at the attractor x_0 and the relaxation and fluctuation paths represent routes to and from the attractor.

A.2 Properties of large deviation Hamiltonian

A.2.1 Conservation law

From now on, we shall regard X and P as vectors so that there are multiple directions in X space. For this case, when the Hamiltonian $\mathcal{H}$ possesses a kind of symmetry, it is related to the conservation law of the system. More specifically, let us suppose that we find a function $\mathcal{C}(X)$ that satisfies

$$\mathcal{H}\left(X, P + \alpha \frac{\partial \mathcal{C}}{\partial X}\right) = \mathcal{H}(X, P) \tag{72}$$

for any X, P, and α. This condition is equivalent to

$$\frac{\partial \mathcal{H}}{\partial P}(X, P) \cdot \frac{\partial \mathcal{C}}{\partial X}(X) = 0$$

for any X and P. Now that $\mathcal{H}(X, \cdot)$ is flat in the direction of $\partial \mathcal{C}/\partial X$, from the property of the Legendre-Fenchel transform [63], the corresponding Lagrangian has a property,

$$\mathcal{L}(X, \dot{X}) = +\infty \quad \text{if} \quad \dot{X} \cdot \frac{\partial \mathcal{C}}{\partial X}(X) \neq 0. \tag{73}$$

This expression indicates that the probability for a path crossing a contour of $\mathcal{C}$ is strictly 0. This constraint applies not only to the optimal path but also to any path with random fluctuations. We thus understand that (72) serves as a condition for $\mathcal{C}$ being an invariant of the system.

A.2.2 Detailed balance

The detailed balance is a property of equilibrium states which asserts time-reversibility of the process, meaning that the probabilities of any trajectory and its reversed counterpart are equal. A basic expression of detailed balance for a stationary stochastic process is $\mathbb{P}_{\Delta t}(y;\ x)\mathbb{P}_S(x) = \mathbb{P}_{\Delta t}(x;\ y)\mathbb{P}_S(y)$, where $\mathbb{P}_{\Delta t}(y;\ x)$ is the transition probability from a state x to another state y during a time interval Δt, and $\mathbb{P}_S(x)$ is the stationary probability distribution.

Since we are now considering a continuous Markov process, it is enough to regard Δt as arbitrary small. For the limit of $\Delta t \to 0$, we may write $y \sim x + \dot{x}\Delta t$ and redefine the transition probability as $\mathbb{P}_{\Delta t}(x, \dot{x}) \sim \mathbb{P}_{\Delta t}(x + \dot{x}\Delta t;\ x)$. Assuming the continuity of $\mathbb{P}$ and $\mathbb{P}_S$, the detailed balance condition is rewritten as

$$\mathbb{P}_{\Delta t}(x, \dot{x})\mathbb{P}_S(x) \sim \mathbb{P}_{\Delta t}(x + \dot{x}\Delta t, -\dot{x})\mathbb{P}_S(x + \dot{x}\Delta t). \tag{74}$$

For the present problem, the probability distribution is specified as $\mathbb{P}_{\Delta t}(x, \dot{x}) \asymp \exp(-\Delta t \mathcal{L}(x, \dot{x})/\epsilon)$ and $\mathbb{P}_S(x) \asymp \exp(-\mathcal{U}(x)/\epsilon)$. Therefore, the detailed

balance condition (74) is rewritten as

$$\mathcal{L}(x,\dot{x}) - \mathcal{L}(x,-\dot{x}) = \dot{x}\cdot\frac{\partial \mathcal{U}}{\partial x}. \tag{75}$$

This condition is modified in terms of the Hamiltonian via the Legendre-Fenchel transform as

$$\mathcal{H}(x,-p) = \mathcal{H}\left(x, p+\frac{\partial \mathcal{U}}{\partial x}\right). \tag{76}$$

Because $\mathcal{H}(x,0)=0$ in the current problem, the stationary Hamilton-Jacobi equation (70) is a necessary (but not sufficient) condition for $\mathcal{U}$ being the quasipotential.

Once the detailed balance condition (76) is verified, we understand that the probabilities of a path and its reverse are related at the large deviation level via the expression,

$$\frac{\mathbb{P}\left[\{X^\epsilon(t)=X(t)\}_{t_i\leqslant t\leqslant t_f}\right]}{\mathbb{P}\left[\{X^\epsilon(t)=X(t_f+t_i-t)\}_{t_i\leqslant t\leqslant t_f}\right]} \underset{\epsilon\to 0}{\asymp} \exp\left(-\frac{\mathcal{U}(x_f)-\mathcal{U}(x_i)}{\epsilon}\right), \tag{77}$$

an equivalent form of the Crooks fluctuation theorem. Another outcome of the detailed balance is that the fluctuation path is the time reverse of the relaxation path. This property is derived from

$$R(X) = \frac{\partial \mathcal{H}}{\partial P}(X,0) = -\frac{\partial \mathcal{H}}{\partial P}\left(X, \frac{\partial \mathcal{U}}{\partial X}\right) = -F(X).$$

B Microcanonical ensemble and quasi-potential for the Schrödinger equation

In this appendix, we consider the microcanonical ensemble of the dynamics governed by the Schrödinger equation, (7). The aim is to compute the quasipotential of the local empirical spectral density, i.e., the large deviation rate function of n^μ, in the small μ limit. We will prove that

$$\mathbb{P}^\mu_{A,m}[n^\mu = n] \underset{\mu\to 0}{\asymp} \exp\left(-\frac{\mathcal{U}_A[n]}{(2\pi\mu)^d}\right), \tag{78}$$

with $\mathbb{P}^\mu_{A,m}$ the probabilities with respect to the microcanonical measure with the constraints,

$$\mathcal{A}_\omega[n^\mu] \equiv \int h\left(\omega - \frac{|\boldsymbol{p}|^2}{2}\right) n^\mu(\boldsymbol{x},\boldsymbol{p}) d\boldsymbol{x} d\boldsymbol{p} = A(\omega), \tag{79}$$

where h is the Heaviside function, and $A : \mathbb{R}^+ \to \mathbb{R}^+$ is a prescribed function. The expected form of the rate function, namely, the quasipotential, is

$$\mathcal{U}_A[n] = \begin{cases} -\int d\boldsymbol{x}d\boldsymbol{p} \log\left(\dfrac{n(\boldsymbol{x},\boldsymbol{p})}{n_h(\boldsymbol{p})}\right) & \text{if} \quad \mathcal{A}_\omega[n] = A(\omega) \\ +\infty & \text{otherwise} \end{cases}, \tag{80}$$

where

$$n_h^A(\boldsymbol{p}) = \frac{A'(|\boldsymbol{p}|^2/2)}{\int d\boldsymbol{x}d\boldsymbol{\eta}\delta(|\boldsymbol{p}|^2/2 - |\boldsymbol{\eta}|^2/2)}. \tag{81}$$

In the following proof, we first define the microcanonical measure and then derive the rate function (80).

If we consider an infinite domain with a finite amount of total wave action, $\int_{\mathbb{R}^d} |\psi^\mu|^2 d\boldsymbol{x} < \infty$, since the wave action density will be diluted to absolutely 0 everywhere, an equilibrium state of the microcanonical ensemble does not make sense. Here, we instead assume a spatial periodicity of the scaler field $\psi^\mu(\boldsymbol{x})$ and concentrate our attention on a d-dimensional cubic domain, $[0, 2\pi)^d \equiv \Gamma \subset \mathbb{R}^d$. In this setting, the empirical local spectral density (10) consists of delta functions in wave-vector space like

$$n^\mu(\boldsymbol{x},\boldsymbol{p}) = \sum_{\boldsymbol{k}\in(1/2)\mathbb{Z}^d} n^\mu_{\boldsymbol{k}}(\boldsymbol{x})\delta(\boldsymbol{p} - \mu\boldsymbol{k}), \tag{82}$$

where $n^\mu_k(\boldsymbol{x})$ is a discrete form of Wigner distribution adapted to periodic domains, defined as

$$n^\mu_{\boldsymbol{k}}(\boldsymbol{x}) = \frac{1}{(2\pi)^d}\int_\Gamma d\boldsymbol{y} e^{-2i\boldsymbol{k}\cdot\boldsymbol{y}} \psi^\mu\left(\boldsymbol{x}+\boldsymbol{y}\right)\psi^{\mu\dagger}\left(\boldsymbol{x}-\boldsymbol{y}\right). \tag{83}$$

We also introduce the Fourier coefficients of ψ^μ as

$$\hat{\psi}^\mu_{\boldsymbol{k}} = \frac{1}{(2\pi)^d}\int_\Gamma e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\psi^\mu(\boldsymbol{x})d\boldsymbol{x} \tag{84a}$$

$$\psi^\mu(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\hat{\psi}^\mu_{\boldsymbol{k}}. \tag{84b}$$

Perceval's theorem allows us to write the wave action density per unit volume in three forms,

$$\frac{1}{(2\pi)^d}\int_{\Gamma\times\mathbb{R}^d} n^\mu(\boldsymbol{x},\boldsymbol{p})d\boldsymbol{x}d\boldsymbol{p} = \frac{1}{(2\pi)^d}\int_\Gamma |\psi^\mu(\boldsymbol{x})|^2 d\boldsymbol{x} = \sum_{\boldsymbol{k}} |\hat{\psi}^\mu_{\boldsymbol{k}}|^2. \tag{85}$$

We shall set $\mu \to 0$ while keeping this action density finite. When we fix a volume element in $\boldsymbol{p}$ space, the number of $\boldsymbol{k}$ vectors which are involved there

increases as $\sim \mu^{-d}$. Therefore, the typical amplitude of the Fourier coefficients depends on μ as $\hat{\psi}^\mu_{\boldsymbol{k}} \sim \mathcal{O}(\mu^{d/2})$. Now we scale $\hat{\psi}^\mu$ by introducing a new coefficient, $a^\mu_{\boldsymbol{p}}$ with $\boldsymbol{p} \in \mu\mathbb{Z}^d$, as $\hat{\psi}^\mu_{\boldsymbol{k}} = \mu^{d/2} a^\mu_{\mu\boldsymbol{k}}$, i.e,

$$a^\mu_{\boldsymbol{p}} = \frac{1}{(2\pi\mu^{1/2})^d} \int_\Gamma e^{-\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{x}/\mu} \psi^\mu(\boldsymbol{x}) d\boldsymbol{x}. \tag{86}$$

Note that $a^\mu_{\boldsymbol{p}}$ remains finite for $\mu \to 0$, but its norm, $\sum_{\boldsymbol{p}\in\mu\mathbb{Z}^d} |a^\mu_{\boldsymbol{p}}|^2 = (2\pi\mu)^{-d} \int_{\Gamma\times\mathbb{R}^d} n^\mu(\boldsymbol{x},\boldsymbol{p}) d\boldsymbol{x}d\boldsymbol{p}$, diverges in the same limit.

To apply equilibrium statistical mechanics, we consider the phase space spanned by the scaled coefficients, $\{a^\mu_{\boldsymbol{p}}\}_{\boldsymbol{p}\in\mu\mathbb{Z}^d}$. In this space, the Lebesgue measure $\mathfrak{m}$ is represented as

$$d\mathfrak{m} = \prod_{\boldsymbol{p}} da^\mu_{\boldsymbol{p}} da^{\mu\dagger}_{\boldsymbol{p}} \equiv \prod_{\boldsymbol{p}} da^r_{\boldsymbol{p}} da^i_{\boldsymbol{p}}, \tag{87}$$

where $a^\mu_{\boldsymbol{p}} = (a^r_{\boldsymbol{p}} + \mathrm{i}a^i_{\boldsymbol{p}})/\sqrt{2}$ is understood. This measure makes sense only when an upper limit of the wave vector is set to truncate the infinite product. As a result, the number of degrees of freedom of ψ^μ in physical space is also restricted. Let us define the bounded set of $\boldsymbol{k}$ as $\mathbb{K}_\Delta \equiv \{-1/(2\Delta)+1, \ldots, 1/(2\Delta)\}^d$, and accordingly that of $\boldsymbol{p}$ as $\mu\mathbb{K}_\Delta$. The number of elements in $\mathbb{K}_\Delta$ is $\mathcal{N} \equiv 1/\Delta^d$. Then, ψ^μ is specified by the values at $\mathcal{N}$ points, $\Gamma_\Delta \equiv \{0, \Delta, \ldots, 2\pi - \Delta\}^d$, and the values in $\Gamma \setminus \Gamma_\Delta$ are determined by interpolation. The Lebesgue measure is now represented in either wave vector or position space as

$$d\mathfrak{m} = \prod_{\boldsymbol{p}\in\mu\mathbb{K}_\Delta} da^\mu_{\boldsymbol{p}} da^{\mu\dagger}_{\boldsymbol{p}} = J^\mu_\Delta \prod_{\boldsymbol{x}\in\Gamma_\Delta} d\psi^\mu(\boldsymbol{x}) d\psi^{\mu\dagger}(\boldsymbol{x}), \tag{88}$$

where J^μ_Δ is the Jacobian of the function that maps $a^\mu_{\boldsymbol{p}}$ to $\hat{\psi}^\mu$. To compute this Jacobian, we consider the integral,

$$\int d\mathfrak{m} \exp\left[-\frac{1}{(2\pi\mu)^d} \int_{\Gamma\times\mathbb{R}^d} n^\mu(\boldsymbol{x},\boldsymbol{p}) d\boldsymbol{x}d\boldsymbol{p}\right] \tag{89}$$

that we express in two different ways:

$$\begin{aligned} &\int \left(\prod_{\boldsymbol{p}\in\mu\mathbb{K}_\Delta} da^\mu_{\boldsymbol{p}} da^{\mu\dagger}_{\boldsymbol{p}}\right) \exp\left[-\sum_{\boldsymbol{p}\in\mu\mathbb{K}_\Delta} |a^\mu_{\boldsymbol{p}}|^2\right] \\ =&J^\mu_\Delta \int \left(\prod_{\boldsymbol{x}\in\Gamma_\Delta} d\psi^\mu(\boldsymbol{x}) d\psi^{\mu\dagger}(\boldsymbol{x})\right) \exp\left[-\left(\frac{\Delta}{2\pi\mu}\right)^d \sum_{\boldsymbol{x}\in\Gamma_\Delta} |\psi^\mu(\boldsymbol{x})|^2\right]. \end{aligned} \tag{90}$$

The left-hand side turns out to be $(2\pi)^{\mathcal{N}}$, and the right-hand side is $J_\Delta^\mu (2\pi)^{\mathcal{N}} (2\pi\mu)^{d\mathcal{N}} \mathcal{N}^{\mathcal{N}}$. We therefore obtain J_Δ^μ and, accordingly,

$$d\mathfrak{m} = \prod_{\boldsymbol{x}\in\Gamma_\Delta} \frac{d\psi^\mu(\boldsymbol{x}) d\psi^{\mu\dagger}(\boldsymbol{x})}{(2\pi\mu)^d \mathcal{N}}. \tag{91}$$

We denote by $\mathbb{E}$ integrals over the Lebesgue measure (91). The microcanonical measure with constraints on A is then defined as

$$d\mathfrak{m}_A = \frac{\Pi_\omega \delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)}{\mathbb{E}\left[\Pi_\omega \delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)\right]} d\mathfrak{m}, \tag{92}$$

where $\Pi_\omega \delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)$ means that we constrain the values of all the invariants $\mathcal{A}_\omega$ for any ω.

Our goal is to compute the probability distribution of n^μ for a microcanonical measure constrained by A, i.e.,

$$\mathbb{P}_{A,m}^\mu[n^\mu = n] = \frac{\mathbb{E}\left[\delta\left(n^\mu - n\right) \Pi_\omega \delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)\right]}{\mathbb{E}\left[\Pi_\omega \delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)\right]}. \tag{93}$$

It will be mathematically convenient, for intermediate computations, to use in the following a normalizable Gaussian measure $d\mathfrak{m}_G$,

$$\begin{aligned} d\mathfrak{m}_G &= \prod_{\boldsymbol{p}\in\mu\mathbb{K}_\Delta} e^{-|a_{\boldsymbol{p}}^\mu|^2} \frac{da_{\boldsymbol{p}}^\mu da_{\boldsymbol{p}}^{\mu\dagger}}{2\pi} \\ &= \exp\left[-\frac{1}{(2\pi\mu)^d} \int_{\Gamma\times\mathbb{R}^d} n^\mu(\boldsymbol{x},\boldsymbol{p}) d\boldsymbol{x} d\boldsymbol{p} - \mathcal{N}\log 2\pi\right] d\mathfrak{m}, \end{aligned} \tag{94}$$

which satisfies $\int d\mathfrak{m}_G = 1$. We note that the $\mathcal{N}$-dependent term diverges in the $\Delta \to 0$ limit, but this divergence will be compensated in $\mathbb{P}_A^\mu$ below. We denote $\mathbb{E}_G$ averages with respect to this Gaussian measure.

Then, (93) can be rewritten as

$$\begin{aligned} &\mathbb{P}_{A,m}^\mu[n^\mu = n] \\ &= \frac{\exp\left[(2\pi\mu)^{-d} \int_{\Gamma\times\mathbb{R}^d} n(\boldsymbol{x},\boldsymbol{p}) d\boldsymbol{x} d\boldsymbol{p}\right] \mathbb{E}_G\left[\delta\left(n^\mu - n\right) \Pi_\omega \delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)\right]}{\mathbb{E}_G\left\{\exp\left[(2\pi\mu)^{-d} \int_{\Gamma\times\mathbb{R}^d} n^\mu(\boldsymbol{x},\boldsymbol{p}) d\boldsymbol{x} d\boldsymbol{p}\right] \Pi_\omega \delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)\right\}}. \end{aligned} \tag{95}$$

In the following, when we consider the Gaussian measure $\mathfrak{m}_G$, the continuous limit $\Delta \to 0$ is always understood.

We look for a large deviation principle

$$\mathbb{E}_G\left[\delta\left(n^\mu - n\right) \Pi_\omega \delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)\right] \underset{\mu\to 0}{\asymp} \exp\left(-\frac{\mathcal{I}_G[n,A]}{(2\pi\mu)^d}\right). \tag{96}$$

Our strategy is to compute its rescaled cumulant generating function and to apply the Gärtner-Ellis theorem. We shall then define a free energy as

$$f_G[\lambda,\beta] \equiv -\lim_{\mu\to 0}(2\pi\mu)^d \log \mathbb{E}_G\left\{\exp\left[-\frac{1}{(2\pi\mu)^d}\int_{\Gamma\times\mathbb{R}^d} d\boldsymbol{x}d\boldsymbol{p} \times\left(\int_{\mathbb{R}^+}\beta(\omega)h\left(\omega-\frac{|\boldsymbol{p}|^2}{2}\right)d\omega+\lambda(\boldsymbol{x},\boldsymbol{p})\right)n^\mu(\boldsymbol{x},\boldsymbol{p})\right]\right\}, \tag{97}$$

where $\beta : \mathbb{R}^+ \to \mathbb{R}$ is a real continuous function representing the chemical potential and $\lambda : \Gamma \times \mathbb{R}^d \to \mathbb{R}$ is also a real continuous function. Because n^μ is quadratic in ψ^μ, the expectation is a Gaussian integral. To make it explicit, the expression in the square brackets is rewritten as

$$-\frac{1}{(2\pi\mu)^d}\int \psi^{\mu\dagger}(\boldsymbol{x})\mathcal{L}_{\tilde{\lambda}}\psi^\mu(\boldsymbol{x})d\boldsymbol{x}, \tag{98}$$

where $\mathcal{L}_{\tilde{\lambda}}$ is a pseudo-differential operator defined as

$$\mathcal{L}_{\tilde{\lambda}}\psi^\mu(\boldsymbol{x}) \equiv \int L_{\tilde{\lambda}}(\boldsymbol{x},\boldsymbol{x}')\psi^\mu(\boldsymbol{x}')d\boldsymbol{x}' \tag{99a}$$

$$L_{\tilde{\lambda}}(\boldsymbol{x},\boldsymbol{x}') = \frac{1}{(2\pi\mu)^d}\int_{\mathbb{R}^d}\tilde{\lambda}\left(\frac{\boldsymbol{x}+\boldsymbol{x}'}{2},\boldsymbol{p}\right)e^{i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}')/\mu}d\boldsymbol{p} \tag{99b}$$

$$\tilde{\lambda}(\boldsymbol{x},\boldsymbol{p}) \equiv \begin{cases}\lambda(\boldsymbol{x},\boldsymbol{p})+\int_{\mathbb{R}^+}\beta(\omega)h(\omega-|\boldsymbol{p}|^2/2)d\omega & (\boldsymbol{x}\in\Gamma)\\ 0 \quad (\boldsymbol{x}\notin\Gamma).\end{cases} \tag{99c}$$

In (98) and (99a), the integration range is unbounded. However, Riemann-Lebesgue lemma applied to (99b) assures that the kernel function $L_{\tilde{\lambda}}$ vanishes in the limit of $\mu \to 0$ except in the vicinity of the points of $\boldsymbol{x} = \boldsymbol{x}'$. Consequently, further taking into account (99c), the range of the integration of (98) and (99a) can be reduced from $\mathbb{R}^d$ to Γ.

To obtain a simpler form of $f_G[\lambda,\beta]$, we need to compute the functional determinant of $\mathcal{L}_{\tilde{\lambda}}$. This is done here straightforwardly:

$$\begin{aligned} f_G[\lambda,\beta] &= -\lim_{\mu\to 0}(2\pi\mu)^d\log\mathbb{E}_G\left\{\exp\left[-\frac{1}{(2\pi\mu)^d}\int_{\mathbb{R}^d}\psi^{\mu\dagger}(\boldsymbol{x})\mathcal{L}_{\tilde{\lambda}}\psi^\mu(\boldsymbol{x})d\boldsymbol{x}\right]\right\}\\ &= -\lim_{\mu\to 0}\lim_{\Delta\to 0}(2\pi\mu)^d\log\int\prod_{\boldsymbol{x}\in\Gamma_\Delta}\frac{d\psi^\mu(\boldsymbol{x})d\psi^{\mu\dagger}(\boldsymbol{x})}{2\pi(2\pi\mu)^d\mathcal{N}}\\ &\quad\times\exp\left[-\frac{\Delta^{2d}}{(2\pi\mu)^d}\sum_{\boldsymbol{x},\boldsymbol{x}'\in\Gamma_\Delta}\psi^{\mu\dagger}(\boldsymbol{x})L_{\tilde{\lambda}}(\boldsymbol{x},\boldsymbol{x}')\psi^\mu(\boldsymbol{x}')-\frac{\Delta^d}{(2\pi\mu)^d}\sum_{\boldsymbol{x}\in\Gamma_\Delta}|\psi^\mu(\boldsymbol{x})|^2\right]\\ &= \lim_{\mu\to 0}\lim_{\Delta\to 0}(2\pi\mu)^d\log\det\left(\boldsymbol{I}+\boldsymbol{L}_{\tilde{\lambda}}^{\Delta}\right). \end{aligned} \tag{100}$$

We have carried out the Gaussian integration. The problem thus reduces to computing the determinant of an $\mathcal{N} \times \mathcal{N}$ matrix, $\boldsymbol{I} + \boldsymbol{L}^{\Delta}_{\tilde{\lambda}}$, where $\boldsymbol{I}$ is a unit matrix and $\boldsymbol{L}^{\Delta}_{\tilde{\lambda}}$ consists of $\left\{\Delta^d L_{\tilde{\lambda}}(\boldsymbol{x}, \boldsymbol{x}') \mid \boldsymbol{x}, \boldsymbol{x}' \in \Gamma_\Delta\right\}$. Let us use the following expression,

$$\begin{aligned}\log\det\left(\boldsymbol{I} + \boldsymbol{L}^{\Delta}_{\tilde{\lambda}}\right) &= \mathrm{tr}\log\left(\boldsymbol{I} + \boldsymbol{L}^{\Delta}_{\tilde{\lambda}}\right) \\ &= -\sum_{j=1}^{\infty} \frac{(-1)^j}{j} \mathrm{tr}\boldsymbol{L}^{\Delta j}_{\tilde{\lambda}},\end{aligned} \tag{101}$$

which holds for sufficiently small $\tilde{\lambda}$. We know that $\boldsymbol{L}^{\Delta j}_{\tilde{\lambda}}$ consists of $\left\{\Delta^d L^j_{\tilde{\lambda}}(\boldsymbol{x}, \boldsymbol{x}') \mid \boldsymbol{x}, \boldsymbol{x}' \in \Gamma_\Delta\right\}$ where $L^j_{\tilde{\lambda}}(\boldsymbol{x}, \boldsymbol{x}')$ corresponds to the kernel function of an operator,

$$\mathcal{L}^j_{\tilde{\lambda}} \equiv \underbrace{\mathcal{L}_{\tilde{\lambda}} \ldots \mathcal{L}_{\tilde{\lambda}}}_{j}, \tag{102}$$

in the small Δ limit. In general, a product between pseudo-differential operators corresponds to a star product, or a Moyal product, between symbols [64, 65]. The star product is expanded in terms of μ with the leading term equivalent to the ordinary product. Accordingly, $\mathcal{L}^j_{\tilde{\lambda}} = \mathcal{L}_{\tilde{\lambda}^j} + \mathcal{O}(\mu)$ holds, where $\tilde{\lambda}^j$ is the jth power of $\tilde{\lambda}$. We thus derive

$$\begin{aligned}\lim_{\Delta\to 0} L^j_{\tilde{\lambda}}(\boldsymbol{x}, \boldsymbol{x}) &= \frac{1}{(2\pi\mu)^d} \int_{\mathbb{R}^d} \tilde{\lambda}^j(\boldsymbol{x}, \boldsymbol{p}) d\boldsymbol{p} + \mathcal{O}(\mu) \\ \therefore \lim_{\Delta\to 0} \mathrm{tr}\boldsymbol{L}^{\Delta j}_{\tilde{\lambda}} &= \frac{1}{(2\pi\mu)^d} \int_{\Gamma\times\mathbb{R}^d} \tilde{\lambda}^j(\boldsymbol{x}, \boldsymbol{p}) d\boldsymbol{x} d\boldsymbol{p} + \mathcal{O}(\mu),\end{aligned} \tag{103}$$

and hence

$$f_G[\lambda, \beta] = \int_{\Gamma\times\mathbb{R}^d} \log\left(\lambda(\boldsymbol{x}, \boldsymbol{p}) + \int_{\mathbb{R}^+} \beta(\omega) h\left(\omega - \frac{|\boldsymbol{p}|^2}{2}\right) d\omega + 1\right) d\boldsymbol{x} d\boldsymbol{p}. \tag{104}$$

From this formula, the Gärtner-Ellis theorem yields the rate function $\mathcal{I}_G$ (96) as

$$\mathcal{I}_G[n, A] = -\inf_{\lambda,\beta}\left\{\int d\boldsymbol{x} d\boldsymbol{p} \lambda(\boldsymbol{x}, \boldsymbol{p}) n(\boldsymbol{x}, \boldsymbol{p}) + \int_{\mathbb{R}^+} d\omega \beta(\omega) A(\omega) - f[\lambda, \beta]\right\}, \tag{105}$$

which is computed as

$$\mathcal{I}_G[n, A] = \begin{cases} \int d\boldsymbol{x} d\boldsymbol{p}\,(n - 1 - \log n) & \text{if} \quad \mathcal{A}_\omega[n] = A(\omega) \\ +\infty & \text{otherwise} \end{cases}. \tag{106}$$

As should have been expected, the minimum value of $\mathcal{I}_G[n, A]$ is 0, which is realized when and only when $n = 1$ and $A(\omega) = \mathcal{A}_\omega[1]$.

Based on the large deviation result (96), the numerator in (95) turns out to be

$$\exp\left[(2\pi\mu)^{-d}\int_{\Gamma\times\mathbb{R}^d} n(\boldsymbol{x},\boldsymbol{p})d\boldsymbol{x}d\boldsymbol{p}\right]\mathbb{E}_G\left[\delta\left(n^\mu - n\right)\Pi_\omega\delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)\right]$$
$$\underset{\mu\to 0}{\asymp}\exp\left(\frac{S_A[n]}{(2\pi\mu)^d}\right) \tag{107}$$

with

$$S_A[n] = \begin{cases}\int d\boldsymbol{x}d\boldsymbol{p}\,(1+\log n) & \text{if} \quad \mathcal{A}_\omega[n] = A(\omega)\\ -\infty & \text{otherwise}\end{cases}. \tag{108}$$

The finite part of this function defines the entropy for a mesoscopic state specified by n, and it coincides with the Lyapunov function (25) for the wave kinetic equation. Because the denominator of (95) is the integration of the numerator over all n, the Laplace's principle enables us to compute it as

$$\mathbb{E}_G\left\{\exp\left[(2\pi\mu)^{-d}\int_{\Gamma\times\mathbb{R}^d} n^\mu(\boldsymbol{x},\boldsymbol{p})d\boldsymbol{x}d\boldsymbol{p}\right]\Pi_\omega\delta\left(\mathcal{A}_\omega\left[n^\mu\right] - A(\omega)\right)\right\}$$
$$\underset{\mu\to 0}{\asymp}\exp\left(\frac{s[A]}{(2\pi\mu)^d}\right), \tag{109}$$

with $s[A] = \sup_n\{S_A[n]\}$. The supremum is achieved when $\mathcal{A}_\omega[n] = A(\omega)$, and

$$n(\boldsymbol{x},\boldsymbol{p}) = n_h^A(\boldsymbol{p}) = N\left(\frac{|\boldsymbol{p}|^2}{2}\right), \tag{110}$$

that is, n is homogeneous in space and depends only on the magnitude of its wave vector. The function $N(\omega)$ is related to $A(\omega)$ by the condition $\mathcal{A}_\omega[n_h^A] = A(\omega)$. This gives formula (81) for n_h^A. We also have

$$s[A] = \int d\boldsymbol{x}d\boldsymbol{p}\left(1+\log n_h^A\right). \tag{111}$$

Finally, starting from (95), and using the two asymptotic relations (107) and (109), as well as (108) and (111), we obtain

$$\mathbb{P}^\mu_{A,m}[n^\mu = n] \underset{\mu\to 0}{\asymp}\exp\left(-\frac{\mathcal{U}_A[n]}{(2\pi\mu)^d}\right) \tag{112}$$

where the quasipotential $\mathcal{U}_A[n]$ is given by equation (80). We have established the announced results.

C Computations of scattering terms

This appendix describes somewhat intricate derivation of the scattering terms that appear in the wave kinetic equation and the large deviation Hamiltonian. For this purpose, we prepare some useful formulae,

$$\int_0^t d\tau_1 e^{i\omega\tau_1/\mu} \int_0^{\tau_1} d\tau_2 e^{-i\omega\tau_2/\mu} = \pi\mu t\delta(\omega) + o(\mu) \tag{113a}$$

$$\int_0^t d\tau_1 e^{i\omega\tau_1/\mu} \int_0^{t} d\tau_2 e^{-i\omega\tau_2/\mu} = 2\pi\mu t\delta(\omega) + o(\mu) \tag{113b}$$

$$\frac{1}{(2\pi\mu)^d} \int_{\mathbb{R}^d} d\boldsymbol{\xi} e^{i\boldsymbol{p}\cdot(\boldsymbol{\xi}-\boldsymbol{x})/\mu} f(\boldsymbol{\xi}) = \sum_{|\alpha|\geq 0} \frac{(-\mathrm{i}\mu)^{|\alpha|}}{\alpha!} \nabla_x^\alpha f(\boldsymbol{x}) \nabla_p^\alpha \delta(\boldsymbol{p}). \tag{113c}$$

Equations (113a) and (113b) are often used in literature of weak turbulence [2]. The residual terms denoted by $o(\mu)$ make negligible contributions in the limit of $\mu \to 0$ compared to the leading-order terms when integrated with respect to ω. In (113c), a multi-index notation is used. In the following computation, integration is always carried out over $\mathbb{R}^d$, except for the basic positional coordinates represented by $\boldsymbol{x}$ whose integration range is Γ.

C.1 Terms appearing in the classical wave kinetic equation

We first compute the scattering terms in the wave kinetic equation, specifically $\mathbb{E}\left[w^\mu(\psi_2^\mu, \psi_0^\mu)\right]$, $\mathbb{E}\left[w^\mu(\psi_0^\mu, \psi_2^\mu)\right]$, and $\mathbb{E}\left[w^\mu(\psi_1^\mu, \psi_1^\mu)\right]$. These terms are common with those linear to λ in the scattering part of the large deviation Hamiltonian, $\mathcal{H}_S$. The computations are slightly involved but mostly straightforward. A detailed procedure is presented only for the $\mathbb{E}\left[w^\mu(\psi_2^\mu, \psi_0^\mu)\right]$ case.

From (16) and (18), we have

$$w^\mu(\psi_2^\mu, \psi_0^\mu) = \frac{1}{-\mu^2(2\pi\mu)^d} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \int d\boldsymbol{y} d\boldsymbol{\xi}_{1234} e^{-i\boldsymbol{p}\cdot\boldsymbol{y}/\mu}$$
$$\times G^\mu\left(\boldsymbol{x} + \frac{\boldsymbol{y}}{2} - \boldsymbol{\xi}_1, t - \tau_1\right) V^\mu(\boldsymbol{\xi}_1) G^\mu(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2, \tau_1 - \tau_2) V^\mu(\boldsymbol{\xi}_2) G^\mu(\boldsymbol{\xi}_2 - \boldsymbol{\xi}_3, \tau_2) \psi^\mu(\boldsymbol{\xi}_3, 0)$$
$$\times G^{\mu\dagger}\left(\boldsymbol{x} - \frac{\boldsymbol{y}}{2} - \boldsymbol{\xi}_4, t\right) \psi^{\mu\dagger}(\boldsymbol{\xi}_4, 0).$$

Taking ensemble average, writing the propagators G^μ as Fourier integrals (17) with wave vectors $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3$ and $\boldsymbol{\eta}_4$ in this order, and setting

$$\mathbb{E}\left[V^\mu(\boldsymbol{\xi}_1) V^\mu(\boldsymbol{\xi}_2)\right] = \int d\boldsymbol{\eta}_5 e^{i\boldsymbol{\eta}_5\cdot(\boldsymbol{\xi}_1-\boldsymbol{\xi}_2)/\mu} \Pi(\boldsymbol{\eta}_5)$$
$$\psi^\mu(\boldsymbol{\xi}_3, 0) \psi^{\mu\dagger}(\boldsymbol{\xi}_4, 0) = \int d\boldsymbol{\eta}_6 e^{i\boldsymbol{\eta}_6\cdot(\boldsymbol{\xi}_3-\boldsymbol{\xi}_4)/\mu} n((\boldsymbol{\xi}_3 + \boldsymbol{\xi}_4)/2, \boldsymbol{\eta}_6),$$

we derive

$$
\begin{aligned}
&\mathbb{E}\left[w^{\mu}(\psi_2^{\mu},\psi_0^{\mu})\right]\\
=&\frac{1}{-\mu^2(2\pi\mu)^{5d}}\int d\boldsymbol{y}d\boldsymbol{\xi}_{1234}d\boldsymbol{\eta}_{123456}e^{-i\boldsymbol{p}\cdot\boldsymbol{y}/\mu}\\
\times&e^{-i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_4|^2)t/2\mu}\int_0^t d\tau_1 e^{i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_2|^2)\tau_1/2\mu}\int_0^{\tau_1}d\tau_2 e^{i(|\boldsymbol{\eta}_2|^2-|\boldsymbol{\eta}_3|^2)\tau_2/2\mu}\\
\times&e^{i\boldsymbol{\eta}_1\cdot(\boldsymbol{x}+\boldsymbol{y}/2-\boldsymbol{\xi}_1)/\mu}e^{i\boldsymbol{\eta}_2\cdot(\boldsymbol{\xi}_1-\boldsymbol{\xi}_2)/\mu}e^{i\boldsymbol{\eta}_3\cdot(\boldsymbol{\xi}_2-\boldsymbol{\xi}_3)/\mu}e^{-i\boldsymbol{\eta}_4\cdot(\boldsymbol{x}-\boldsymbol{y}/2-\boldsymbol{\xi}_4)/\mu}\\
\times&\Pi(\boldsymbol{\eta}_5)e^{i\boldsymbol{\eta}_5\cdot(\boldsymbol{\xi}_1-\boldsymbol{\xi}_2)}n((\boldsymbol{\xi}_3+\boldsymbol{\xi}_4)/2,\boldsymbol{\eta}_6)e^{i\boldsymbol{\eta}_6\cdot(\boldsymbol{\xi}_3-\boldsymbol{\xi}_4)/\mu}.
\end{aligned}
$$

Integration of this expression with respect to $\boldsymbol{y}$, $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ yields

$$
\begin{aligned}
&\mathbb{E}\left[w^{\mu}(\psi_2^{\mu},\psi_0^{\mu})\right]\\
=&\frac{1}{-\mu^2(2\pi\mu)^{2d}}\int d\boldsymbol{\xi}_{34}d\boldsymbol{\eta}_{123456}\\
\times&e^{-i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_4|^2)t/2\mu}\int_0^t d\tau_1 e^{i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_2|^2)\tau_1/2\mu}\int_0^{\tau_1}d\tau_2 e^{i(|\boldsymbol{\eta}_2|^2-|\boldsymbol{\eta}_3|^2)\tau_2/2\mu}\\
\times&e^{i(\boldsymbol{\eta}_1\cdot\boldsymbol{x}-\boldsymbol{\eta}_3\cdot\boldsymbol{\xi}_3-\boldsymbol{\eta}_4\cdot\boldsymbol{x}+\boldsymbol{\eta}_4\cdot\boldsymbol{\xi}_4+\boldsymbol{\eta}_6\cdot\boldsymbol{\xi}_3-\boldsymbol{\eta}_6\cdot\boldsymbol{\xi}_4)/\mu}\\
\times&\delta((\boldsymbol{\eta}_1+\boldsymbol{\eta}_4)/2-\boldsymbol{p})\delta(\boldsymbol{\eta}_1-\boldsymbol{\eta}_2-\boldsymbol{\eta}_5)\delta(\boldsymbol{\eta}_3-\boldsymbol{\eta}_2-\boldsymbol{\eta}_5)\\
\times&\Pi(\boldsymbol{\eta}_5)n((\boldsymbol{\xi}_3+\boldsymbol{\xi}_4)/2,\boldsymbol{\eta}_6).
\end{aligned}
$$

We change the variables as

$$
\boldsymbol{X}=\frac{\boldsymbol{\xi}_3+\boldsymbol{\xi}_4}{2},\quad \boldsymbol{Y}=\boldsymbol{\xi}_3-\boldsymbol{\xi}_4,
$$

and carry out the integration with respect to $\boldsymbol{Y}$ to get

$$
\begin{aligned}
&\mathbb{E}\left[w^{\mu}(\psi_2^{\mu},\psi_0^{\mu})\right]\\
=&\frac{1}{-\mu^2(2\pi\mu)^{d}}\int d\boldsymbol{X}d\boldsymbol{\eta}_{123456}\\
\times&e^{-i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_4|^2)t/2\mu}\int_0^t d\tau_1 e^{i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_2|^2)\tau_1/2\mu}\int_0^{\tau_1}d\tau_2 e^{i(|\boldsymbol{\eta}_2|^2-|\boldsymbol{\eta}_3|^2)\tau_2/2\mu}\\
\times&e^{i(\boldsymbol{\eta}_1-\boldsymbol{\eta}_4)\cdot(\boldsymbol{x}-\boldsymbol{X})/\mu}\delta((\boldsymbol{\eta}_3+\boldsymbol{\eta}_4)/2-\boldsymbol{\eta}_6)\\
\times&\delta((\boldsymbol{\eta}_1+\boldsymbol{\eta}_4)/2-\boldsymbol{p})\delta(\boldsymbol{\eta}_1-\boldsymbol{\eta}_2-\boldsymbol{\eta}_5)\delta(\boldsymbol{\eta}_3-\boldsymbol{\eta}_2-\boldsymbol{\eta}_5)\\
\times&\Pi(\boldsymbol{\eta}_5)n(\boldsymbol{X},\boldsymbol{\eta}_6).
\end{aligned}
$$

We understand that $\boldsymbol{\eta}_1=\boldsymbol{\eta}_3$ and $\boldsymbol{\eta}_6=\boldsymbol{p}$ hold in the integrand. Integration with respect to $\boldsymbol{\eta}_3$ and $\boldsymbol{\eta}_6$ yields

$$
\mathbb{E}\left[w^{\mu}(\psi_2^{\mu},\psi_0^{\mu})\right]
$$

$$
\begin{aligned}
&=\frac{1}{-\mu^2(2\pi\mu)^d}\int d\boldsymbol{X}d\boldsymbol{\eta}_{1245}\\
&\times e^{-i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_4|^2)t/2\mu}\int_0^t d\tau_1 e^{i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_2|^2)\tau_1/2\mu}\int_0^{\tau_1} d\tau_2 e^{i(|\boldsymbol{\eta}_2|^2-|\boldsymbol{\eta}_1|^2)\tau_2/2\mu}\\
&\times e^{i(\boldsymbol{\eta}_1-\boldsymbol{\eta}_4)\cdot(\boldsymbol{x}-\boldsymbol{X})/\mu}\delta((\boldsymbol{\eta}_1+\boldsymbol{\eta}_4)/2-\boldsymbol{p})\delta(\boldsymbol{\eta}_1-\boldsymbol{\eta}_2-\boldsymbol{\eta}_5)\\
&\times\Pi(\boldsymbol{\eta}_5)n(\boldsymbol{X},\boldsymbol{p}).
\end{aligned}
$$

We use (113a) to derive

$$
\int_0^t d\tau_1 e^{i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_2|^2)\tau_1/2\mu}\int_0^{\tau_1} d\tau_2 e^{i(|\boldsymbol{\eta}_2|^2-|\boldsymbol{\eta}_1|^2)\tau_2/2\mu}=\pi\mu t\delta\left(\frac{|\boldsymbol{\eta}_1|^2}{2}-\frac{|\boldsymbol{\eta}_2|^2}{2}\right)+o(\mu).
$$

We also use (113c) to derive

$$
\begin{aligned}
&\frac{1}{(2\pi\mu)^d}\int d\boldsymbol{X}d\boldsymbol{\eta}_4 e^{-i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_4|^2)t/2\mu}e^{i(\boldsymbol{\eta}_1-\boldsymbol{\eta}_4)\cdot(\boldsymbol{x}-\boldsymbol{X})/\mu}\delta((\boldsymbol{\eta}_1+\boldsymbol{\eta}_4)/2-\boldsymbol{p})n(\boldsymbol{X},\boldsymbol{p})\\
=&\delta(\boldsymbol{\eta}_1-\boldsymbol{p})n(\boldsymbol{X},\boldsymbol{p})+\mathcal{O}(\mu,t).
\end{aligned}
$$

Consequently, we obtain

$$
\mathbb{E}\left[w^\mu(\psi_2^\mu,\psi_0^\mu)\right]=-\frac{t}{2\mu}\int d\boldsymbol{\eta}\sigma(\boldsymbol{p},\boldsymbol{\eta})n(\boldsymbol{x},\boldsymbol{p})+o\left(\frac{t}{\mu}\right). \tag{114}
$$

Because $w^\mu(\psi_0^\mu,\psi_2^\mu)=[w^\mu(\psi_2^\mu,\psi_0^\mu)]^\dagger$, and σ and n are real functions, we also have

$$
\mathbb{E}\left[w^\mu(\psi_0^\mu,\psi_2^\mu)\right]=-\frac{t}{2\mu}\int d\boldsymbol{\eta}\sigma(\boldsymbol{p},\boldsymbol{\eta})n(\boldsymbol{x},\boldsymbol{p})+o\left(\frac{t}{\mu}\right). \tag{115}
$$

Finally, we consider $\mathbb{E}[w^\mu(\psi_1^\mu,\psi_1^\mu)]$. From (16) and (18), we have

$$
\begin{aligned}
w^\mu(\psi_1^\mu,\psi_1^\mu)=&\frac{1}{\mu^2(2\pi\mu)^d}\int_0^t d\tau_1\int_0^t d\tau_2\int d\boldsymbol{y}d\boldsymbol{\xi}_{1234}e^{-i\boldsymbol{p}\cdot\boldsymbol{y}/\mu}\\
&\times G^\mu\left(\boldsymbol{x}+\frac{\boldsymbol{y}}{2}-\boldsymbol{\xi}_1,t-\tau_1\right)V^\mu(\boldsymbol{\xi}_1)G^\mu(\boldsymbol{\xi}_1-\boldsymbol{\xi}_2,\tau_1)\psi^\mu(\boldsymbol{\xi}_2,0)\\
&\times G^{\mu\dagger}\left(\boldsymbol{x}-\frac{\boldsymbol{y}}{2}-\boldsymbol{\xi}_3,t-\tau_2\right)V^\mu(\boldsymbol{\xi}_3)G^{\mu\dagger}(\boldsymbol{\xi}_3-\boldsymbol{\xi}_4,\tau_2)\psi^{\mu\dagger}(\boldsymbol{\xi}_4,0).
\end{aligned}
$$

Taking the ensemble average, introducing the Fourier integrals, and integrating some variables in the same manner as the previous case, we derive

$$
\begin{aligned}
&\mathbb{E}\left[w^\mu(\psi_1^\mu,\psi_1^\mu)\right]\\
=&\frac{1}{\mu^2(2\pi\mu)^d}\int d\boldsymbol{X}d\boldsymbol{\eta}_{123456}
\end{aligned}
$$

$$\times e^{-i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_3|^2)t/2\mu}\int_0^t d\tau_1 e^{i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_2|^2)\tau_1/2\mu}\int_0^t d\tau_2 e^{-i(|\boldsymbol{\eta}_3|^3-|\boldsymbol{\eta}_4|^2)\tau_2/2\mu}$$
$$\times e^{i(\boldsymbol{\eta}_1-\boldsymbol{\eta}_3)\cdot(\boldsymbol{x}-\boldsymbol{X})/\mu}\delta((\boldsymbol{\eta}_2+\boldsymbol{\eta}_4)/2-\boldsymbol{\eta}_6)$$
$$\times\delta((\boldsymbol{\eta}_1+\boldsymbol{\eta}_3)/2-\boldsymbol{p})\delta(\boldsymbol{\eta}_1-\boldsymbol{\eta}_2-\boldsymbol{\eta}_5)\delta(\boldsymbol{\eta}_3-\boldsymbol{\eta}_4-\boldsymbol{\eta}_5)$$
$$\times\Pi(\boldsymbol{\eta}_5)n(\boldsymbol{X},\boldsymbol{\eta}_6).$$

Applying (113c), understanding $\boldsymbol{\eta}_1=\boldsymbol{\eta}_3$ and $\boldsymbol{\eta}_2=\boldsymbol{\eta}_4$ in the integrand, using (113b), and Integrating all the possible variables, we finally obtain

$$\mathbb{E}\left[w^\mu(\psi_1^\mu,\psi_1^\mu)\right]=\frac{t}{\mu}\int d\boldsymbol{\eta}_2\sigma(\boldsymbol{p},\boldsymbol{\eta}_2)n(\boldsymbol{x},\boldsymbol{\eta}_2)+o\left(\frac{t}{\mu}\right). \tag{116}$$

C.2 Quadratic terms in the Hamiltonian

We compute the terms in $\mathcal{H}_S$ quadratic in λ, originating from four expressions,

$$\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^\mu(\psi_1^\mu,\psi_0^\mu)(\boldsymbol{x}_1,\boldsymbol{p}_1)w^\mu(\psi_0^\mu,\psi_1^\mu)(\boldsymbol{x}_2,\boldsymbol{p}_2)] \tag{117a}$$

$$\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^\mu(\psi_0^\mu,\psi_1^\mu)(\boldsymbol{x}_1,\boldsymbol{p}_1)w^\mu(\psi_1^\mu,\psi_0^\mu)(\boldsymbol{x}_2,\boldsymbol{p}_2)] \tag{117b}$$

$$\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^\mu(\psi_1^\mu,\psi_0^\mu)(\boldsymbol{x}_1,\boldsymbol{p}_1)w^\mu(\psi_1^\mu,\psi_0^\mu)(\boldsymbol{x}_2,\boldsymbol{p}_2)] \tag{117c}$$

$$\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^\mu(\psi_0^\mu,\psi_1^\mu)(\boldsymbol{x}_1,\boldsymbol{p}_1)w^\mu(\psi_0^\mu,\psi_1^\mu)(\boldsymbol{x}_2,\boldsymbol{p}_2)]. \tag{117d}$$

Among these, the two pairs, (117a)-(117b) and (117c)-(117d), are complex conjugate, respectively. Therefore, we need to compute only two expressions. Although the number of factors involved in the integration is greater than those in the linear terms, the computation procedures are largely the same. For the sake of conciseness, we denote time t instead of Δt.

From (16) and (18), we write the Wigner transforms in the integrand of (117a) as

$$\begin{aligned}w^\mu(\psi_1^\mu,\psi_0^\mu)(\boldsymbol{x}_1,\boldsymbol{p}_1)=&\frac{1}{i\mu(2\pi\mu)^d}\int_0^t d\tau d\boldsymbol{y}d\boldsymbol{\xi}_{123}e^{-i\boldsymbol{p}_1\cdot\boldsymbol{y}/\mu}\\&\times G^\mu\left(\boldsymbol{x}_1+\frac{\boldsymbol{y}}{2}-\boldsymbol{\xi}_1,t-\tau\right)V^\mu(\boldsymbol{\xi}_1)G^\mu(\boldsymbol{\xi}_1-\boldsymbol{\xi}_2,\tau)\psi^\mu(\boldsymbol{\xi}_2,0)\\&\times G^{\mu\dagger}\left(\boldsymbol{x}_1-\frac{\boldsymbol{y}}{2}-\boldsymbol{\xi}_3,t\right)\psi^{\mu\dagger}(\boldsymbol{\xi}_3,0)\end{aligned}$$

$$
\begin{aligned}
w^{\mu}(\psi_0^{\mu},\psi_1^{\mu})(\boldsymbol{x}_2,\boldsymbol{p}_2) =& \frac{1}{-i\mu(2\pi\mu)^d}\int_0^t d\tau d\boldsymbol{y} d\boldsymbol{\xi}_{123} e^{-i\boldsymbol{p}_2\cdot\boldsymbol{y}/\mu}\\
&\times G^{\mu}\left(\boldsymbol{x}_2+\frac{\boldsymbol{y}}{2}-\boldsymbol{\xi}_1,t\right)\psi^{\mu}(\boldsymbol{\xi}_1,0)\\
&\times G^{\mu\dagger}\left(\boldsymbol{x}_2-\frac{\boldsymbol{y}}{2}-\boldsymbol{\xi}_2,t-\tau\right)V^{\mu}(\boldsymbol{\xi}_2)G^{\mu\dagger}(\boldsymbol{\xi}_2-\boldsymbol{\xi}_3,\tau)\psi^{\mu\dagger}(\boldsymbol{\xi}_3,0).
\end{aligned}
$$

Taking the ensemble average of the product of these expressions, it would be straightforward to have

$$
\begin{aligned}
&\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^{\mu}(\psi_1^{\mu},\psi_0^{\mu})(\boldsymbol{x}_1,\boldsymbol{p}_1)w^{\mu}(\psi_0^{\mu},\psi_1^{\mu})(\boldsymbol{x}_2,\boldsymbol{p}_2)]\\
=&\frac{1}{\mu^2(2\pi\mu)^{4d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}d\boldsymbol{X}_{12}d\boldsymbol{\eta}_{123456789}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\\
&\times e^{-i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_3|^2)t/2\mu}\int_0^t d\tau_1 e^{i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_2|^2)\tau_1/2\mu}\\
&\times e^{-i(|\boldsymbol{\eta}_4|^2-|\boldsymbol{\eta}_5|^2)t/2\mu}\int_0^t d\tau_2 e^{-i(|\boldsymbol{\eta}_5|^2-|\boldsymbol{\eta}_6|^2)\tau_2/2\mu}\\
&\times\delta(\boldsymbol{\eta}_1/2+\boldsymbol{\eta}_3/2-\boldsymbol{p}_1)\delta(\boldsymbol{\eta}_4/2+\boldsymbol{\eta}_5/2-\boldsymbol{p}_2)\delta(\boldsymbol{\eta}_1-\boldsymbol{\eta}_2-\boldsymbol{\eta}_7)\delta(\boldsymbol{\eta}_5-\boldsymbol{\eta}_6-\boldsymbol{\eta}_7)\\
&\times\delta(\boldsymbol{\eta}_2/2+\boldsymbol{\eta}_6/2-\boldsymbol{\eta}_8)\delta(\boldsymbol{\eta}_3/2+\boldsymbol{\eta}_4/2-\boldsymbol{\eta}_9)\\
&\times e^{i(\boldsymbol{\eta}_1-\boldsymbol{\eta}_3)\cdot\boldsymbol{x}_1/\mu}e^{i(\boldsymbol{\eta}_4-\boldsymbol{\eta}_5)\cdot\boldsymbol{x}_2/\mu}e^{-i(\boldsymbol{\eta}_2-\boldsymbol{\eta}_6)\cdot\boldsymbol{X}_1/\mu}e^{-i(\boldsymbol{\eta}_4-\boldsymbol{\eta}_3)\cdot\boldsymbol{X}_2/\mu}\\
&\times\Pi(\boldsymbol{\eta}_7)n(\boldsymbol{X}_1,\boldsymbol{\eta}_8)n(\boldsymbol{X}_2,\boldsymbol{\eta}_9).
\end{aligned}
$$

Applying (113c) and (113a) and integrating all the possible variables yields

$$
\begin{aligned}
&\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^{\mu}(\psi_1^{\mu},\psi_0^{\mu})(\boldsymbol{x}_1,\boldsymbol{p}_1)w^{\mu}(\psi_0^{\mu},\psi_1^{\mu})(\boldsymbol{x}_2,\boldsymbol{p}_2)]\\
=&\frac{t}{\mu(2\pi\mu)^d}\int d\boldsymbol{x}_1 d\boldsymbol{p}_1 d\boldsymbol{\eta}_2\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)^2\sigma(\boldsymbol{p}_1,\boldsymbol{\eta}_2)n(\boldsymbol{x}_1,\boldsymbol{p}_1)n(\boldsymbol{x}_1,\boldsymbol{\eta}_2)+o\left(\frac{t}{\mu(2\pi\mu)^d}\right).
\end{aligned}
\tag{118}
$$

From the the condition of complex conjugate, we also have

$$
\begin{aligned}
&\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^{\mu}(\psi_0^{\mu},\psi_1^{\mu})(\boldsymbol{x}_1,\boldsymbol{p}_1)w^{\mu}(\psi_1^{\mu},\psi_0^{\mu})(\boldsymbol{x}_2,\boldsymbol{p}_2)]\\
=&\frac{t}{\mu(2\pi\mu)^d}\int d\boldsymbol{x}_1 d\boldsymbol{p}_1 d\boldsymbol{\eta}_2\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)^2\sigma(\boldsymbol{p}_1,\boldsymbol{\eta}_2)n(\boldsymbol{x}_1,\boldsymbol{p}_1)n(\boldsymbol{x}_1,\boldsymbol{\eta}_2)+o\left(\frac{t}{\mu(2\pi\mu)^d}\right).
\end{aligned}
\tag{119}
$$

For the computation of (117c), we write,

$$
w^{\mu}(\psi_1^{\mu},\psi_0^{\mu})(\boldsymbol{x}_1,\boldsymbol{p}_1) = \frac{1}{i\mu(2\pi\mu)^d}\int_0^t d\tau d\boldsymbol{y} d\boldsymbol{\xi}_{123} e^{-i\boldsymbol{p}_1\cdot\boldsymbol{y}/\mu}
$$

$$\times G^{\mu}\left(\boldsymbol{x}_1+\frac{\boldsymbol{y}}{2}-\boldsymbol{\xi}_1,t-\tau\right)V^{\mu}(\boldsymbol{\xi}_1)G^{\mu}(\boldsymbol{\xi}_1-\boldsymbol{\xi}_2,\tau)\psi^{\mu}(\boldsymbol{\xi}_2,0)$$
$$\times G^{\mu\dagger}\left(\boldsymbol{x}_1-\frac{\boldsymbol{y}}{2}-\boldsymbol{\xi}_3,t\right)\psi^{\mu\dagger}(\boldsymbol{\xi}_3,0)$$

Taking the ensemble average of the product of this expression, it would be straightforward to have

$$\begin{aligned}
&\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^{\mu}(\psi_1^{\mu},\psi_0^{\mu})(\boldsymbol{x}_1,\boldsymbol{p}_1)w^{\mu}(\psi_1^{\mu},\psi_0^{\mu})(\boldsymbol{x}_2,\boldsymbol{p}_2)]\\
=&\frac{1}{-\mu^2(2\pi\mu)^{4d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}d\boldsymbol{X}_{12}d\boldsymbol{\eta}_{123456789}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\\
&\times e^{-i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_3|^2)t/2\mu}\int_0^t d\tau_1 e^{i(|\boldsymbol{\eta}_1|^2-|\boldsymbol{\eta}_2|^2)\tau_1/2\mu}\\
&\times e^{-i(|\boldsymbol{\eta}_4|^2-|\boldsymbol{\eta}_6|^2)t/2\mu}\int_0^t d\tau_2 e^{i(|\boldsymbol{\eta}_4|^2-|\boldsymbol{\eta}_5|^2)\tau_2/2\mu}\\
&\times\delta(\boldsymbol{\eta}_1/2+\boldsymbol{\eta}_3/2-\boldsymbol{p}_1)\delta(\boldsymbol{\eta}_4/2+\boldsymbol{\eta}_6/2-\boldsymbol{p}_2)\delta(\boldsymbol{\eta}_1-\boldsymbol{\eta}_2-\boldsymbol{\eta}_7)\delta(\boldsymbol{\eta}_4-\boldsymbol{\eta}_5+\boldsymbol{\eta}_7)\\
&\times\delta(\boldsymbol{\eta}_2/2+\boldsymbol{\eta}_6/2-\boldsymbol{\eta}_8)\delta(\boldsymbol{\eta}_5/2+\boldsymbol{\eta}_3/2-\boldsymbol{\eta}_9)\\
&\times e^{i(\boldsymbol{\eta}_1-\boldsymbol{\eta}_3)\cdot\boldsymbol{x}_1/\mu}e^{i(\boldsymbol{\eta}_4-\boldsymbol{\eta}_6)\cdot\boldsymbol{x}_2/\mu}e^{-i(\boldsymbol{\eta}_2-\boldsymbol{\eta}_6)\cdot\boldsymbol{X}_1/\mu}e^{-i(\boldsymbol{\eta}_5-\boldsymbol{\eta}_3)\cdot\boldsymbol{X}_2/\mu}\\
&\times\Pi(\boldsymbol{\eta}_7)n(\boldsymbol{X}_1,\boldsymbol{\eta}_8)n(\boldsymbol{X}_2,\boldsymbol{\eta}_9).
\end{aligned}$$

Applying (113c) and (113a) and integrating all the possible variables yields

$$\begin{aligned}
&\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^{\mu}(\psi_1^{\mu},\psi_0^{\mu})(\boldsymbol{x}_1,\boldsymbol{p}_1)w^{\mu}(\psi_1^{\mu},\psi_0^{\mu})(\boldsymbol{x}_2,\boldsymbol{p}_2)]\\
=&-\frac{t}{\mu(2\pi\mu)^d}\int d\boldsymbol{x}_1d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x},\boldsymbol{p}_2)\sigma(\boldsymbol{p}_1,\boldsymbol{p}_2)n(\boldsymbol{x},\boldsymbol{p}_1)n(\boldsymbol{x},\boldsymbol{p}_2)+o\left(\frac{t}{\mu(2\pi\mu)^d}\right).
\end{aligned}\tag{120}$$

Finally, from the the condition of complex conjugate, we have

$$\begin{aligned}
&\frac{1}{(2\pi\mu)^{2d}}\int d\boldsymbol{x}_{12}d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x}_2,\boldsymbol{p}_2)\mathbb{E}[w^{\mu}(\psi_0^{\mu},\psi_1^{\mu})(\boldsymbol{x}_1,\boldsymbol{p}_1)w^{\mu}(\psi_0^{\mu},\psi_1^{\mu})(\boldsymbol{x}_2,\boldsymbol{p}_2)]\\
=&-\frac{t}{\mu(2\pi\mu)^d}\int d\boldsymbol{x}_1d\boldsymbol{p}_{12}\lambda(\boldsymbol{x}_1,\boldsymbol{p}_1)\lambda(\boldsymbol{x},\boldsymbol{p}_2)\sigma(\boldsymbol{p}_1,\boldsymbol{p}_2)n(\boldsymbol{x},\boldsymbol{p}_1)n(\boldsymbol{x},\boldsymbol{p}_2)+o\left(\frac{t}{\mu(2\pi\mu)^d}\right).
\end{aligned}\tag{121}$$

D Derivation of the Hamiltonian in the diffusive limit

In this appendix, we show how to obtain the large deviation Hamiltonian Eq. (58) from Eq. (42). Since we are dealing with the scattering term that is local

in position space, one can omit the dependence on $\boldsymbol{x}$. As explained in the main text, the important point is to show how the diffusion kernel $\Sigma[n]$ transforms when only infinitesimal deviation from incoming wave vector are allowed by the cross section. For this specific purpose, we introduce a parameter ν and write the typical correlation length of V as μ/ν. Our strategy is to expand the Hamiltonian in terms of ν.

Because the potential spectrum $\Pi(\boldsymbol{p})$ is supposed to have a finite support with the extent of $\mathcal{O}(\nu)$, it is convenient to rewrite the cross section as $\sigma(\boldsymbol{p}_1, \boldsymbol{p}_2) = \nu^{-d}\tilde{\sigma}((\boldsymbol{p}_1 - \boldsymbol{p}_2)/\nu;\ \boldsymbol{p}_1)$. For any test functions, f and g, the diffusive kernel (39) satisfies

$$\begin{aligned}&\int d\boldsymbol{p}_1 d\boldsymbol{p}_2 f(\boldsymbol{p}_1)\Sigma(\boldsymbol{p}_1, \boldsymbol{p}_2)g(\boldsymbol{p}_2)\\ =&c\int d\boldsymbol{p}d\boldsymbol{q} f(\boldsymbol{p})n(\boldsymbol{p})\tilde{\sigma}(\boldsymbol{q};\ \boldsymbol{p})n(\boldsymbol{p}-\nu\boldsymbol{q})\left[g(\boldsymbol{p}) - g(\boldsymbol{p}-\nu\boldsymbol{q})\right].\end{aligned} \tag{122}$$

Taylor expanding the integrand with respect to ν, one obtains

$$\begin{aligned}&\int d\boldsymbol{p}_1 d\boldsymbol{p}_2 f(\boldsymbol{p}_1)\Sigma(\boldsymbol{p}_1, \boldsymbol{p}_2)g(\boldsymbol{p}_2)\\ =&c\int d\boldsymbol{p} f(\boldsymbol{p})n(\boldsymbol{p})\Bigg\{\nu n(\boldsymbol{p})\sum_{n=1}^{d}\left(\int dq_n \tilde{\sigma}(\boldsymbol{q};\ \boldsymbol{p})q_n\right)\partial_{p_n} g(\boldsymbol{p})\\ &-\nu^2\sum_{n=1}^{d}\sum_{m=1}^{d}\left(\int dq_n dq_m \tilde{\sigma}(\boldsymbol{q};\ \boldsymbol{p})q_n q_m\right)\partial_{p_n} n(\boldsymbol{p})\partial_{p_m} g(\boldsymbol{p})\\ &-\frac{\nu^2}{2}n(\boldsymbol{p})\sum_{n=1}^{d}\sum_{m=1}^{d}\left(\int dq_n dq_m \tilde{\sigma}(\boldsymbol{q};\ \boldsymbol{p})q_n q_m\right)\partial_{p_n}\partial_{p_m} g(\boldsymbol{p}) + \mathcal{O}\left(\nu^3\right)\Bigg\}.\end{aligned} \tag{123}$$

To evaluate the integrations with respect to $\boldsymbol{q}$, we shall exponentiate the Dirac-δ in the definition of the cross section (23) and use the inverse Fourier transform of Π in (57). The resulting expression is

$$\begin{aligned}&\tilde{\sigma}(\boldsymbol{q};\ \boldsymbol{p}) = \frac{1}{\nu(2\pi)^d}\int d\boldsymbol{y}\int_{\mathbb{R}} ds \mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{y}} R^{\nu}\left(s\boldsymbol{p} + \boldsymbol{y} - \frac{\nu}{2}s\boldsymbol{q}\right)\\ &\text{and}\quad R^{\nu}(\boldsymbol{y}) = \int d\boldsymbol{p}\Pi(\boldsymbol{p})\mathrm{e}^{i\boldsymbol{p}\cdot\boldsymbol{y}/\nu},\end{aligned} \tag{124}$$

where $R^{\nu}(\boldsymbol{y}) = R(\boldsymbol{y}/\nu)$ (see Eq. (57)) is a scaled correlation function of the random potential with the correlation length of $\mathcal{O}(1)$. Again Taylor expanding Eq. (124) with respect to ν, one gets

$$c\nu\int d\boldsymbol{q}\tilde{\sigma}(\boldsymbol{q};\ \boldsymbol{p})q_n = -\nu\sum_{m=1}^{d}\partial_{p_m} D^{\nu}_{nm}(\boldsymbol{p}) + \mathcal{O}(\nu^2)$$

$$c\nu^2 \int d\boldsymbol{q}\tilde{\sigma}(\boldsymbol{q};\ \boldsymbol{p})q_n q_m = 2\nu D^{\nu}_{nm}(\boldsymbol{p}) + \mathcal{O}(\nu^2)$$

with $\boldsymbol{D}^{\nu}(\boldsymbol{p}) = -(c/2)\int_{\mathbb{R}} \nabla\otimes\nabla R^{\nu}(s\boldsymbol{p})ds$. Here, $\int_{\mathbb{R}} \nabla R^{\nu}(s\boldsymbol{p})ds$ vanishes because of the point symmetry in R^{ν}. Inserting these expressions into (123), one obtains

$$\begin{aligned}
&\int d\boldsymbol{p}_1 d\boldsymbol{p}_2 f(\boldsymbol{p}_1)\Sigma(\boldsymbol{p}_1,\boldsymbol{p}_2)g(\boldsymbol{p}_2)\\
&= -c\nu \int d\boldsymbol{p} f n \sum_{n=1}^{d}\sum_{m=1}^{d} \{n\partial_{p_n} D^{\nu}_{nm}\partial_{p_m} g + 2D^{\nu}_{nm}\partial_{p_n} n\partial_{p_m} g + nD^{\nu}_{nm}\partial_{p_n}\partial_{p_m} g\}\\
&+\mathcal{O}(\nu^2). \qquad (125)
\end{aligned}$$

Evidently from this result, the dominant term is of order $\mathcal{O}(\nu)$ in the present scaling. Therefore, in order for diffusion in wave-vector space to be comparable to the free propagation in position space, one needs to rescale the position coordinate. However, for the sake of simplicity here, we shall set $\nu = 1$ while ignoring $\mathcal{O}(\nu^2)$ terms. Consequently, after reorganizing products of derivatives and performing integration by parts, we confirm that (125) transforms into Eq. (56).

References

[1] Bouchet, F.: Is the Boltzmann equation reversible? A large deviation perspective on the irreversibility paradox. Journal of Statistical Physics **181**(2), 515–550 (2020)

[2] Guioth, J., Bouchet, F., Eyink, G.L.: Path large deviations for the kinetic theory of weak turbulence. Journal of Statistical Physics **189**(20) (2022)

[3] Cardy, J., Falkovich, G., Gawedzki, K.: Non-equilibrium Statistical Mechanics and Turbulence. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, U.K. (2008). https://doi.org/10.1017/CBO9780511812149

[4] Chetrite, R., Delannoy, J.-Y., Gawedzki, K.: Kraichnan flow in a square: an example of integrable chaos. Journal of Statistical Physics **126**(6), 1165–1200 (2007)

[5] Gawedzki, K.: Stochastic processes in turbulent transport. arXiv preprint arXiv:0806.1949 (2008)

[6] Gawedzki, K.: Fluctuation relations in stochastic thermodynamics. arXiv preprint arXiv:1308.1518 (2013)

[7] Chetrite, R., Gawędzki, K.: Eulerian and Lagrangian pictures of non-equilibrium diffusions. Journal of Statistical Physics **137**(5), 890–916 (2009)

[8] Bouchet, F., Gawędzki, K., Nardini, C.: Perturbative calculation of quasi-potential in non-equilibrium diffusions: a mean-field example. Journal of Statistical Physics **163**(5), 1157–1210 (2016)

[9] Ryzhik, L., Papanicolaou, G., Keller, J.B.: Transport equations for elastic and other waves in random media. Wave Motion **24**(4), 327–370 (1996)

[10] Bal, G.: Kinetics of scalar wave fields in random media. Wave Motion **43**(2), 132–157 (2005)

[11] Powell, J., Vanneste, J.: Transport equations for waves in randomly perturbed Hamiltonian systems, with application to Rossby waves. Wave Motion **42**(4), 289–308 (2005)

[12] Danioux, E., Vanneste, J.: Near-inertial-wave scattering by random flows. Physical Review Fluids **1**(3), 033701 (2016)

[13] Savva, M.A., Vanneste, J.: Scattering of internal tides by barotropic quasigeostrophic flows. Journal of Fluid Mechanics **856**, 504–530 (2018)

[14] Savva, M.A., Kafiabad, H.A., Vanneste, J.: Inertia-gravity-wave scattering by three-dimensional geostrophic turbulence. Journal of Fluid Mechanics **916** (2021)

[15] Boury, S., Bühler, O., Shatah, J.: Wave scattering and irreversible wave capturing by two-dimensional turbulent flow. in preparation for PRE.

[16] Hasselmann, K.: On the non-linear energy transfer in a gravity-wave spectrum Part 1. General theory. Journal of Fluid Mechanics **12**(4), 481–500 (1962)

[17] Kafiabad, H.A., Savva, M.A., Vanneste, J.: Diffusion of inertia-gravity waves by geostrophic turbulence. Journal of Fluid Mechanics **869** (2019)

[18] Bôas, A.B.V., Young, W.R.: Directional diffusion of surface gravity wave action by ocean macroturbulence. Journal of Fluid Mechanics **890** (2020)

[19] Dong, W., Bühler, O., Smith, K.S.: Frequency diffusion of waves by unsteady flows. Journal of Fluid Mechanics **905** (2020)

[20] Whalen, C.B., MacKinnon, J.A., Talley, L.D., Waterhouse, A.F.: Estimating the mean diapycnal mixing using a finescale strain parameterization. Journal of Physical Oceanography **45**(4), 1174–1188 (2015)

[21] Monsalve, E., Brunet, M., Gallet, B., Cortet, P.-P.: Quantitative experimental observation of weak inertial-wave turbulence. Physical Review Letters **125**(25), 254502 (2020)

[22] Rodda, C., Savaro, C., Davis, G., Reneuve, J., Augier, P., Sommeria, J., Valran, T., Viboud, S., Mordant, N.: Experimental observations of internal wave turbulence transition in a stratified fluid. Physical Review Fluids **7**(9), 094802 (2022)

[23] Michel, G., Bonnefoy, F., Ducrozet, G., Falcon, E.: Statistics of rogue waves in isotropic wave fields. Journal of Fluid Mechanics **943**, 26 (2022)

[24] Falcon, E., Michel, G., Prabhudesai, G., Cazaubiel, A., Berhanu, M., Mordant, N., Aumaître, S., Bonnefoy, F.: Saturation of the inverse cascade in surface gravity-wave turbulence. Physical Review Letters **125**(13), 134501 (2020)

[25] Cazaubiel, A., Mawet, S., Darras, A., Grosjean, G., Van Loon, J., Dorbolo, S., Falcon, E.: Wave turbulence on the surface of a fluid in a high-gravity environment. Physical Review Letters **123**(24), 244501 (2019)

[26] Banks, J., Buckmaster, T., Korotkevich, A., Kovačič, G., Shatah, J.: Direct verification of the kinetic description of wave turbulence for finite-size systems dominated by interactions among groups of six waves. Physical review letters **129**(3), 034101 (2022)

[27] Zhu, Y., Semisalov, B., Krstulovic, G., Nazarenko, S.: Testing wave turbulence theory for the Gross-Pitaevskii system. Physical Review E **106**(1), 014205 (2022)

[28] Du, R., Bühler, O.: Domain dependence of wave turbulence theory for the Majda-McLaughlin-Tabak (MMT) model. To be submitted to PRE. (2022)

[29] Buckmaster, T., Germain, P., Hani, Z., Shatah, J.: Onset of the wave turbulence description of the longtime behavior of the nonlinear Schrödinger equation. Inventiones mathematicae **225**(3), 787–855 (2021)

[30] Deng, Y., Hani, Z.: On the derivation of the wave kinetic equation for NLS. In: Forum of Mathematics, Pi, vol. 9 (2021). Cambridge University Press

[31] Collot, C., Germain, P.: On the derivation of the homogeneous kinetic wave equation. arXiv preprint arXiv:1912.10368 (2019)

[32] Collot, C., Germain, P.: Derivation of the homogeneous kinetic wave equation: longer time scales. arXiv preprint arXiv:2007.03508 (2020)

[33] Faou, E.: Linearized wave turbulence convergence results for three-wave systems. Communications in Mathematical Physics **378**(2), 807–849 (2020)

[34] Ampatzoglou, I., Collot, C., Germain, P.: Derivation of the kinetic wave equation for quadratic dispersive problems in the inhomogeneous setting. arXiv preprint arXiv:2107.11819 (2021)

[35] Deng, Y., Hani, Z.: Propagation of chaos and the higher order statistics in the wave kinetic theory. arXiv (2022). https://doi.org/10.48550/arXiv.2110.04565

[36] Deng, Y., Ionescu, A.D., Pusateri, F.: On the wave turbulence theory of 2D gravity waves, I: deterministic energy estimates. arXiv preprint arXiv:2211.10826 (2022)

[37] Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Macroscopic fluctuation theory. Reviews of Modern Physics **87**(2), 593 (2015)

[38] Léonard, C.: On large deviations for particle systems associated with spatially homogeneous Boltzmann type equations. Probability Theory and Related Fields **101**(1), 1–44 (1995)

[39] Rezakhanlou, F.: Large deviations from a kinetic limit. The Annals of Probability **26**(3), 1259–1340 (1998)

[40] Bodineau, T., Gallagher, I., Saint-Raymond, L., Simonella, S.: Fluctuation theory in the Boltzmann–Grad limit. Journal of Statistical Physics **180**(1), 873–895 (2020)

[41] Heydecker, D.: Large deviations of Kac's conservative particle system and energy non-conserving solutions to the Boltzmann equation: a counterexample to the predicted rate function. arXiv preprint arXiv:2103.14550 (2021)

[42] Basile, G., Benedetto, D., Bertini, L., Caglioti, E.: Asymptotic probability of energy increasing solutions to the homogeneous Boltzmann equation. arXiv preprint arXiv:2202.07311 (2022)

[43] Feliachi, O., Bouchet, F.: Dynamical large deviations for plasmas below the Debye length and the Landau equation. Journal of Statistical Physics **183**(3), 1–58 (2021)

[44] Feliachi, O., Bouchet, F.: Dynamical large deviations for homogeneous systems with long range interactions and the Balescu–Guernsey–Lenard equation. Journal of Statistical Physics **186**(2) (2022)

[45] Mielke, A., Peletier, M.A., Renger, D.M.: On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. Potential Analysis **41**(4), 1293–1327 (2014)

[46] Bodineau, T., Gallagher, I., Saint-Raymond, L., Simonella, S.: Statistical dynamics of a hard sphere gas: fluctuating Boltzmann equation and large deviations. arXiv preprint arXiv:2008.10403 (2020)

[47] Dematteis, G., Grafke, T., Onorato, M., Vanden-Eijnden, E.: Experimental evidence of hydrodynamic instantons: the universal route to rogue waves. Physical Review X **9**(4), 041057 (2019)

[48] Garrido, M., Grande, R., Kurianski, K., Staffilani, G.: Large Deviations Principle for the Cubic NLS Equation (2021). https://doi.org/10.48550/arXiv.2110.15748

[49] Erdős, L., Yau, H.-T.: Linear Boltzmann equation as the weak coupling limit of a random Schrödinger equation. Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences **53**(6), 667–735 (2000)

[50] Bal, G., Komorowski, T., Ryzhik, L.: Kinetic limits for waves in a random medium. Kinetic & Related Models **3**(4), 529 (2010)

[51] Gu, Y., Ryzhik, L.: The random Schrödinger equation: homogenization in time-dependent potentials. Multiscale Modeling & Simulation **14**(1), 323–363 (2016). https://doi.org/10.1137/15M1024986

[52] Zhang, N., Bal, G.: Homogenization of the schrödinger equation with large, random potential. Stochastics and Dynamics **14**(01), 1350013 (2014). https://doi.org/10.1142/S0219493713500135

[53] Lee, H.-W.: Theory and application of the quantum phase-space distribution functions. Physics Reports **259**(3), 147–211 (1995)

[54] Nazarenko, S.: Wave Turbulence. Lecture Notes in Physics, vol. 825. Springer Berlin Heidelberg, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-15942-8

[55] Eyink, G.L., Shi, Y.-K.: Kinetic wave turbulence. Physica D: Nonlinear Phenomena **241**(18), 1487–1511 (2012). https://doi.org/10.1016/j.physd.2012.05.015

[56] Lagendijk, A., Tiggelen, B.v., Wiersma, D.S.: Fifty years of Anderson localization. Physics Today **62**(8), 24–29 (2009). https://doi.org/10.1063/1.3206091

[57] Sheng, P.: Introduction to Wave Scattering, Localization, and Mesoscopic Phenomena, 2nd edn. Springer series in materials science, vol. 88. Springer, Berlin; New York (2006)

[58] Lanford, O.E.: Time evolution of large classical systems. In: Ehlers, J., Hepp, K., Weidenmüller, H.A., Beiglböck, W., Moser, J. (eds.) Dynamical Systems, Theory and Applications vol. 38, pp. 1–111. Springer, Berlin, Heidelberg (1975). https://doi.org/10.1007/3-540-07171-7_1

[59] Gallagher, I., Saint-Raymond, L., Texier, B.: From Newton to Boltzmann: Hard Spheres and Short-range Potentials. European Mathematical Society Zürich, Switzerland, ??? (2013)

[60] Freidlin, M.I., Wentzell, A.D.: Random Perturbations of Dynamical Systems. Grundlehren der mathematischen Wissenschaften, vol. 260. Springer, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-25847-3

[61] Graham, R.: Macroscopic potentials, bifurcations and noise in dissipative systems. In: Moss, F., McClintock, P.V.E. (eds.) Noise in Nonlinear Dynamical Systems, 1st edn., pp. 225–278. Cambridge University Press, Cambridge, U.K. (1989). https://doi.org/10.1017/CBO9780511897818.009

[62] Guioth, J., Onuki, Y., Bouchet, F.: Path large deviations for inhomogeneous weak wave turbulence. To be submitted to J. Stat. Phys.

[63] Touchette, H.: The large deviation approach to statistical mechanics. Physics Reports **478**(1-3), 1–69 (2009)

[64] Polkovnikov, A.: Phase space representation of quantum dynamics. Annals of Physics **325**(8), 1790–1852 (2010)

[65] Onuki, Y.: Quasi-local method of wave decomposition in a slowly varying medium. Journal of Fluid Mechanics **883** (2020)