Transport equations for waves in randomly perturbed Hamiltonian systems, with application to Rossby waves

J. M. Powell and J. Vanneste [*]

School of Mathematics, University of Edinburgh, Edinburgh, UK

Abstract

Transport equations for linear waves in randomly perturbed media are derived for a general class of Hamiltonian systems. These equations govern the evolution in position–wavevector space of the energy density of the waves; they describe the modulations caused by large-scale inhomogeneities of the medium, and the scattering due to weak random perturbations with spatial scales comparable to the wavelengths. Particular attention is paid to the conservative properties of the transport equations inherited from the Hamiltonian structure of the systems considered. The general results are used to derive transport equations for Rossby waves propagating in a two-layer model with random topography.

[*] J. Vanneste, School of Mathematics, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom.

Email addresses: `jpowell@maths.ed.ac.uk`, `J.Vanneste@ed.ac.uk` (J. M. Powell and J. Vanneste).

1 Introduction

The propagation of waves in inhomogeneous conservative media is, in some circumstances, governed by a relatively simple equation describing the transport in the position–wavevector phase space $(\boldsymbol{x}, \boldsymbol{k})$ of a scalar energy density $a(\boldsymbol{x}, \boldsymbol{k}, t)$. Transport equations of this type are obtained for waves in media whose deterministic properties vary slowly compared to the typical wavelengths, and they can include the scattering effect of random perturbations of the medium. If $\epsilon \ll 1$ characterises the scale separation between the medium and the waves, and if the random perturbations have an $O(\epsilon^{1/2})$ amplitude and spatial scales comparable to the wavelengths, then transport equations of the form

$$\partial_t a(\boldsymbol{x}, \boldsymbol{k}, t) + \nabla_{\boldsymbol{k}} \omega(\boldsymbol{x}, \boldsymbol{k}) \cdot \nabla_{\boldsymbol{x}} a(\boldsymbol{x}, \boldsymbol{k}, t) - \nabla_{\boldsymbol{x}} \omega(\boldsymbol{x}, \boldsymbol{k}) \cdot \nabla_{\boldsymbol{k}} a(\boldsymbol{x}, \boldsymbol{k}, t)$$
$$= \int \sigma(\boldsymbol{x}, \boldsymbol{k}, \boldsymbol{k}') a(\boldsymbol{x}, \boldsymbol{k}', t) \, \mathrm{d}\boldsymbol{k}' - \Sigma(\boldsymbol{x}, \boldsymbol{k}) a(\boldsymbol{x}, \boldsymbol{k}, t) \tag{1.1}$$

are obtained at leading-order in ϵ. Here, ω is the frequency, σ is the differential scattering cross-section, i.e, the rate at which wave energy with wavevector $\boldsymbol{k}'$ is converted to wavevector $\boldsymbol{k}$, and

$$\Sigma(\boldsymbol{x}, \boldsymbol{k}) = \int \sigma(\boldsymbol{x}, \boldsymbol{k}', \boldsymbol{k}) \, \mathrm{d}\boldsymbol{k}'$$

is the total scattering cross-section. Equation (1.1) is valid when the dispersion relation relating frequency ω to k is single valued. When it has several branches, i.e. when several wave modes coexist, their propagation is governed by a set of transport equations similar to (1.1) coupled by the scattering terms on the right-hand side. The degeneracies in the dispersion relation that are associated with polarised waves lead to a further complication, with additional transport equations describing the evolution of the polarisation state. The validity of (1.1) also requires that the random perturbations be not too

strongly anisotropic: for highly anisotropic perturbations, and in particular for one-dimensional perturbations, localization effects prevent wave propagation [e.g. 1, 6].

Transport equations of the form (1.1) or its generalisations have been derived for many different types of waves using a variety of (essentially equivalent) techniques. In view of the common features of these derivations, it is of interest to consider more general theories which provide transport equations for large classes of systems. Ryzhik, Papanicolaou & Keller [10] developed such a general theory for a class of symmetric hyperbolic systems, and they presented applications to several systems. We refer the reader to that paper for further background on transport equations, for details of applications, and for references.

Ryzhik et al. [10] used a powerful formalism, based on the Wigner function, which lends itself naturally to further generalisations. Guo & Wang [3] extended their results by considering vector Schrödinger equations of the form $i\partial_t \boldsymbol{u} = H(\boldsymbol{x}, \partial_{\boldsymbol{x}})\boldsymbol{u}$, where $H(\boldsymbol{x}, \partial_{\boldsymbol{x}})$ is a pseudodifferential operator. The present paper continues this generalisation effort and derives transport equations for waves in systems governed by pseudodifferential equations of the form

$$\partial_t \boldsymbol{u} = J(\boldsymbol{x}, \partial_{\boldsymbol{x}})H(\boldsymbol{x}, \partial_{\boldsymbol{x}})\boldsymbol{u}, \tag{1.2}$$

where J and H are, respectively, skew-adjoint and self-adjoint pseudodifferential (matrix) operators. This is the most general form for linear systems conserving an energy-like quadratic quantity, here

$$\mathcal{H} = \frac{1}{2}\int \boldsymbol{u}^* H(\boldsymbol{x}, \partial_{\boldsymbol{x}})\boldsymbol{u}\,\mathrm{d}\boldsymbol{x}, \tag{1.3}$$

3

with a non-degenerate $H(\boldsymbol{x}, \partial_{\boldsymbol{x}})$. Written as

$$\partial_t \boldsymbol{u} = J(\boldsymbol{x}, \partial_{\boldsymbol{x}})\frac{\delta \mathcal{H}}{\delta \boldsymbol{u}^*}, \tag{1.4}$$

(1.2) is recognised as a general linear non-canonical Hamiltonian (or Poisson) system. In particular, for canonical Hamiltonian systems, $\boldsymbol{u}$ is real and J is given by the canonical $2n \times 2n$ symplectic matrix

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix},$$

where I is the n-dimensional identity matrix, while for (conservative) Schrödinger equations $J = -\mathrm{i}I$.

We emphasise that linear Poisson systems of the form (1.2) emerge naturally from nonlinear Poisson systems given by

$$\partial_t \boldsymbol{v} = \mathcal{J}(\boldsymbol{v}, \partial_{\boldsymbol{x}})\frac{\delta \mathcal{E}}{\delta \boldsymbol{v}}, \tag{1.5}$$

where $\mathcal{J}$ is a skew-adjoint operator which may depend on (pseudodifferentials of) $\boldsymbol{v}$ and satisfies a Jacobi identity, and $\mathcal{E}$ a functional of $\boldsymbol{v}$ [e.g. 8, Chapter 6]. This is obvious when the linearisation is about $\boldsymbol{v} = 0$, but perhaps less so when it is about a nontrivial equilibrium $\boldsymbol{v} = \boldsymbol{V}(\boldsymbol{x})$. In the latter case, the equilibrium condition is crucial in ensuring that

$$\mathcal{J}(\boldsymbol{V}, \partial_{\boldsymbol{x}}) \left.\frac{\delta \mathcal{E}}{\delta \boldsymbol{v}}\right|_{\boldsymbol{v}=\boldsymbol{V}(\boldsymbol{x})} = 0.$$

In turn, this ensures that $\delta\mathcal{E}/\delta\boldsymbol{v} = 0$ for $\boldsymbol{v} = \boldsymbol{V}(\boldsymbol{x})$ or, if $\mathcal{J}$ is degenerate, that a (Casimir) functional $\mathcal{C}(\boldsymbol{v})$ satisfying $\mathcal{J}(\boldsymbol{v}, \partial_{\boldsymbol{x}})\delta C/\delta\boldsymbol{v} = 0$ for all $\boldsymbol{v}$ can

be found such that

$$\left.\frac{\delta(\mathcal{E} + \mathcal{C})}{\delta \boldsymbol{v}}\right|_{\boldsymbol{v}=\boldsymbol{V}(\boldsymbol{x})} = 0. \tag{1.6}$$

Introducing the expansion $\boldsymbol{v} = \boldsymbol{V} + \boldsymbol{u}$ into (1.5), linearising for $\boldsymbol{u}$ and taking (1.6) into account leads to a system of the form (1.2), with

$$J(\boldsymbol{x}, \partial_{\boldsymbol{x}}) = \mathcal{J}(\boldsymbol{V}(\boldsymbol{x}), \partial_{\boldsymbol{x}}) \quad \text{and} \quad H(\boldsymbol{x}, \partial_{\boldsymbol{x}}) = \left.\frac{\delta^2(\mathcal{E} + \mathcal{C})}{\delta \boldsymbol{v}^2}\right|_{\boldsymbol{v}=\boldsymbol{V}(\boldsymbol{x})} \tag{1.7}$$

[see, e.g., 7, for details]. The conserved quadratic invariant $\mathcal{H}$ in (1.3) associated with the operator $H(\boldsymbol{x}, \partial_{\boldsymbol{x}})$ constructed in this way is sometimes referred to as pseudoenergy; here we follow another standard usage and term it wave energy.

We emphasise the construction (1.5)–(1.7) because it provides a important route leading to conservative linear systems of the form (1.2) with spatially dependent coefficients. In particular, since the equations governing the dynamics of inviscid fluids are naturally cast in the Poisson form (1.5) [7], the propagation of waves in shear flows can be conveniently studied within the formalism of the present paper.

The plan of this paper is as follows. The detailed structure of the Hamiltonian linear systems considered is described in section 2: the scaling assumptions are specified, and the relevant expansion of (1.2) in powers of the small parameter ϵ is given. Particular attention is paid to the manner in which the Hamiltonian structure of (1.2) is reflected in the terms of this expansion. The correlation tensors which define the random perturbation are also defined in section 2. Section 3 provides the definition of the Wigner matrix, and the derivation of an evolution equation for this matrix. A multiple-scale expansion of this equation leads to the desired transport equations; this is described in section

4. The reader not interested in derivation details can skip most of section 4: subsection 4.1 provides all the definitions needed to use the complete form of the transport equations and scattering cross-sections which are given by (4.25)–(4.27). The application to Rossby waves is treated in section 5. The paper concludes with a Discussion in section 6. Details of the (sometimes lengthy) derivations are given in an Appendix.

2 Formulation

We consider evolution equations of the form (1.2) for a n-dimensional complex variable $\boldsymbol{u}(\boldsymbol{x}, t)$, with $\boldsymbol{x} \in \mathbb{R}^d$. In this expression, the operators $J(\boldsymbol{x}, \partial_{\boldsymbol{x}})$ and $H(\boldsymbol{x}, \partial_{\boldsymbol{x}})$ are the pseudodifferential operators associated with the corresponding $n \times n$ matrices $J(\boldsymbol{x}, \mathrm{i}\boldsymbol{k})$ and $H(\boldsymbol{x}, \mathrm{i}\boldsymbol{k})$. We use the standard (Kohn–Nirenberg) correspondence in which the differentiations are on the right of the $\boldsymbol{x}$-dependence. Thus, for instance,

$$H(\boldsymbol{x}, \partial_{\boldsymbol{x}})\boldsymbol{u}(\boldsymbol{x}, t) = \int H(\boldsymbol{x}, \mathrm{i}\boldsymbol{k})\hat{\boldsymbol{u}}(\boldsymbol{k}, t)\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{k},$$

where $\hat{\boldsymbol{u}}(\boldsymbol{k}, t)$ denotes the Fourier transform of $\boldsymbol{u}(\boldsymbol{x}, t)$, with

$$\boldsymbol{u}(\boldsymbol{x}, t) = \int \hat{\boldsymbol{u}}(\boldsymbol{k}, t)\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{k} \quad \text{and} \quad \hat{\boldsymbol{u}}(\boldsymbol{k}, t) = \frac{1}{(2\pi)^d}\int \boldsymbol{u}(\boldsymbol{x}, t)\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}.$$

We have chosen this interpretation of pseudodifferential operators rather than the Weyl correspondence chosen by Guo & Wang [3] for its simplicity, even though it makes our derivation of the transport equation somewhat less elegant; it is straightforward, if sometimes cumbersome, to translate between the two interpretations [see, e.g., 2]. With the assumption that $J(\boldsymbol{x}, \partial_{\boldsymbol{x}})$ and $H(\boldsymbol{x}, \partial_{\boldsymbol{x}})$ are, respectively, skew-adjoint and self-adjoint, the conservation of the wave energy (1.3) is readily established.

Our interest is in wave-like solutions to (1.2) with wavelengths much shorter than the typical scale of deterministic variations of J and H. To make this explicit, we introduce the scale-separation parameter $\epsilon \ll 1$ and, redefining time and space variables according $t \mapsto t/\epsilon$ and $\boldsymbol{x} \mapsto \boldsymbol{x}/\epsilon$, we rewrite (1.2) in terms of slow variables as

$$\epsilon \partial_t \boldsymbol{u} = J(\boldsymbol{x}, \epsilon \partial_{\boldsymbol{x}}) H(\boldsymbol{x}, \epsilon \partial_{\boldsymbol{x}}) \boldsymbol{u}. \tag{2.1}$$

In addition to the $O(1)$ slowly-varying contribution to J and H indicated in (2.1), we consider a second contribution that is random with zero average and varies over the spatial scale of the waves $\boldsymbol{x}/\epsilon$. The distinguished limit in which this random contribution has an effect comparable to that of the slow variations of J and H is achieved when the amplitude of the random perturbation scales like $\epsilon^{1/2}$ [cf. 10]. We adopt this scaling here and, correspondingly, expand J and H according to

$$J(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) = J_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) + \epsilon^{1/2} J_{1/2}(\boldsymbol{x}/\epsilon, \mathrm{i}\boldsymbol{k}) + \epsilon J_1(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) + \cdots \tag{2.2}$$

$$\text{and} \quad H(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) = H_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) + \epsilon^{1/2} H_{1/2}(\boldsymbol{x}/\epsilon, \mathrm{i}\boldsymbol{k}) + \epsilon H_1(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) + \cdots, \tag{2.3}$$

where $J_{1/2}$ and $H_{1/2}$ are zero-average random matrices. For simplicity we have assumed here that $J_{1/2}$ and $H_{1/2}$ depend on space only through $\boldsymbol{x}/\epsilon$; an additional, slow dependence on $\boldsymbol{x}$ could in fact be included without significant changes. We note that the (deterministic) $O(\epsilon)$ terms J_1 and H_1 appear automatically when J and H are expanded; these terms are crucial, in particular to ensure wave-energy conservation, and they are related to J_0 and H_0. Indeed, the self-adjointess of H and skew-adjointness of J imply at leading order that

$$H_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) = H_0^*(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) \quad \text{and} \quad J_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) = -J_0^*(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}),$$

where $*$ denotes the Hermitian adjoint, and at $O(\epsilon)$ that

$$H_1(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) - H_1^*(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) = -\mathrm{i}\nabla_{\boldsymbol{x}} \cdot \nabla_{\boldsymbol{k}} H_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) \tag{2.4}$$
$$\text{and} \quad J_1(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) + J_1^*(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) = -\mathrm{i}\nabla_{\boldsymbol{x}} \cdot \nabla_{\boldsymbol{k}} J_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}). \tag{2.5}$$

The last two equalities determine the skew-adjoint and self-adjoint parts of H_1 and J_1, respectively [cf. 15].

The transport equations to be derived turn out to depend on the random operators $J_{1/2}$ and $H_{1/2}$ only through their correlation tensors, which we now define. Denoting by

$$\hat{J}_{1/2}(\boldsymbol{l}, \boldsymbol{k}) = \frac{1}{(2\pi)^d} \int J_{1/2}(\boldsymbol{\xi}, \mathrm{i}\boldsymbol{k}) \mathrm{e}^{-\mathrm{i}\boldsymbol{l}\cdot\boldsymbol{\xi}} \, \mathrm{d}\boldsymbol{\xi}$$

the Fourier transform of $J_{1/2}(\boldsymbol{x}/\epsilon, \mathrm{i}\boldsymbol{k})$ with respect to its first argument, and similarly for $H_{1/2}$, we define the correlation 4-tensors J, H and K by

$$\langle \hat{J}_{1/2}^{\alpha\beta}(\boldsymbol{l}, \boldsymbol{k}+\boldsymbol{m}) \hat{J}_{1/2}^{\gamma\delta}(\boldsymbol{m}, \boldsymbol{n}) \rangle = -\mathsf{J}^{\alpha\beta\gamma\delta}(\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{n})\delta(\boldsymbol{l}+\boldsymbol{m}) \tag{2.6}$$
$$\langle \hat{H}_{1/2}^{\alpha\beta}(\boldsymbol{l}, \boldsymbol{k}+\boldsymbol{m}) \hat{H}_{1/2}^{\gamma\delta}(\boldsymbol{m}, \boldsymbol{n}) \rangle = \mathsf{H}^{\alpha\beta\gamma\delta}(\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{n})\delta(\boldsymbol{l}+\boldsymbol{m}) \tag{2.7}$$
$$\langle \hat{J}_{1/2}^{\alpha\beta}(\boldsymbol{l}, \boldsymbol{k}+\boldsymbol{m}) \hat{H}_{1/2}^{\gamma\delta}(\boldsymbol{m}, \boldsymbol{n}) \rangle = \mathsf{K}^{\alpha\beta\gamma\delta}(\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{n})\delta(\boldsymbol{l}+\boldsymbol{m}), \tag{2.8}$$

where $\langle \cdot \rangle$ denotes ensemble average and $\alpha, \beta, \gamma, \delta = 1, 2, \cdots, n$ denote the components. We have here assumed that $J_{1/2}$ and $H_{1/2}$ are statistically homogeneous, hence the presence of the Dirac distributions on the right-hand sides. For future reference, we note that the skew-adjointness and self-adjointness of $J_{1/2}$ and $H_{1/2}$ imply that

$$\hat{J}_{1/2}^*(\boldsymbol{l}, \boldsymbol{k}) = -\hat{J}_{1/2}(-\boldsymbol{l}, \boldsymbol{k}+\boldsymbol{l}) \quad \text{and} \quad \hat{H}_{1/2}^*(\boldsymbol{l}, \boldsymbol{k}) = \hat{H}_{1/2}(-\boldsymbol{l}, \boldsymbol{k}+\boldsymbol{l}). \tag{2.9}$$

Simple manipulations then show that

$$\mathsf{J}^{\alpha\beta\gamma\delta*}(\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{n}) = \mathsf{J}^{\delta\gamma\beta\alpha}(\boldsymbol{l}, \boldsymbol{n}, \boldsymbol{k}), \quad \mathsf{H}^{\alpha\beta\gamma\delta*}(\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{n}) = \mathsf{H}^{\delta\gamma\beta\alpha}(\boldsymbol{l}, \boldsymbol{n}, \boldsymbol{k}),$$
$$\text{and} \quad \mathsf{K}^{\alpha\beta\gamma\delta*}(\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{n}) = -\mathsf{K}^{\beta\alpha\delta\gamma}(-\boldsymbol{l}, \boldsymbol{k}-\boldsymbol{l}, \boldsymbol{n}-\boldsymbol{l}). \tag{2.10}$$

8

3 Wigner matrix

Following Ryzhik et al. [10], we define the $n \times n$ (Hermitian) Wigner matrix associated with $\boldsymbol{u}(\boldsymbol{x}, t)$ by

$$W_\epsilon(\boldsymbol{x}, \boldsymbol{k}, t) = \frac{1}{(2\pi)^d} \int \boldsymbol{u}(\boldsymbol{x} - \epsilon \boldsymbol{y}/2, t) \boldsymbol{u}^*(\boldsymbol{x} + \epsilon \boldsymbol{y}/2, t) \mathrm{e}^{\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{y}} \, \mathrm{d}\boldsymbol{y}, \tag{3.1}$$

and we note the dual expression

$$W_\epsilon(\boldsymbol{x}, \boldsymbol{k}, t) = \frac{1}{\epsilon^d} \int \hat{\boldsymbol{u}}(\boldsymbol{k}/\epsilon + \boldsymbol{l}/2, t) \hat{\boldsymbol{u}}^*(\boldsymbol{k}/\epsilon - \boldsymbol{l}/2, t) \, \mathrm{e}^{\mathrm{i}\boldsymbol{l} \cdot \boldsymbol{x}} \, \mathrm{d}\boldsymbol{l} \tag{3.2}$$

in terms of the Fourier transform $\hat{\boldsymbol{u}}$ of $\boldsymbol{u}$. It is useful to relate the conserved wave energy $\mathcal{H}$ given by (1.3) to the Wigner matrix; a short calculation detailed in Appendix A.1 gives

$$\mathcal{H} = \frac{1}{2} \operatorname{Tr} \iint H(\boldsymbol{x}, \mathrm{i}\boldsymbol{k} + \epsilon \partial_{\boldsymbol{x}}/2) W_\epsilon(\boldsymbol{x}, \boldsymbol{k}, t) \, \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{k}, \tag{3.3}$$

where Tr denotes the trace.

Let us now derive an evolution equation for the Wigner matrix. From (2.1) and (3.1), we have

$$\epsilon \partial_t W_\epsilon = S_0 + \epsilon^{1/2} S_{1/2} + \epsilon S_1 + \cdots, \tag{3.4}$$

where the terms on the right-hand side can be written as

$$S_p = \frac{1}{(2\pi)^d} \int \left[\boldsymbol{u}(\boldsymbol{x} - \epsilon \boldsymbol{y}/2, t) \left(L_p \boldsymbol{u} \right)^* (\boldsymbol{x} + \epsilon \boldsymbol{y}/2, t) \right. \tag{3.5}$$

$$\left. + (L_p \boldsymbol{u}) (\boldsymbol{x} - \epsilon \boldsymbol{y}/2, t) \boldsymbol{u}(\boldsymbol{x} + \epsilon \boldsymbol{y}/2, t) \right] \mathrm{e}^{\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{y}} \, \mathrm{d}\boldsymbol{y}, \tag{3.6}$$

for $p = 0, 1/2, 1$. Here, the pseudodifferential operators $L_p = L_p(\boldsymbol{x} \pm \epsilon \boldsymbol{y}/2, \epsilon \partial_{\boldsymbol{x}})$ are defined by the corresponding matrices

9

$$L_0(\boldsymbol{x}, i\boldsymbol{k}) = J_0(\boldsymbol{x}, i\boldsymbol{k}) H_0(\boldsymbol{x}, i\boldsymbol{k}), \quad L_{1/2}(\boldsymbol{x}, i\boldsymbol{k}) = J_0(\boldsymbol{x}, i\boldsymbol{k}) H_{1/2}(\boldsymbol{x}/\epsilon, i\boldsymbol{k}) + J_{1/2}(\boldsymbol{x}/\epsilon, i\boldsymbol{k}) H_0(\boldsymbol{x}, i\boldsymbol{k}),$$
$$\text{and} \quad L_1(\boldsymbol{x}, i\boldsymbol{k}) = J_0(\boldsymbol{x}, i\boldsymbol{k}) H_1(\boldsymbol{x}, i\boldsymbol{k}) + J_1(\boldsymbol{x}, i\boldsymbol{k}) H_0(\boldsymbol{x}, i\boldsymbol{k}) + J_{1/2}(\boldsymbol{x}/\epsilon, i\boldsymbol{k}) H_{1/2}(\boldsymbol{x}/\epsilon, i\boldsymbol{k}).$$

Note that although the matrix L_0 is independent of ϵ, the term S_0 is not, as a result of the ϵ dependence of the arguments $\boldsymbol{x} \pm \epsilon \boldsymbol{y}/2$. Furthermore, $L_{1/2}$ and L_1 depend on ϵ through the argument $\boldsymbol{x}/\epsilon$ of $J_{1/2}$ and $H_{1/2}$.

To reduce (3.4) to a closed equation for the Wigner matrix, we use the Fourier transforms of the matrices L_p, defined by

$$\hat{L}_p(\boldsymbol{l}, \boldsymbol{k}) = \frac{1}{(2\pi)^d} \int L_p(\boldsymbol{x}, i\boldsymbol{k}) e^{-i\boldsymbol{l}\cdot\boldsymbol{x}} \, d\boldsymbol{x}.$$

Introducing this transform into (3.6) along with the Fourier transforms of $\boldsymbol{u}(\boldsymbol{x} \pm \epsilon \boldsymbol{y}/2, t)$ leads, on using (3.2), to

$$S_p = \frac{1}{(2\pi)^d} \iiint \left[W_\epsilon(\boldsymbol{y}, \boldsymbol{k} - \epsilon\boldsymbol{l}/2, t) \hat{L}_p^*(\boldsymbol{l}, \boldsymbol{k} - \epsilon\boldsymbol{l}/2 - \epsilon\boldsymbol{m}/2) \right. \tag{3.7}$$
$$\left. + \hat{L}_p(-\boldsymbol{l}, \boldsymbol{k} + \epsilon\boldsymbol{l}/2 + \epsilon\boldsymbol{m}/2) W_\epsilon(\boldsymbol{y}, \boldsymbol{k} + \epsilon\boldsymbol{l}/2, t) \right] e^{i[\boldsymbol{m}\cdot(\boldsymbol{x}-\boldsymbol{y}) - \boldsymbol{l}\cdot\boldsymbol{x}]} \, d\boldsymbol{y} \, d\boldsymbol{l} \, d\boldsymbol{m}.$$

It is now straightforward, if tedious, to expand S_p in powers of ϵ. Expanding S_0 yields

$$S_0 = Q_{00} W_\epsilon(\boldsymbol{x}, \boldsymbol{k}, t) + \epsilon Q_{01} W_\epsilon(\boldsymbol{x}, \boldsymbol{k}, t) + O(\epsilon^2), \tag{3.8}$$

where

$$Q_{00} W = W L_0^* + L_0 W,$$
$$Q_{01} W = \frac{i}{2} \left[\nabla_{\boldsymbol{x}} W \cdot \nabla_{\boldsymbol{k}} L_0^* - \nabla_{\boldsymbol{k}} W \cdot \nabla_{\boldsymbol{x}} L_0^* - W \nabla_{\boldsymbol{k}} \cdot \nabla_{\boldsymbol{x}} L_0^* \right]$$
$$\quad - \frac{i}{2} \left[\nabla_{\boldsymbol{k}} L_0 \cdot \nabla_{\boldsymbol{x}} W - \nabla_{\boldsymbol{x}} L_0 \cdot \nabla_{\boldsymbol{k}} W - \nabla_{\boldsymbol{k}} \cdot \nabla_{\boldsymbol{x}} L_0 W \right],$$

for any $W = W(\boldsymbol{x}, \boldsymbol{k}, t)$. In the expansion

$$S_{1/2} = Q_{1/2,0} W_\epsilon(\boldsymbol{x}, \boldsymbol{k}, t) + O(\epsilon), \tag{3.9}$$

10

of $S_{1/2}$, $Q_{1/2,0}$ is given by

$$Q_{1/2,0}W = 2^d \int \left[W(\boldsymbol{x}, \boldsymbol{k}+\boldsymbol{l}, t) V^*(\boldsymbol{x}, -2\boldsymbol{l}, \boldsymbol{k}+\boldsymbol{l}) + V(\boldsymbol{x}, 2\boldsymbol{l}, \boldsymbol{k}-\boldsymbol{l}) W(\boldsymbol{x}, \boldsymbol{k}-\boldsymbol{l}, t) \right] \mathrm{e}^{2\mathrm{i}\boldsymbol{l}\cdot\boldsymbol{x}/\epsilon} \, \mathrm{d}\boldsymbol{l},$$

where we have introduced the matrix

$$V(\boldsymbol{x}, \boldsymbol{l}, \boldsymbol{k}) = \hat{J}_{1/2}(\boldsymbol{l}, \boldsymbol{k}) H_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) + J_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k} + \mathrm{i}\boldsymbol{l}) \hat{H}_{1/2}(\boldsymbol{l}, \boldsymbol{k}). \tag{3.10}$$

To obtain this result, the scaling relating the Fourier transform with respect to $\boldsymbol{x}$ (involved in the definition of $\hat{L}_p$) to the Fourier transform with respect to $\boldsymbol{x}/\epsilon$ (involved in the definition of $\hat{J}_{1/2}$ and $\hat{H}_{1/2}$) must be taken into account.

Finally, we have

$$S_1 = Q_{10} W_\epsilon(\boldsymbol{x}, \boldsymbol{k}, t) + O(\epsilon), \tag{3.11}$$

where

$$\begin{aligned}
Q_{10}W = {} & W \left[J_1(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) H_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) + J_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) H_1(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) \right]^* \\
& + \left[J_1(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) H_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) + J_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) H_1(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) \right] W \\
& + 4^d \iint \left\{ W(\boldsymbol{x}, \boldsymbol{k}+\boldsymbol{l}+\boldsymbol{m}, t) \left[\hat{J}_{1/2}(-2\boldsymbol{m}, \boldsymbol{k}-\boldsymbol{l}+\boldsymbol{m}) \hat{H}_{1/2}(-2\boldsymbol{l}, \boldsymbol{k}+\boldsymbol{l}+\boldsymbol{m}) \right]^* \right. \\
& \left. + \left[\hat{J}_{1/2}(2\boldsymbol{m}, \boldsymbol{k}+\boldsymbol{l}-\boldsymbol{m}) \hat{H}_{1/2}(2\boldsymbol{l}, \boldsymbol{k}-\boldsymbol{l}-\boldsymbol{m}) \right] W(\boldsymbol{x}, \boldsymbol{k}-\boldsymbol{l}-\boldsymbol{m}, t) \right\} \mathrm{e}^{2\mathrm{i}(\boldsymbol{l}+\boldsymbol{m})\cdot\boldsymbol{x}/\epsilon} \, \mathrm{d}\boldsymbol{l}\,\mathrm{d}\boldsymbol{m}.
\end{aligned}$$

Together with (3.8)–(3.11), (3.4) provides a closed evolution equation for the Wigner matrix $W_\epsilon(\boldsymbol{x}, \boldsymbol{k}, t)$, accurate to $O(\epsilon)$, which can be solved perturbatively. Because the random terms in $S_{1/2}$ and S_1 depend on $\boldsymbol{x}/\epsilon$, this requires to use a multiple-scale method. We describe the multiple-scale calculations, which lead to transport equations, in the next section.

11

4 Derivation of the transport equations

We expand the Wigner matrix according to

$$W_\epsilon(\boldsymbol{x}, \boldsymbol{k}, t) = W_0(\boldsymbol{x}, \boldsymbol{k}, t) + \epsilon^{1/2} W_{1/2}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{k}, t) + \epsilon W_1(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{k}, t) + O(\epsilon^{3/2}). \quad (4.1)$$

Here, we have introduced the fast spatial variable $\boldsymbol{\xi} = \boldsymbol{x}/\epsilon$, and we have anticipated the fact that the (deterministic) leading-order approximation W_0 to W_ϵ is independent of $\boldsymbol{\xi}$. When applied to differential operators, the multiple-scale method is efficiently implemented using the substitution $\partial_{\boldsymbol{x}} \mapsto \partial_{\boldsymbol{x}} + \epsilon^{-1} \partial_{\boldsymbol{\xi}}$ which follows from the chain rule [e.g. 4]. The analogous substitution for pseudodifferential operators is conveniently carried out in the Fourier representation. If an operator $Q(\partial_{\boldsymbol{x}})$ is given by

$$Q(\partial_{\boldsymbol{x}}) \boldsymbol{u}(\boldsymbol{x}) = \frac{1}{(2\pi)^d} \iint Q(\mathrm{i}\boldsymbol{m}) \boldsymbol{u}(\boldsymbol{y}) \mathrm{e}^{\mathrm{i}\boldsymbol{m}\cdot(\boldsymbol{x}-\boldsymbol{y})} \, \mathrm{d}\boldsymbol{y} \mathrm{d}\boldsymbol{m}$$

when acting on $\boldsymbol{u}(\boldsymbol{x})$, its multiple-scale version, which we will denote by a tilde for clarity, acts on $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\xi})$ and is given by

$$\tilde{Q}(\partial_{\boldsymbol{x}} + \epsilon^{-1} \partial_{\boldsymbol{\xi}}) \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\xi}) = \frac{1}{(2\pi)^{2d}} \iiiint Q(\mathrm{i}\boldsymbol{m} + \mathrm{i}\boldsymbol{n}/\epsilon) \boldsymbol{u}(\boldsymbol{y}, \boldsymbol{\eta}) \mathrm{e}^{\mathrm{i}[\boldsymbol{m}\cdot(\boldsymbol{x}-\boldsymbol{y}) + \boldsymbol{n}\cdot(\boldsymbol{\xi}-\boldsymbol{\eta})]} \, \mathrm{d}\boldsymbol{y} \mathrm{d}\boldsymbol{m} \mathrm{d}\boldsymbol{\eta} \mathrm{d}\boldsymbol{n}.$$

This provides a simple rule to compute the multiple-scale versions of S_p in (3.7) and, by expansion in powers of ϵ, of the operators $\tilde{Q}_{pq}$.

Introducing (4.1) into (3.4) leads to a sequence of equations for the Wigner matrices W_p, $p = 0, 1/2, 1, \cdots$. We now detail each of these equations, up to that for W_1 whose solvability conditions yield transport equations of the form (1.1).

12

4.1 $O(1)$ equation

At leading order, we find that

$$Q_{00}W_0(\boldsymbol{x}, \boldsymbol{k}, t) = 0. \tag{4.2}$$

To solve this equation, we consider the eigenvalues $-\mathrm{i}\omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})$ and (right) eigenvectors $\boldsymbol{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k})$ of $L_0(\boldsymbol{x}, \boldsymbol{k})$. These satisfy

$$L_0(\boldsymbol{x}, \boldsymbol{k})\boldsymbol{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) = -\mathrm{i}\omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})\boldsymbol{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}), \quad s = 1, 2, \cdots, n \tag{4.3}$$

and respectively define the dispersion relation and polarisation relations for waves propagating in the system (2.1). The sign convention for the frequencies is the usual one: seeking approximate solutions to the leading-order approximation

$$\epsilon\partial_t\boldsymbol{u} = J_0(\boldsymbol{x}, \partial_{\boldsymbol{x}})H_0(\boldsymbol{x}, \partial_{\boldsymbol{x}})\boldsymbol{u}$$

of (2.1) in the WKB form

$$\boldsymbol{u}(\boldsymbol{x}, t) \propto \mathrm{e}^{\mathrm{i}\vartheta(\boldsymbol{x},t)/\epsilon}\,\boldsymbol{e}$$

leads to the eigenvalue problem (4.3), with the usual relations

$$\boldsymbol{k} = \nabla_{\boldsymbol{x}}\vartheta \quad \text{and} \quad \omega = -\partial_t\vartheta.$$

We assume (i) that the frequencies $\omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})$ are all real, and (ii) that they all have multiplicity one. The first assumption, which amounts to the linear stability of the system, is satisfied in particular if $H_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k})$ is sign definite; the second assumption excludes polarised waves. We now review some properties

of the eigenvalue problem (4.3) which will be needed in what follows. The left eigenvectors $\hat{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k})$ of $L_0(\boldsymbol{x}, \boldsymbol{k})$, which satisfy

$$L_0^*(\boldsymbol{x}, \boldsymbol{k})\hat{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) = \mathrm{i}\omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})\hat{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}), \tag{4.4}$$

can be related to the right eigenvectors according to

$$\hat{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) = H_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k})e_{(s)}(\boldsymbol{x}, \boldsymbol{k}). \tag{4.5}$$

The relationship

$$\hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k})J_0(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) = -\mathrm{i}\omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})e^*_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \tag{4.6}$$

follows readily. We choose to normalise the eigenvectors so that the orthogonality between left and right eigenvectors read

$$\hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k})e_{(t)}(\boldsymbol{x}, \boldsymbol{k}) = \delta_{ts}. \tag{4.7}$$

With this normalisation, the n-dimensional identity matrix has the expansion

$$I = \sum_s e_{(s)}(\boldsymbol{x}, \boldsymbol{k})\hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k}). \tag{4.8}$$

Returning to (4.2), we note that the null space of Q_{00} is spanned by the matrices $e_{(s)}(\boldsymbol{x}, \boldsymbol{k})e^*_{(s)}(\boldsymbol{x}, \boldsymbol{k})$. Thus, the solution of (4.2) takes the general form

$$W_0(\boldsymbol{x}, \boldsymbol{k}, t) = \sum_s a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t)e_{(s)}(\boldsymbol{x}, \boldsymbol{k})e^*_{(s)}(\boldsymbol{x}, \boldsymbol{k}). \tag{4.9}$$

for some amplitudes $a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t)$. These amplitudes, whose transport equations we seek to derive, can be interpreted as the wave-energy density of mode s in the $(\boldsymbol{x}, \boldsymbol{k})$ phase space. Indeed, introducing (4.1) into (3.3) and taking (4.7)

14

and (4.9) into account gives the approximation for the wave energy

$$\mathcal{H} = \frac{1}{2}\sum_s \iint a_{(s)}(\boldsymbol{x},\boldsymbol{k},t)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{k} + O(\epsilon^{1/2}). \tag{4.10}$$

4.2 $O(\epsilon^{1/2})$ equation

At $O(\epsilon^{1/2})$, the evolution equation (3.4) for the Wigner matrix gives

$$\tilde{Q}_{00}W_{1/2}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t) + Q_{1/2,0}W_0(\boldsymbol{x},\boldsymbol{k},t) = 0, \tag{4.11}$$

where

$$\tilde{Q}_{00}W_{1/2}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t) = \frac{1}{(2\pi)^d}\iint \Big[W_{1/2}(\boldsymbol{x},\boldsymbol{\eta},\boldsymbol{k},t)L_0^*(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{n}/2)$$
$$+\, L_0(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{n}/2)W_{1/2}(\boldsymbol{x},\boldsymbol{\eta},\boldsymbol{k},t)\Big]\,\mathrm{e}^{\mathrm{i}\boldsymbol{n}\cdot(\boldsymbol{\xi}-\boldsymbol{\eta})}\,\mathrm{d}\boldsymbol{\eta}\mathrm{d}\boldsymbol{n}.$$

To solve this equation, we define the $n \times n$ matrix $F(\boldsymbol{x},\boldsymbol{m},\boldsymbol{k},t)$ by

$$W_{1/2}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t) = \int F(\boldsymbol{x},\boldsymbol{m}/2,\boldsymbol{k},t)\mathrm{e}^{\mathrm{i}\boldsymbol{m}\cdot\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{m},$$

and introduce this expression into (4.11) to find that

$$F(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k},t)L_0^*(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l}) + L_0(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})F(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k},t)$$
$$= -\big[W_0(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l},t)V^*(\boldsymbol{x},-2\boldsymbol{l},\boldsymbol{k}+\boldsymbol{l}) + V(\boldsymbol{x},2\boldsymbol{l},\boldsymbol{k}-\boldsymbol{l})W_0(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l},t)\big]. \tag{4.12}$$

We then expand F according to

$$F(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k},t) = \sum_{s,t}\mathcal{P}_{(s,t)}(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k},t)\boldsymbol{e}_{(s)}(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})\boldsymbol{e}_{(t)}^*(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l}), \tag{4.13}$$

where the scalar coefficients $\mathcal{P}_{(s,t)}$ remain to be determined. This is achieved by introducing the expansion (4.13) into (4.12), left-multiplying by $\hat{\boldsymbol{e}}_{(s)}^*(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})$ and right-multiplying by $\hat{\boldsymbol{e}}_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})$. Calculations detailed in Appendix A.2 yield the result

15

$$\mathrm{i}\left[\omega_{(s)}(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})-\omega_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})\right]\mathcal{P}_{(s,t)}(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k},t)$$

$$=-[a_{(s)}(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l},t)-a_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l},t)]\hat{e}_{(s)}^{*}(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})\hat{J}_{1/2}(2\boldsymbol{l},\boldsymbol{k}-\boldsymbol{l})\hat{e}_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})$$

$$+\mathrm{i}[\omega_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})a_{(s)}(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l},t)-\omega_{(s)}(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})a_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l},t)]$$

$$\times e_{(s)}^{*}(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})\hat{H}_{1/2}(2\boldsymbol{l},\boldsymbol{k}-\boldsymbol{l})e_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l}) \tag{4.14}$$

which completes the determination of $W_{1/2}$.

4.3 O(ϵ) equation

The $O(\epsilon)$ equation for the Wigner matrix reads

$$\partial_{t}W_{0}(\boldsymbol{x},\boldsymbol{k},t)=\tilde{Q}_{00}W_{1}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t)+(Q_{01}+Q_{10})W_{0}(\boldsymbol{x},\boldsymbol{k},t)+\tilde{Q}_{1/2,0}W_{1/2}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t),\tag{4.15}$$

where

$$\tilde{Q}_{1/2,0}W_{1/2}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t)=\left(\frac{2}{\pi}\right)^{d}\iiint\left[W_{1/2}(\boldsymbol{x},\boldsymbol{\eta},\boldsymbol{k}+\boldsymbol{l},t)V^{*}(x,-2\boldsymbol{l},\boldsymbol{k}+\boldsymbol{l}-\boldsymbol{m})\right.$$

$$\left.+V(x,2\boldsymbol{l},\boldsymbol{k}-\boldsymbol{l}+\boldsymbol{m})W_{1/2}(\boldsymbol{x},\boldsymbol{\eta},\boldsymbol{k}-\boldsymbol{l},t)\right]\mathrm{e}^{2\mathrm{i}[\boldsymbol{m}\cdot(\boldsymbol{\xi}-\boldsymbol{\eta})+\boldsymbol{l}\cdot\boldsymbol{\xi}]}\,\mathrm{d}\boldsymbol{\eta}\mathrm{d}\boldsymbol{l}\mathrm{d}\boldsymbol{m}$$

$$=2^{d}\iint[F(\boldsymbol{x},\boldsymbol{m},\boldsymbol{k}+\boldsymbol{l},t)V^{*}(\boldsymbol{x},-2\boldsymbol{l},\boldsymbol{k}+\boldsymbol{l}-\boldsymbol{m})$$

$$+V(\boldsymbol{x},2\boldsymbol{l},\boldsymbol{k}-\boldsymbol{l}+\boldsymbol{m})F(\boldsymbol{x},\boldsymbol{m},\boldsymbol{k}-\boldsymbol{l},t)]\,\mathrm{e}^{2\mathrm{i}(\boldsymbol{m}+\boldsymbol{l})\cdot\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{l}\mathrm{d}\boldsymbol{m}.$$

The transport equations are deduced as solvability conditions for (4.15). We first average this equation and note that

$$\langle\tilde{Q}_{00}W_{1}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t)\rangle=\tilde{Q}_{00}\langle W_{1}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t)\rangle=Q_{00}\langle W_{1}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t)\rangle,$$

since $\tilde{Q}_{00}$ is deterministic. The solvability conditions follow by left- and right-multiplying by $\hat{e}_{(s)}^{*}(\boldsymbol{x},\boldsymbol{k})$ and $\hat{e}_{(s)}(\boldsymbol{x},\boldsymbol{k})$, respectively, for $s=1,2,\cdots,n$. This cancels the first term on the right-hand side of (4.15), leading to

$$\partial_{t}a_{(s)}(\boldsymbol{x},\boldsymbol{k},t)=\hat{e}_{(s)}^{*}(\boldsymbol{x},\boldsymbol{k})\left[(Q_{01}+\langle Q_{10}\rangle)W_{0}(\boldsymbol{x},\boldsymbol{k},t)+\langle\tilde{Q}_{1/2,0}W_{1/2}(\boldsymbol{x},\boldsymbol{\xi},\boldsymbol{k},t)\rangle\right]\hat{e}_{(s)}(\boldsymbol{x},\boldsymbol{k})\tag{4.16}$$

after using (4.7) and (4.9). Because $W_{1/2}$ can be expressed in terms of W_0 and hence of $a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t)$, these n equations are closed and, when simplified, provide the sought transport equations. The simplification starts by considering the deterministic terms, which are obtained by setting $J_{1/2} = H_{1/2} = 0$; in this case, $Q_{1/2,0}$ vanishes and Q_{10} simplifies. We show in Appendix A.3 that (4.16) then reduces to transport equations of the Liouville form

$$\partial_t a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) + \nabla_{\boldsymbol{k}}\omega(\boldsymbol{x}, \boldsymbol{k}) \cdot \nabla_{\boldsymbol{x}} a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) - \nabla_{\boldsymbol{x}}\omega(\boldsymbol{x}, \boldsymbol{k}) \cdot \nabla_{\boldsymbol{k}} a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) = 0. \quad (4.17)$$

We now turn to the random contributions to (4.16). We write the last term as

$$\hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k})\langle \tilde{Q}_{1/2,0} W_{1/2}(\boldsymbol{x}, \boldsymbol{\xi}, \boldsymbol{k}, t)\rangle \hat{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k})$$
$$= 2^d \hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k})\langle \iint V(2\boldsymbol{l}, \boldsymbol{k} - \boldsymbol{l} + \boldsymbol{m}) F(\boldsymbol{x}, \boldsymbol{m}, \boldsymbol{k} - \boldsymbol{l}, t)\mathrm{e}^{2\mathrm{i}(\boldsymbol{m}+\boldsymbol{l})\cdot\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{l}\mathrm{d}\boldsymbol{m}\rangle \hat{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) + \text{c.c.}$$
$$= I_{\mathsf{J}} + I_{\mathsf{H}} + I_{\mathsf{K}} + \text{c.c.}, \quad (4.18)$$

where c.c. denotes the complex conjugate of the previous terms. In the last line of (4.18), we have separated three contributions involving, respectively, products of the form $\hat{J}_{1/2} \otimes \hat{J}_{1/2}$, $\hat{H}_{1/2} \otimes \hat{H}_{1/2}$, and $\hat{J}_{1/2} \otimes \hat{H}_{1/2}$; thus, in terms of the correlation tensors defined in (2.6)-(2.8), I_{J} involves J, I_{H} involves H, and I_{K} involves K. We detail in Appendix A.4 the derivation of each of these terms. The first is found to be

$$I_{\mathsf{J}} = 2\pi \sum_t \int \Xi^{\mathsf{J}}_{(s,t)}(\boldsymbol{x}, \boldsymbol{l}, \boldsymbol{k})\delta\left[\omega_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}) - \omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})\right]$$
$$\times \left[a_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}, t) - a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t)\right]\,\mathrm{d}\boldsymbol{l}, \quad (4.19)$$

where

$$\Xi^{\mathsf{J}}_{(s,t)}(\boldsymbol{x}, \boldsymbol{l}, \boldsymbol{k}) = \mathsf{J}^{\alpha\beta\gamma\delta}(\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{k})\hat{e}^{\alpha*}_{(s)}(\boldsymbol{x}, \boldsymbol{k})\hat{e}^{\beta}_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l})\hat{e}^{\gamma*}_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l})\hat{e}^{\delta}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \quad (4.20)$$

and a summation over the repeated superscripts is understood. The second term is given by

17

$$I_{\mathsf{H}} = 2\pi\omega_{(s)}^2(\boldsymbol{x},\boldsymbol{k}) \sum_t \int \Xi_{(s,t)}^{\mathsf{H}}(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k})\delta\left[\omega_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l}) - \omega_{(s)}(\boldsymbol{x},\boldsymbol{k})\right]$$

$$\times \left[a_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l},t) - a_{(s)}(\boldsymbol{x},\boldsymbol{k},t)\right]\,\mathrm{d}\boldsymbol{l}, \quad (4.21)$$

where

$$\Xi_{(s,t)}^{\mathsf{H}}(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k}) = \mathsf{H}^{\alpha\beta\gamma\delta}(\boldsymbol{l},\boldsymbol{k},\boldsymbol{k})e_{(s)}^{\alpha*}(\boldsymbol{x},\boldsymbol{k})e_{(t)}^{\beta}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})e_{(t)}^{\gamma*}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})e_{(s)}^{\delta}(\boldsymbol{x},\boldsymbol{k}). \quad (4.22)$$

Finally, the third term is given by

$$I_{\mathsf{K}} = 2\pi\omega_{(s)}(\boldsymbol{x},\boldsymbol{k}) \sum_t \int \Xi_{(s,t)}^{\mathsf{K}}(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k})\delta\left[\omega_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l}) - \omega_{(s)}(\boldsymbol{x},\boldsymbol{k})\right]$$

$$\times \left[a_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l},t) - a_{(s)}(\boldsymbol{x},\boldsymbol{k},t)\right]\,\mathrm{d}\boldsymbol{l}$$

$$- \left[a_{(s)}(\boldsymbol{x},\boldsymbol{k},t)\hat{e}_{(s)}^{\alpha*}(\boldsymbol{x},\boldsymbol{k})e_{(s)}^{\delta}(\boldsymbol{x},\boldsymbol{k})\int \mathsf{K}^{\alpha\beta\beta\delta}(\boldsymbol{l},\boldsymbol{k},\boldsymbol{k})\,\mathrm{d}\boldsymbol{l} + \text{c.c.}\right], \quad (4.23)$$

where

$$\Xi_{(s,t)}^{\mathsf{K}}(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k}) = \mathrm{i}\left[\mathsf{K}^{\alpha\beta\gamma\delta}(\boldsymbol{l},\boldsymbol{k},\boldsymbol{k})\hat{e}_{(s)}^{\alpha*}(\boldsymbol{x},\boldsymbol{k})\hat{e}_{(t)}^{\beta}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})e_{(t)}^{\gamma*}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})e_{(s)}^{\delta}(\boldsymbol{x},\boldsymbol{k})\right. \quad (4.24)$$

$$\left. + \mathsf{K}^{\gamma\delta\alpha\beta}(-\boldsymbol{l},\boldsymbol{k}-\boldsymbol{l},\boldsymbol{k}-\boldsymbol{l})e_{(s)}^{\alpha*}(\boldsymbol{x},\boldsymbol{k})e_{(t)}^{\beta}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})\hat{e}_{(t)}^{\gamma*}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})\hat{e}_{(s)}^{\delta}(\boldsymbol{x},\boldsymbol{k})\right]$$

is real.

There is a further random contribution in (4.15), namely that stemming from the random term in $Q_{10}W_0$. This contribution is given by

$$4^d\,\hat{e}_{(s)}^*(\boldsymbol{x},\boldsymbol{k}) \iint \left\{W_0(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l}+\boldsymbol{m},t)\langle\hat{J}_{1/2}(-2\boldsymbol{m},\boldsymbol{k}-\boldsymbol{l}+\boldsymbol{m})\hat{H}_{1/2}(-2\boldsymbol{l},\boldsymbol{k}+\boldsymbol{l}+\boldsymbol{m})\rangle^*\right.$$

$$\left. + \langle\hat{J}_{1/2}(2\boldsymbol{m},\boldsymbol{k}+\boldsymbol{l}-\boldsymbol{m})\hat{H}_{1/2}(2\boldsymbol{l},\boldsymbol{k}-\boldsymbol{l}-\boldsymbol{m})\rangle W_0(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l}-\boldsymbol{m},t)\right\}\mathrm{e}^{2\mathrm{i}(\boldsymbol{l}+\boldsymbol{m})\cdot\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{l}\mathrm{d}\boldsymbol{m}\,\hat{e}_{(s)}(\boldsymbol{x},\boldsymbol{k}).$$

Taking (2.8) into account, this term can be shown to be identical, up to the sign, to the second line in the expression (4.23) of I^{K}. Therefore, combining (4.19)–(4.24) and letting $\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{l}$ leads to the transport equations of the form

$$\partial_t a_{(s)}(\boldsymbol{x},\boldsymbol{k},t) + \nabla_{\boldsymbol{k}}\omega(\boldsymbol{x},\boldsymbol{k})\cdot\nabla_{\boldsymbol{x}}a_{(s)}(\boldsymbol{x},\boldsymbol{k},t) - \nabla_{\boldsymbol{x}}\omega(\boldsymbol{x},\boldsymbol{k})\cdot\nabla_{\boldsymbol{k}}a_{(s)}(\boldsymbol{x},\boldsymbol{k},t)$$

$$= \sum_t \int \sigma_{(s,t)}(\boldsymbol{x}, \boldsymbol{k}, \boldsymbol{k}')a_{(t)}(\boldsymbol{x}, \boldsymbol{k}', t)\mathrm{d}\boldsymbol{k}' - \Sigma_{(s)}a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t). \quad (4.25)$$

The differential scattering cross-section is found to be given by

$$
\begin{aligned}
\sigma_{(s,t)}(\boldsymbol{x}, \boldsymbol{k}, \boldsymbol{k}') = {}& 2\pi\delta\Big[\omega_{(t)}(\boldsymbol{x}, \boldsymbol{k}') - \omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})\Big] \\
& \times \Big[\omega_{(s)}^2(\boldsymbol{x}, \boldsymbol{k})\mathsf{H}^{\alpha\beta\gamma\delta}(\boldsymbol{k} - \boldsymbol{k}', \boldsymbol{k}, \boldsymbol{k})e_{(s)}^{\alpha*}(\boldsymbol{x}, \boldsymbol{k})e_{(t)}^{\beta}(\boldsymbol{x}, \boldsymbol{k}')e_{(t)}^{\gamma*}(\boldsymbol{x}, \boldsymbol{k}')e_{(s)}^{\delta}(\boldsymbol{x}, \boldsymbol{k}) \\
& \quad + \mathsf{J}^{\alpha\beta\gamma\delta}(\boldsymbol{k} - \boldsymbol{k}', \boldsymbol{k}, \boldsymbol{k})\hat{e}_{(s)}^{\alpha*}(\boldsymbol{x}, \boldsymbol{k})\hat{e}_{(t)}^{\beta}(\boldsymbol{x}, \boldsymbol{k}')\hat{e}_{(t)}^{\gamma*}(\boldsymbol{x}, \boldsymbol{k}')\hat{e}_{(s)}^{\delta}(\boldsymbol{x}, \boldsymbol{k}) \quad (4.26) \\
& \quad + \mathrm{i}\omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})\mathsf{K}^{\alpha\beta\gamma\delta}(\boldsymbol{k} - \boldsymbol{k}', \boldsymbol{k}, \boldsymbol{k})\hat{e}_{(s)}^{\alpha*}(\boldsymbol{x}, \boldsymbol{k})\hat{e}_{(t)}^{\beta}(\boldsymbol{x}, \boldsymbol{k}')e_{(t)}^{\gamma*}(\boldsymbol{x}, \boldsymbol{k}')e_{(s)}^{\delta}(\boldsymbol{x}, \boldsymbol{k}) \\
& \quad + \mathrm{i}\omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})\mathsf{K}^{\gamma\delta\alpha\beta}(\boldsymbol{k}' - \boldsymbol{k}, \boldsymbol{k}', \boldsymbol{k}')e_{(s)}^{\alpha*}(\boldsymbol{x}, \boldsymbol{k})e_{(t)}^{\beta}(\boldsymbol{x}, \boldsymbol{k}')\hat{e}_{(t)}^{\gamma*}(\boldsymbol{x}, \boldsymbol{k}')\hat{e}_{(s)}^{\delta}(\boldsymbol{x}, \boldsymbol{k})\Big]
\end{aligned}
$$

and is real and non-negative. The total scattering cross section is

$$\Sigma_{(s)} = \sum_t \int \sigma_{(s,t)}(\boldsymbol{x}, \boldsymbol{k}, \boldsymbol{k}')\,\mathrm{d}\boldsymbol{k}'. \quad (4.27)$$

From the definitions (2.6)–(2.8) particularized to the case $\boldsymbol{n} = \boldsymbol{k}$, it can be shown that

$$\mathsf{H}^{\alpha\beta\gamma\delta}(\boldsymbol{k} - \boldsymbol{k}', \boldsymbol{k}, \boldsymbol{k}) = \mathsf{H}^{\gamma\delta\alpha\beta}(\boldsymbol{k}' - \boldsymbol{k}, \boldsymbol{k}', \boldsymbol{k}'),$$

with similar properties for the tensors J and H. This implies the symmetry property

$$\sigma_{(s,t)}(\boldsymbol{x}, \boldsymbol{k}, \boldsymbol{k}') = \sigma_{(t,s)}(\boldsymbol{x}, \boldsymbol{k}', \boldsymbol{k})$$

which ensures energy conservation.

The transport equations (4.25), together with the explicit expressions (4.26) and (4.27) for the differential and total scattering cross-sections, are the main results of this paper. They generalise the results of Ryzhik et al. [10] and of Guo & Wang [3] (in the conservative case) to the large class of randomly perturbed linear Hamiltonian systems of the form (1.2). With these results, the derivation of transport equations for particular systems is reduced to the

straightforward, algorithmic computation of the left and right eigenvectors $\hat{\boldsymbol{e}}_{(s)}(\boldsymbol{x}, t)$ and $\boldsymbol{e}_{(s)}(\boldsymbol{x}, t)$ and of their products with the correlation 4-tensors defined in (2.6)–(2.8). This is illustrated in the next section where we obtain transport equations for a type of waves of importance in geophysical fluid dynamics, namely Rossby waves.

5 Application to Rossby waves

Rossby waves are planetary-scale waves which propagate in the atmosphere and oceans as a result of the earth's rotation and curvature. In the last ten years or so, the satellite observation of oceanic Rossby waves has stimulated numerous works, several of which consider the interaction between Rossby waves and the bottom topography, sometimes modelled as a random function [see, e.g., 9, 5, 14, and references therein]. Scaling hypotheses of various kinds have been used to examine this interaction asymptotically, but not, as far as we are aware, the scattering scaling considered in the present paper. Here we apply the results of section 4 to obtain a transport equation for Rossby waves over a random two-dimensional topography. Since one of the aspects of physical interest is the way in which topography affects the vertical structure of the waves, we consider the simplest model in which the vertical structure is represented, namely the two-layer quasi-geostrophic model [e.g. 11, section 12].

5.1 Equations of motion

With $\boldsymbol{x} \in \mathbb{R}^2$, the natural dynamical variables for the two-layer quasi-geostrophic model are $q_1(\boldsymbol{x}, t)$ and $q_2(\boldsymbol{x}, t)$, the potential-vorticity perturbations in the top

and bottom layers, respectively. These are related to the streamfunctions of each layer according to

$$q_1 = \nabla^2 \psi_1 - F_1(\psi_1 - \psi_2), \tag{5.1}$$
$$q_2 = \nabla^2 \psi_2 - F_2(\psi_2 - \psi_1). \tag{5.2}$$

In these expressions,

$$F_1 = \frac{f^2}{g'D_1} \quad \text{and} \quad F_2 = \frac{f^2}{g'D_2},$$

where f is the mean Coriolis parameter, D_1 and D_2 are the layer depths, and $g' = g(\rho_2 - \rho_1)/\rho_2$ is the so-called reduced gravity (with g the gravity constant and ρ_1, ρ_2 the density of each layer). The linear evolution equations for q_1 and q_2 are

$$\partial_t q_1 + \beta \partial_{x_1} \psi_1 = 0, \tag{5.3}$$
$$\partial_t q_2 + \beta \partial_{x_1} \psi_2 - \frac{f}{D_2} \nabla h \times \nabla \psi_2 = 0, \tag{5.4}$$

where β is the North–South gradient of the Coriolis parameter, $h(\boldsymbol{x})$ is the topography height, and $\times$ denotes the cross product in two dimensions (i.e., $\boldsymbol{a} \times \boldsymbol{b} = a_1 b_2 - a_2 b_1$).

The Hamiltonian form of (5.3)–(5.4) is found by linearising the Poisson structure (1.5) of the corresponding nonlinear equation. This structure is given by

$$\mathcal{E} = -\frac{1}{2} \int (D_1 \psi_1 q_1 + D_2 \psi_2 q_2) \, \mathrm{d}\boldsymbol{x}, \tag{5.5}$$

and

$$\mathcal{J} = \begin{pmatrix} -D_1^{-1} \partial (q_1 + \beta x_2, \cdot) & 0 \\ & \\ 0 & -D_2^{-1} \partial (q_2 + \beta x_2 + fh/D_2, \cdot) \end{pmatrix}, \tag{5.6}$$

21

where we have introduced the Jacobian operator $\partial(a, b) = \nabla a \times \nabla b$ [cf. 12]. Since the Hamiltonian is quadratic, the linearisation is immediate, and the operators H and J follow. The form of H is derived from (5.5) by noting that

$$
H(\boldsymbol{x}, \partial_{\boldsymbol{x}}) \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} \delta\mathcal{E}/\delta q_1 \\ \delta\mathcal{E}/\delta q_2 \end{pmatrix} = - \begin{pmatrix} D_1\psi_1 \\ D_2\psi_2 \end{pmatrix}. \tag{5.7}
$$

Using (5.1)–(5.2), the corresponding matrix is written as

$$
H(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) = \frac{1}{D} \begin{pmatrix} D_1^2 K^{-2} + D_1 D_2 L^{-2} & D_1 D_2 (K^{-2} - L^{-2}) \\ D_1 D_2 (K^{-2} - L^{-2}) & D_2^2 K^{-2} + D_1 D_2 L^{-2} \end{pmatrix}, \tag{5.8}
$$

where we have defined

$$
D = D_1 + D_2, \quad \boldsymbol{k} = (k_1, k_2), \quad K^2 = k_1^2 + k_2^2, \quad \text{and} \quad L^2 = K^2 + F_1 + F_2.
$$

The matrix form of J is directly derived from (5.6) as

$$
J(\boldsymbol{x}, \mathrm{i}\boldsymbol{k}) = \begin{pmatrix} \mathrm{i}k_1\beta/D_1 & 0 \\ 0 & \mathrm{i}k_1\beta/D_2 - \mathrm{i}f\nabla h(\boldsymbol{x}) \times \boldsymbol{k}/D_2^2 \end{pmatrix}. \tag{5.9}
$$

Taking (5.7)–(5.9) into account, it is clear that the linear equations of motion (5.3)–(5.4) have the Hamiltonian form (1.2).

5.2   Scattering by topography

We consider a small-amplitude random topography $h(\boldsymbol{x})$. Rescaling $h(\boldsymbol{x})$ by $\epsilon^{1/2}$, we find that the only non-zero matrices in the expansions (2.2)–(2.3) of

22

J and H to $O(\epsilon)$ are

$$J_0(\mathrm{i}\boldsymbol{k}) = \begin{pmatrix} \mathrm{i}k_1\beta/D_1 & 0 \\ & \\ 0 & \mathrm{i}k_1\beta/D_2 \end{pmatrix}, \qquad J_{1/2}(\boldsymbol{\xi}, \mathrm{i}\boldsymbol{k}) = \begin{pmatrix} 0 & 0 \\ & \\ 0 & \mathrm{i}f\nabla h(\boldsymbol{\xi}) \times \boldsymbol{k}/D_2^2 \end{pmatrix},$$

and $H_0 = H$ given in (5.8). Taking the Fourier transform of $J_{1/2}$ with respect to $\boldsymbol{\xi}$, and denoting the power spectrum of $h(\boldsymbol{\xi})$ by $R(\boldsymbol{l})$, with

$$\langle \hat{h}(\boldsymbol{l})\hat{h}(\boldsymbol{m}) \rangle = R(\boldsymbol{l})\delta(\boldsymbol{l} + \boldsymbol{m}),$$

we compute the only non-zero entry of the correlation 4-tensor (2.6) as

$$\mathsf{J}^{2222}(\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{n}) = (\boldsymbol{l} \times \boldsymbol{k})(\boldsymbol{l} \times \boldsymbol{n})f^2 R(\boldsymbol{l})/D_2^4.$$

We now have all the elements needed to derive transport equations for Rossby waves. Solving the eigenvalue problem (4.3) with (5.8)–(5.9) gives the two frequencies

$$\omega_{(1)} = -\beta k_1/K^2 \quad \text{and} \quad \omega_{(2)} = -\beta k_1/L^2, \tag{5.10}$$

which can be recognised as the frequencies of barotropic and baroclinic Rossby waves, respectively [e.g. 11]. The corresponding right and left eigenvectors, suitably normalised, are

$$\boldsymbol{e}_{(1)} = \frac{K}{D^{1/2}}\begin{pmatrix} 1 \\ \\ 1 \end{pmatrix}, \quad \hat{\boldsymbol{e}}_{(1)} = \frac{1}{KD^{1/2}}\begin{pmatrix} D_1 \\ \\ D_2 \end{pmatrix},$$

$$\boldsymbol{e}_{(2)} = \frac{L}{(DD_1D_2)^{1/2}}\begin{pmatrix} D_2 \\ \\ -D_1 \end{pmatrix}, \quad \hat{\boldsymbol{e}}_{(2)} = \frac{(D_1D_2)^{1/2}}{LD^{1/2}}\begin{pmatrix} 1 \\ \\ -1 \end{pmatrix}.$$

23

With these results, a straightforward computation provides the scattering sections defined in (4.26) in the form

$$\sigma(\boldsymbol{k}, \boldsymbol{k}') = \frac{2\pi f^2}{D^2} (\boldsymbol{k} \times \boldsymbol{k}')^2 R(\boldsymbol{k} - \boldsymbol{k}') \tag{5.11}$$

$$\times \begin{pmatrix} (KK')^{-2}\delta(\omega_{(1)}(\boldsymbol{k}) - \omega_{(1)}(\boldsymbol{k}')) & (D_1/D_2)(KL')^{-2}\delta(\omega_{(1)}(\boldsymbol{k}) - \omega_{(2)}(\boldsymbol{k}')) \\ (D_1/D_2)(LK')^{-2}\delta(\omega_{(2)}(\boldsymbol{k}) - \omega_{(1)}(\boldsymbol{k}')) & (D_1/D_2)^2(LL')^{-2}\delta(\omega_{(2)}(\boldsymbol{k}) - \omega_{(2)}(\boldsymbol{k}')) \end{pmatrix}.$$

Here, we have written the four elements $\sigma_{(s,t)}$, $s, t = 1, 2$ in matrix form; the $(1, 1)$ entry controls the energy transfer between barotropic waves, the $(1, 2)$ entry the transfer from baroclinic waves to barotropic waves, the $(2, 1)$ entry the transfer from barotropic to baroclinic waves, and the $(2, 2)$ entry the transfer between baroclinic waves. Note that $\sigma_{(1,1)}$ also gives the scattering cross-section for Rossby waves in a single-layer quasi-geostrophic model.

5.3 Physical implications

Qualitative conclusions about the influence of topography on oceanic Rossby waves can be drawn from the form of $\sigma_{(s,t)}$. We first note from (5.10) that the frequency of the barotropic and baroclinic modes have different ranges:

$$-\infty < \omega_{(1)} < \infty \quad \text{and} \quad -\omega_{(2)}^{\max} < \omega_{(2)} < \omega_{(2)}^{\max} = \frac{\beta}{2(F_1 + F_2)^{1/2}}.$$

As a result, the barotropic modes with $|\omega_{(1)}| > \omega_{(2)}^{\max}$ do not interact with the baroclinic modes.

Next, we estimate typical orders of magnitudes for the total cross sections

$$\Sigma_{(s,t)}(\boldsymbol{k}) = \int \sigma_{(s,t)}(\boldsymbol{k}, \boldsymbol{k}') \, \mathrm{d}\boldsymbol{k}', \quad \text{for} \quad s, t = 1, 2. \tag{5.12}$$

24

These provide (inverse) time scales for the energy transfer induced by topography from barotropic/baroclinic modes to barotropic/baroclinic modes. Changing the integration variable in (5.12), we estimate the cross section $\Sigma_{(1,1)}(\boldsymbol{k})$ as follows:

$$\begin{aligned}\Sigma_{(1,1)}(\boldsymbol{k}) &= \frac{2\pi f^2}{D^2 K^2}\int \frac{(\boldsymbol{k}\times\boldsymbol{k}')^2}{|\boldsymbol{k}+\boldsymbol{k}'|^2}R(\boldsymbol{k}')\delta(\omega_{(1)}(\boldsymbol{k})-\omega_{(1)}(\boldsymbol{k}+\boldsymbol{k}'))\,\mathrm{d}\boldsymbol{k}'\\ &\sim \frac{2\pi f^2}{D^2 K^2 c_{(1)}^{\mathrm{g}}(\boldsymbol{k})}\int k_\perp^2 R(\boldsymbol{k})\,\mathrm{d}k_\perp,\end{aligned}$$

where $c_{(1)}^{\mathrm{g}} = |\boldsymbol{c}_{(1)}^{\mathrm{g}}|$ denotes the magnitude of the group velocity of the barotropic mode, and $k_\perp$ denotes the component of $\boldsymbol{k}'$ perpendicular to $\boldsymbol{c}_{(1)}^{\mathrm{g}}(\boldsymbol{k})$ (i.e. along the dispersion curve). For a topography characterised by its correlation length γ^{-1} and variance $\langle h^2\rangle$, this leads to the estimate

$$\Sigma_{(1,1)}(\boldsymbol{k}) \sim \frac{2\pi f^2\langle h^2\rangle\gamma}{D^2 K^2 c_{(1)}^{\mathrm{g}}(\boldsymbol{k})}. \tag{5.13}$$

This is a crude estimate, which neglects the strong anisotropy of Rossby waves; it nonetheless provides a first measure of effect of topographic scattering. Reasoning similarly, we estimate the other total cross sections as

$$\Sigma_{(1,2)}(\boldsymbol{k}) \sim \frac{2\pi D_1 f^2\langle h^2\rangle\gamma}{D_2 D^2(F_1+F_2)c_{(2)}^{\mathrm{g}}(\boldsymbol{k})}, \quad \Sigma_{(2,1)}(\boldsymbol{k}) \sim \frac{2\pi D_1 f^2\langle h^2\rangle\gamma}{D_2 D^2(F_1+F_2)c_{(1)}^{\mathrm{g}}(\boldsymbol{k})},$$
$$\text{and} \quad \Sigma_{(2,2)}(\boldsymbol{k}) \sim \frac{2\pi D_1^2 K^2 f^2\langle h^2\rangle\gamma}{D_2^2 D^2(F_1+F_2)^2 c_{(2)}^{\mathrm{g}}(\boldsymbol{k})}, \tag{5.14}$$

where we have approximated $L \sim F_1 + F_2$ as is relevant except for very short waves. When the interface between the two fluids models the thermocline, as is usual, $D_1 \ll D_2$. Thus, most of the scattering occurs between barotropic modes, and the baroclinic modes are affected by topography mainly through their interactions with barotropic modes. This is of course consistent with the fact that the baroclinic modes have a weak signature in the lower layer.

We now compute the order of magnitude of the cross sections estimated in (5.13)–(5.14). We choose typical North-Atlantic values for the parameters which define the two-layer quasi-geostrophic model:

$$D_1 = 500\,\text{m},\ D_2 = 3000\,\text{m},\ f = 10^{-4}\text{s}^{-1},$$
$$\beta = 2 \times 10^{-11}\,\text{m}^{-1}\text{s}^{-1} \quad \text{and} \quad g' = 0.02\,\text{m}\,\text{s}^{-2}.$$

This choice yields $F_1 = 10^{-9}\,\text{m}^{-2}$, $F_2 \sim 2 \times 10^{-10}\,\text{m}^{-2}$, and hence an internal radius of deformation $(F_1 + F_2)^{-1/2} \sim 30\,\text{km}$ [cf. 11, section 13]. Because it assumes that the topography correlation length has the same order of magnitude as the wavelength, the transport theory of this paper best applies to moderately long Rossby waves. To fix ideas, we consider waves with $K = 10^{-5}\,\text{m}^{-1}$, i.e. wavelength of 600 km or so; this is consistent with the reasonable estimate for the topography correlation length given by $\gamma^{-1} = 100\,\text{km}$ [13]. The scaling adopted is relevant only to regions where the topography is shallow enough to affect the waves less than the β-effect. Specifically, the condition

$$\frac{f(\langle h^2 \rangle)^{1/2}\gamma}{D_2} \ll \beta \tag{5.15}$$

needs to be satisfied (cf. (5.4)). This is realistic only in regions with rather low topographic features, for which we take $(\langle h^2 \rangle)^{1/2} = 20\,\text{m}$, thus satisfying (5.15) only marginally. The group velocity of Rossby waves depends strongly on the angle between their wavevector $\boldsymbol{k}$ and the North–South direction. Here, since we limit our discussion to order-of-magnitude considerations, we take the crude approximation

$$c_{(1)}^{\text{g}} \sim \beta/K^2 \sim 0.2\,\text{m}\,\text{s}^{-1} \quad \text{and} \quad c_{(2)}^{\text{g}} \sim \beta/(F_1 + F_2) \sim 0.02\text{m}\,\text{s}^{-1},$$

for the barotropic and baroclinic mode, respectively.

26

With all these numerical values, we obtain

$$\Sigma_{(1,1)} \sim 10^{-6}\,\mathrm{s}^{-1}, \quad \Sigma_{(1,2)} \sim 2 \times 10^{-7}\,\mathrm{s}^{-1}, \quad \Sigma_{(2,1)} \sim 2 \times 10^{-8}\,\mathrm{s}^{-1}, \quad \Sigma_{(2,2)} \sim 3 \times 10^{-9}\,\mathrm{s}^{-1}.$$

This indicates a large difference in the scattering time scales between the barotropic and baroclinic modes, with the former expected to relax much more rapidly toward than the latter to an equilibrium distribution (see below). Also, the scattering between baroclinic modes is essentially negligible. This is due to two factors: the small depth ratio $D_1/D_2 \ll 1$, and the large group velocity ratio $c^{\mathrm{g}}_{(1)}/c^{\mathrm{g}}_{(2)} \gg 1$.

The corresponding mean free path Σ/c, which gives the typical propagation distance over which the effect of scattering is relevant, follows: it is of the order of 200 km for the barotropic mode, and 1000 km for the baroclinic mode. The fact that the mean free path for the barotropic mode is only marginally larger than the wavelength suggests that the transport theory has only a limited applicability for the barotropic mode, even for the small topographic height assumed here. In contrast, there is a clear separation between wavelength and mean free path for the baroclinic mode.

It is interesting to examine the energy densities $a_{(1)}(\boldsymbol{x}, \boldsymbol{k}, t)$ and $a_{(2)}(\boldsymbol{x}, \boldsymbol{k}, t)$ that can be expected at equilibrium. Assuming space independence, it is easily seen from (4.25)–(4.27) that a time-dependent solution is obtained for

$$a_{(1)}(\boldsymbol{k}) = \phi(\omega_{(1)}(\boldsymbol{k})) \quad \text{and} \quad a_{(2)}(\boldsymbol{k}) = \phi(\omega_{(2)}(\boldsymbol{k})) \tag{5.16}$$

for some function of a single variable $\phi(\cdot)$. Thus, viewed in the $\boldsymbol{k}$-plane, the energy densities for both the barotropic and baroclinic modes are constant (and equal) along constant-frequency curves. In an initial-value situation, this equilibrium distribution is attained after a transient adjustment stage on a

time scale fixed by the total cross sections estimated above. The ratio of baroclinic to barotropic energy at equilibrium is given by

$$
r = \frac{\displaystyle\int \phi(\omega_{(2)}(\boldsymbol{k}))\,\mathrm{d}\boldsymbol{k}}{\displaystyle\int \phi(\omega_{(1)}(\boldsymbol{k}))\,\mathrm{d}\boldsymbol{k}} = \frac{\displaystyle\int \phi(\Omega)\,\mathrm{d}\Omega \int\limits_{\omega_{(2)}=\Omega} \frac{\mathrm{d}l}{c^{\mathrm{g}}_{(2)}}}{\displaystyle\int \phi(\Omega)\,\mathrm{d}\Omega \int\limits_{\omega_{(1)}=\Omega} \frac{\mathrm{d}l}{c^{\mathrm{g}}_{(1)}}},
$$

where l denotes the arclength along constant frequency curve. Remarkably, in spite of the dispersive and anisotropic nature of the Rossby waves, this ratio is simply unity in the baroclinic frequency range. This is because

$$
\int\limits_{\omega_{(1)}=\Omega} \frac{\mathrm{d}l}{c^{\mathrm{g}}_{(1)}} = \int\limits_{\omega_{(2)}=\Omega} \frac{\mathrm{d}l}{c^{\mathrm{g}}_{(2)}} \quad \left(= \frac{\pi\beta^2}{\Omega^3}\right),
$$

as a simple computation shows. Thus, according to transport theory, the scattering by topography leads to energy equipartition between barotropic and baroclinic modes.

To conclude this discussion of Rossby-wave scattering, we note that the equilibrium distribution (5.16) obtained here under the assumption of space independence remains relevant when there are spatial variations over a scale much larger than the mean free path. In this case, the function ϕ depends (slowly) on $\boldsymbol{x}$ and t in addition to the frequency, and it obeys a diffusion equation [cf. 10]. The derivation of the corresponding anisotropic, $\boldsymbol{k}$-dependent diffusion tensor would be of interest.

6   Discussion

In this paper, we have extended the derivation of transport equations for waves in randomly perturbed conservative media due to Ryzhik et al. [10] to

the general class of linear Hamiltonian systems (1.2). The Hamiltonian nature of the systems considered results in scattering terms that are clearly conservative, unlike those of Guo & Wang [3] who treated general, non-Hamiltonian Schrödinger equations.

Our results are limited to systems with non-degenerate dispersion relations. In the presence of degeneracies, that is, in the presence of polarised waves, the expansion (4.9) of the Wigner function is not complete, and additional amplitudes capturing cross-polarisation effects must be introduced [see 10]. There should be no difficulty in extending our results in this manner to derive transport equations for polarised waves.

We conclude by noting that, to date, the transport equations derived in general contexts are valid for time-independent media. It would be of interest to consider time-dependent media, with a deterministic dependence that is slow compared to the wave periods, and (possibly) a random time dependence with a time scale that is comparable to the wave periods. In the absence of randomness, transport equations of the Liouville form (4.17) can be expected but for the density of wave action rather than wave energy. It is worth pointing out that the transport equations will not be exact conservations as in the time-independent case, but rather approximate conservations, expressing the adiabatic invariance of the wave-action density. This adiabatic invariance cannot be expected to hold for arbitrary time dependence of the operator J, however: as work about the spatial density of wave action (as opposed to the phase-space density considered here) indicates [15] the adiabatic invariance of wave action is valid only for special time dependences of J such as arise when linearising Hamiltonian systems near exact time-dependent solutions. Including time-dependent random inhomogeneities will lead to scattering; unlike that examined here, this scattering will involve waves with different frequencies,

with scattering cross-sections that are controlled by the frequency spectrum of the inhomogeneities.

A Derivation details

A.1 Wave energy and Wigner matrix

Introducing the Fourier expansion of $\boldsymbol{u}(\boldsymbol{x},t)$ into (1.3) leads to

$$\begin{aligned}\mathcal{H} &= \frac{1}{2}\iiint \hat{\boldsymbol{u}}^*(\boldsymbol{l},t)H(\boldsymbol{x},\mathrm{i}\epsilon\boldsymbol{k})\hat{\boldsymbol{u}}(\boldsymbol{k},t)\mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{l})\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{k}\mathrm{d}\boldsymbol{l}\\ &= \frac{1}{2}\operatorname{Tr}\iiint H(\boldsymbol{x},\mathrm{i}\epsilon\boldsymbol{k})\hat{\boldsymbol{u}}(\boldsymbol{k},t)\hat{\boldsymbol{u}}^*(\boldsymbol{l},t)\mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{l})\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{k}\mathrm{d}\boldsymbol{l}.\end{aligned}$$

Changing the variables of integration and using (3.2) then gives

$$\begin{aligned}\mathcal{H} &= \frac{1}{2(2\pi)^d}\operatorname{Tr}\iiiint H(\boldsymbol{x},\mathrm{i}\boldsymbol{k}+\mathrm{i}\epsilon\boldsymbol{l}/2)W_\epsilon(\boldsymbol{y},\boldsymbol{k},t)\mathrm{e}^{\mathrm{i}\boldsymbol{l}\cdot(\boldsymbol{x}-\boldsymbol{y})}\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{k}\mathrm{d}\boldsymbol{l}\\ &= \frac{1}{2}\operatorname{Tr}\iint H(\boldsymbol{x},\mathrm{i}\boldsymbol{k}+\epsilon\partial_{\boldsymbol{x}}/2)W_\epsilon(\boldsymbol{x},\boldsymbol{k},t)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{k},\end{aligned}$$

A.2 Calculation of $\mathcal{P}_{(s,t)}$

We introduce (4.13) into (4.12), and left- and right-multiply by $\hat{\boldsymbol{e}}_{(s)}^*(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})$ and $\hat{\boldsymbol{e}}_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})$, respectively. Using (4.3), (4.7) and (4.9), we obtain

$$\begin{aligned}&\mathrm{i}\left[\omega_{(s)}(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})-\omega_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})\right]\mathcal{P}_{(s,t)}(\boldsymbol{x},\boldsymbol{l},\boldsymbol{k},t)\\ &= a_{(s)}(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l},t)\boldsymbol{e}_{(s)}^*(\boldsymbol{x},\boldsymbol{k}+\boldsymbol{l})V^*(\boldsymbol{x},-2\boldsymbol{l},\boldsymbol{k}+\boldsymbol{l})\hat{\boldsymbol{e}}_{(t)}(\boldsymbol{x},\boldsymbol{k}-\boldsymbol{l})\end{aligned}$$

$$+ a_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}, t) \hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k} + \boldsymbol{l}) V(\boldsymbol{x}, 2\boldsymbol{l}, \boldsymbol{k} - \boldsymbol{l}) \boldsymbol{e}_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}).$$

Using the definition (3.10) of V, (4.5) and (4.6), this can be rewritten as

$$
\begin{aligned}
& \mathrm{i} \left[\omega_{(s)}(\boldsymbol{x}, \boldsymbol{k} + \boldsymbol{l}) - \omega_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}) \right] \mathcal{P}_{(s,t)}(\boldsymbol{x}, \boldsymbol{l}, \boldsymbol{k}, t) \\
& = a_{(s)}(\boldsymbol{x}, \boldsymbol{k} + \boldsymbol{l}, t) \left[\hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k} + \boldsymbol{l}) \hat{J}^*_{1/2}(-2\boldsymbol{l}, \boldsymbol{k} + \boldsymbol{l}) \hat{e}_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}) \right. \\
& \qquad\qquad \left. + \mathrm{i}\omega_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}) e^*_{(s)}(\boldsymbol{x}, \boldsymbol{k} + \boldsymbol{l}) \hat{H}^*_{1/2}(-2\boldsymbol{l}, \boldsymbol{k} + \boldsymbol{l}) \boldsymbol{e}_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}) \right] \\
& \quad + a_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}, t) \left[\hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k} + \boldsymbol{l}) \hat{J}_{1/2}(2\boldsymbol{l}, \boldsymbol{k} - \boldsymbol{l}) \hat{e}_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}) \right. \\
& \qquad\qquad \left. - \mathrm{i}\omega_{(s)}(\boldsymbol{x}, \boldsymbol{k} + \boldsymbol{l}) e^*_{(s)}(\boldsymbol{x}, \boldsymbol{k} + \boldsymbol{l}) \hat{H}_{1/2}(2\boldsymbol{l}, \boldsymbol{k} - \boldsymbol{l}) \boldsymbol{e}_{(t)}(\boldsymbol{x}, \boldsymbol{k} - \boldsymbol{l}) \right].
\end{aligned}
$$

This expression reduces to (4.14) on using (2.9).

A.3 Deterministic contributions to the transport equations

The deterministic terms in (4.16), namely $\hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k})(Q_{01} + Q_{10}) W_0(\boldsymbol{x}, \boldsymbol{k}, t) \boldsymbol{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k})$, can in principle be simplified by introducing the form of Q_{10} and Q_{01}, and the expansion (4.9) of W_0, and by making extensive use of the eigenvalue problem (4.3) and of its consequences. The relationships (2.4)–(2.5) between J_1, H_1 and J_0, H_0 are also crucial. Here, we circumvent most of the tedious computations that this entails by exploiting the conservation of energy.

We first note that, in the absence of random terms, (4.16) takes the form

$$
\begin{aligned}
\partial_t a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) + \mathbf{f}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \cdot \nabla_{\boldsymbol{x}} a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) - \mathbf{g}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \cdot \nabla_{\boldsymbol{k}} a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) \\
= \sum_t h_{(s,t)}(\boldsymbol{x}, \boldsymbol{k}) a_{(t)}(\boldsymbol{x}, \boldsymbol{k}, t), \qquad\qquad (\mathrm{A.1})
\end{aligned}
$$

when (4.7) is used. Here, we have defined

$$
\begin{aligned}
\mathbf{f}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) &= \frac{\mathrm{i}}{2} \left[\hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \nabla_{\boldsymbol{k}} L_0(\boldsymbol{x}, \boldsymbol{k}) \boldsymbol{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) - e^*_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \nabla_{\boldsymbol{k}} L^*_0(\boldsymbol{x}, \boldsymbol{k}) \hat{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \right] \\
\mathbf{g}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) &= \frac{\mathrm{i}}{2} \left[\hat{e}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \nabla_{\boldsymbol{x}} L_0(\boldsymbol{x}, \boldsymbol{k}) \boldsymbol{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) - e^*_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \nabla_{\boldsymbol{x}} L^*_0(\boldsymbol{x}, \boldsymbol{k}) \hat{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \right],
\end{aligned}
$$

and we have grouped all the terms involving undifferentiated amplitudes $a_{(t)}(\boldsymbol{x}, \boldsymbol{k}, t)$ on the right-hand side. The scalars $h_{(s,t)}$ can be written in terms of L_0, $\boldsymbol{e}_{(s)}$, $\hat{\boldsymbol{e}}_{(s)}$, $\boldsymbol{e}_{(t)}$, $\hat{\boldsymbol{e}}_{(t)}$ and their derivatives with respect to $\boldsymbol{x}$ and $\boldsymbol{k}$ (as well as H_1 and J_1). Now, differentiating (4.3) with respect to $\boldsymbol{k}$ and $\boldsymbol{x}$, respectively, leads, after left-multiplying by $\hat{\boldsymbol{e}}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k})$ and using (4.7), to

$$\hat{\boldsymbol{e}}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \nabla_{\boldsymbol{k}} L_0(\boldsymbol{x}, \boldsymbol{k}) \boldsymbol{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) = -\mathrm{i} \nabla_{\boldsymbol{k}} \omega_{(s)}(\boldsymbol{x}, \boldsymbol{k})$$
$$\text{and} \quad \hat{\boldsymbol{e}}^*_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \nabla_{\boldsymbol{x}} L_0(\boldsymbol{x}, \boldsymbol{k}) \boldsymbol{e}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) = -\mathrm{i} \nabla_{\boldsymbol{x}} \omega_{(s)}(\boldsymbol{x}, \boldsymbol{k}).$$

Therefore,

$$\mathbf{f}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) = \nabla_{\boldsymbol{k}} \omega_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \quad \text{and} \quad \mathbf{g}_{(s)}(\boldsymbol{x}, \boldsymbol{k}) = \nabla_{\boldsymbol{x}} \omega_{(s)}(\boldsymbol{x}, \boldsymbol{k}).$$

We now argue that the scalars $h_{(s,t)}(\boldsymbol{x}, \boldsymbol{k}) \equiv 0$ for all s and t. For $s \neq t$, we have

$$h_{(s,t)} = -\frac{\mathrm{i}}{2} \hat{\boldsymbol{e}}^*_{(s)} \nabla_{\boldsymbol{k}} L_0 \, \boldsymbol{e}_{(t)} \cdot \nabla_{\boldsymbol{x}} \boldsymbol{e}^*_{(t)} \, \hat{\boldsymbol{e}}_{(s)} + \frac{\mathrm{i}}{2} \hat{\boldsymbol{e}}^*_{(s)} \nabla_{\boldsymbol{x}} L_0 \, \boldsymbol{e}_{(t)} \cdot \nabla_{\boldsymbol{k}} \boldsymbol{e}^*_{(t)} \, \hat{\boldsymbol{e}}_{(s)} + \text{c.c.},$$

where we have omitted the dependence on $(\boldsymbol{x}, \boldsymbol{k})$ for simplicity. Differentiating (4.3) with respect to $\boldsymbol{k}$ or $\boldsymbol{x}$ yields

$$\hat{\boldsymbol{e}}^*_{(s)} \nabla_{\boldsymbol{k}} L_0 \boldsymbol{e}_{(t)} = \mathrm{i}(\omega_{(s)} - \omega_{(t)}) \hat{\boldsymbol{e}}^*_{(s)} \nabla_{\boldsymbol{k}} \boldsymbol{e}_{(t)} \quad \text{and} \quad \hat{\boldsymbol{e}}^*_{(s)} \nabla_{\boldsymbol{x}} L_0 \boldsymbol{e}_{(t)} = \mathrm{i}(\omega_{(s)} - \omega_{(t)}) \hat{\boldsymbol{e}}^*_{(s)} \nabla_{\boldsymbol{x}} \boldsymbol{e}_{(t)}$$

for $s \neq t$. Taking this into account, we find that $h_{(s,t)} = 0$ for $s \neq t$ and thus that (A.1) reduces to

$$\partial_t a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) + \nabla_{\boldsymbol{k}} \omega_{(s)}(\boldsymbol{x}, \boldsymbol{k}) \cdot \nabla_{\boldsymbol{x}} a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) - \nabla_{\boldsymbol{x}} \omega_{(s)} \cdot \nabla_{\boldsymbol{k}} a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t)$$
$$= h_{(s,s)}(\boldsymbol{x}, \boldsymbol{k}) a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t).$$

It is now easy to show that $h_{(s,s)}(\boldsymbol{x}, \boldsymbol{k}) \equiv 0$: in view of (4.10), the conservation of the energy $\mathcal{H}$ implies at leading order in ϵ that

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_s \iint a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) \, \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{k} = \sum_s \iint h_{(s,s)}(\boldsymbol{x}, \boldsymbol{k}) a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t) \, \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{k} = 0$$

Since this holds for arbitrary $a_{(s)}(\boldsymbol{x}, \boldsymbol{k}, t)$, $h_{(s,s)}(\boldsymbol{x}, \boldsymbol{k}) \equiv 0$ and the Liouville form (4.17) of the transport equations follows.

A.4 Derivation of I_{J}, I_{H} and I_{K}

In this Appendix, we omit the dependence of the various objects on $\boldsymbol{x}$ and t. We introduce (4.13)–(4.14) into (4.18) and use (2.6)–(2.8) to obtain explicit expressions for the three terms I_{J}, I_{H} and I_{K}. A simple computation using (4.5) and (4.7) gives

$$I_{\mathsf{J}} = -2^d \mathrm{i} \sum_t \int \frac{a_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - a_{(s)}(\boldsymbol{k})}{\omega_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - \omega_{(s)}(\boldsymbol{k})} \Xi^{\mathsf{J}}_{(s,t)}(2\boldsymbol{l}, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{l} + \text{c.c.},$$

with $\Xi^{\mathsf{J}}_{(s,t)}$ defined in (4.20). It is necessary to regularise the denominator appearing in this expression [cf. 10]. Causality indicates that this is achieved by adding $-\mathrm{i}\theta$ to the denominator and taking the limit $\theta \downarrow 0$ or, in other words, by making the substitution

$$\frac{1}{\omega_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - \omega_{(s)}(\boldsymbol{k})} \mapsto PV \frac{1}{\omega_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - \omega_{(s)}(\boldsymbol{k})} + \mathrm{i}\pi\delta[\omega_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - \omega_{(s)}(\boldsymbol{k})], \quad \text{(A.2)}$$

where PV denotes the Cauchy principal value. Using the property (2.10) of J, it is easy to show that $\Xi^{\mathsf{J}}_{s,t}(2\boldsymbol{l}, \boldsymbol{k})$ is real. Since the energy density and frequency are also real, the only contribution to I_{J} comes from the regularisation of the denominator, leading to

$$I_{\mathsf{J}} = 2^d \pi \sum_t \int \Xi^{\mathsf{J}}_{(s,t)}(2\boldsymbol{l}, \boldsymbol{k})\delta[\omega_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - \omega_{(s)}(\boldsymbol{k})][a_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - a_{(s)}(\boldsymbol{k})] \, \mathrm{d}\boldsymbol{l} + \text{c.c.},$$

Changing the integration variable yields (4.19). The treatment of I_{H} is similar: using (4.6) and (4.7), it can be written in the form

$$I_{\mathsf{H}} = -2^d \mathrm{i}\omega_{(s)}(\boldsymbol{k}) \sum_t \int \frac{\omega_{(s)}(\boldsymbol{k})a_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - \omega_{(t)}(\boldsymbol{k} - 2\boldsymbol{l})a_{(s)}(\boldsymbol{k})}{\omega_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - \omega_{(s)}(\boldsymbol{k})} \Xi^{\mathsf{H}}_{(s,t)}(2\boldsymbol{l}, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{l} + \text{c.c.},$$

with $\Xi^{\mathsf{H}}_{(s,t)}$ defined in (4.22). Using (2.10) shows that $\Xi^{\mathsf{H}}_{(s,t)}$ is real, so that only the regularisation of the denominator contributes to I^{H}. Using (A.2) and changing the variable of integration leads to (4.21).

The third term, I^{K} is computed in a similar, though lengthier, manner. Using (4.5), (4.6) and (4.7), it can be written as

$$I_{\mathsf{K}} = -2^d \left[i\omega_{(s)}(\boldsymbol{k}) \sum_t \int \frac{a_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - a_{(s)}(\boldsymbol{k})}{\omega_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) - \omega_{(s)}(\boldsymbol{k})} \Xi^{\mathsf{H}}_{(s,t)}(2\boldsymbol{l}, \boldsymbol{k}) \, \mathrm{d}\boldsymbol{l} \right.$$
$$\left. + a_{(s)}(\boldsymbol{k}) \sum_t \int \mathsf{K}^{\alpha\beta\gamma\delta}(2\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{k}) \hat{e}^{\alpha*}_{(s)}(\boldsymbol{k}) \hat{e}^{\beta}_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) e^{\gamma*}_{(t)}(\boldsymbol{k} - 2\boldsymbol{l}) e^{\delta}_{(s)}(\boldsymbol{k}) \, \mathrm{d}\boldsymbol{l} \right] + \mathrm{c.c.},$$

where $\Xi^{\mathsf{K}}_{(s,t)}$ defined in (4.24). Taking (4.8) into account, the term on the second line can be simplified into

$$-2^d a_{(s)}(\boldsymbol{k}) \int \mathsf{K}^{\alpha\beta\beta\delta}(2\boldsymbol{l}, \boldsymbol{k}, \boldsymbol{k}) \hat{e}^{\alpha*}_{(s)}(\boldsymbol{k}) e^{\delta}_{(s)}(\boldsymbol{k}) \, \mathrm{d}\boldsymbol{l} + \mathrm{c.c.}.$$

Using (2.10), it can be shown that $\Xi^{\mathsf{K}}_{(s,t)}$ is real. Thus, when substituting (A.2) into the first line of I^{K}, only the regularisation remains, leading to (4.23) after changing the variable of integration.

References

[1] Asch, M., Kholer, W., Papanicolaou, G., Postel, M. & Sheng, P. 1991, Frequency content of randomly scattered signals, *SIAM Review* **33**, 519–625.

[2] Folland, G. B. 1989, *Harmonic analysis in phase space*, Princeton University Press.

[3] Guo, M. & Wang, X.-P. 1999, Transport equations for a general class of evolution equations with random perturbations, *J. Math. Phys.* **40**, 4828–4858.

[4] Hinch, E. J. 1991, *Perturbation methods*, Cambridge University Press.

[5] Klyatskin, V. I., Gryanik, N. V. & Gurarie, D. 1998, Propagation and localization of Rossby waves over random topography (two-layer model), *Wave Motion* **28**, 333–352.

[6] Molchanov, S. A. 1991, Ideas in the theory of random media, *Acta Appl. Math.* **22**, 139–282.

[7] Morrison, P. J. 1998, Hamiltonian description of the ideal fluid, *Rev. Mod. Phys.* **70**, 467–521.

[8] Olver, P. J. 1993, *Application of Lie groups to differential equations*, 2nd edn, Springer–Verlag.

[9] Reznik, G. M. & Tsybaneva, T. B. 1999, Planetary waves in a stratified ocean of variable depth. Part 1. Two-layer model, *J. Fluid Mech.* **388**, 115–145.

[10] Ryzhik, L. R., Papanicolaou, G. & Keller, J. B. 1996, Transport equations for elastic and other waves in random media, *Wave Motion* **24**, 327–370.

[11] Salmon, R. 1998, *Lectures on geophysical fluid dynamics*, Oxford University.

[12] Shepherd, T. G. 1990, Symmetries, conservation laws and Hamiltonian structure in geophysical fluid dynamics, *Adv. Geophys.* **32**, 287–338.

[13] Thomson, R. E. 1975, The propagation of planetary waves over random topography, *J. Fluid Mech.* **70**, 267–285.

[14] Vanneste, J. 2003, Nonlinear dynamics over rough topography: barotropic and stratified quasi-geostrophic theory, *J. Fluid Mech.* **474**, 299–318.

[15] Vanneste, J. & Shepherd, T. G. 1999, On wave action and phase in the non-canonical Hamiltonian formulation, *Proc. R. Soc. Lond. A* **455**, 3–21.