

Generalized transport characterizations for short oceanic internal waves in a sea of long waves

Yuri V. Lvov^{1,†} and Kurt L. Polzin²

¹Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

²Woods Hole Oceanographic Institution, MS#21, Woods Hole, MA 02543, USA

(Received 14 August 2023; revised 23 February 2024; accepted 2 March 2024)

Wave turbulence provides a conceptual framework for weakly nonlinear interactions in dispersive media. Dating from five decades ago, applications of wave turbulence theory to oceanic internal waves assigned a leading-order role to interactions characterized by a near equivalence between the group velocity of high-frequency internal waves with the phase velocity of near-inertial waves. This scale-separated interaction leads to a Fokker–Planck (generalized diffusion) equation. More recently, starting four decades ago, this scale-separated paradigm has been investigated using ray tracing methods. These ray methods characterize spectral transport of energy by counting the amplitude and net velocity of wave packets in phase space past a high-wavenumber gate prior to ‘breaking’. This explicitly advective characterization is based on an intuitive assignment and lacks theoretical underpinning. When one takes an estimate of the net spectral drift from the wave turbulence derivation and makes the corresponding assessment, one obtains a prediction of spectral transport that is an order of magnitude larger than either observations or reported ray tracing estimates. Motivated by this contradiction, we report two parallel derivations for transport equations describing the refraction of high-frequency internal waves in a sea of random inertial waves. The first uses standard wave turbulence techniques and the second is an ensemble-averaged packet transport equation characterized by the dispersion of wave packets about a mean drift in the spectral domain. The ensemble-averaged transport equation for ray tracing differs in that it contains the intuitively motivated advective term. We conclude that the aforementioned contradiction between theory, numerics and observations needs to be taken at face value and present a pathway for resolving this contradiction.

Key words: internal waves, ocean processes, stratified flows

1. Introduction

Internal waves are a fascinating phenomenon, ubiquitous in the ocean and characterized by the oscillation of the invisible surfaces of constant densities of a stratified water column.

† Email address for correspondence: lvov@rpi.edu

Internal waves carry a significant fraction of ocean kinetic energy and are an important intermediary in transferring energy and momentum to smaller scales where they are dissipated. There is renewed theoretical interest in investigating the internal waves in the ocean due to recent developments in our ability to perform high-resolution numerical modelling, with internal waves permitting global ocean simulations (Arbic *et al.* 2018) and regional numerical models (Pan *et al.* 2020).

The wave turbulence kinetic equation has been used extensively to describe processes of spectral energy transfers between internal waves, see Müller *et al.* (1986) and Polzin & Lvov (2011) for reviews. Early on, special types of resonant three-wave triads characterized by extreme scale separations were identified to play an important role in these spectral energy transfers (McComas & Bretherton 1977; McComas & Müller 1981a), yet the details and delicate counterbalance of the nonlinear transfers remain an enigma. An important feature of the internal wave kinetic equation is that it diverges for almost all spectral power-law indices in the internal wave spectrum (Lvov *et al.* 2010). This is a mathematical manifestation of the lack of locality in internal wave interactions: nonlinear transfers have significant contributions under extreme scale-separated conditions. The weighting of these extreme scale-separated interactions with spectral power-law assumptions results in divergent integrals. Our goal is to understand and characterize this nonlinear transfer process in this extreme scale-separated limit.

An alternative approach to the wave turbulence kinetic equation is proposed in Henyey, Wright & Flatté (1986), where ray-tracing (or eikonal) techniques are used to describe spectral energy transfers. In this paradigm, the energy cascade is assessed as a net drift $\langle \dot{\mathbf{p}} \rangle$ of wave packets toward high wavenumber, where $\mathbf{p}$ is the momentum of a wave and $\langle \dots \rangle$ represents an ensemble average. Such studies (Henyey *et al.* 1986; Sun & Kunze 1999b; Ijichi & Hibiya 2017) provide metrics of the net drift rate at a high-wavenumber gate, beyond which waves are considered to ‘break’. These numerical simulations are conducted in a ‘kitchen sink’ manner in which scale separations in vertical and horizontal wavenumber are viewed as tuneable parameters to arrive at downscale transport estimates that align with observations. The alignment requires that the background have similar scales as the wave packet and creates a thematic issue for an asymptotic theory such as ray tracing. A further issue is that a rigorous description of this ensemble average transport $\langle \dot{\mathbf{p}} \rangle$ is an open question that we address here.

Motivation for our efforts comes from a comparison with the empirical metrics of ocean mixing referred to as ‘fine-scale parameterizations’ (Gregg 1989; Polzin, Toole & Schmitt 1995); see Polzin *et al.* (2014) for a review and McComas (1977), McComas & Müller (1981b), Polzin & Lvov (2017) and Dematteis, Polzin & Lvov (2022) for descriptions of the pivotal role that extreme scale-separated interactions play in interpreting the oceanic internal wave spectrum. The most glaring incompatibility of wave turbulence and ray tracing is presented in § 4: if the mean-drift rate in vertical wavenumber is identified as the corresponding gradient of diffusivity as derived from the kinetic equation, then the predicted downscale energy transport is an order of magnitude larger than that supported by the observations. This method parallels assessments for downscale energy transport in ray-tracing numerics (Henyey *et al.* 1986; Sun & Kunze 1999b; Ijichi & Hibiya 2017), but is similarly ten times larger than those numerical results. This disparity has led us to a systematic examination and physical interpretation of the assumptions within both kinetic equation and ray-path approaches to pinpoint the multiple junctures that might underpin a systematic difference between observation and theory concerning extreme scale-separated interactions.

Despite claims by McComas & Bretherton (1977) and Nazarenko, Newell & Galtier (2001) that ray tracing should reduce to the resonant manifold, our understanding is that the kinetic equation and ray tracing differ on fundamental levels. The wave kinetic equation represents the internal wave field as a system of amplitude-modulated waves having constant wavenumber and frequency linked through a dispersion relation. Ray tracing represents a wave packet as a frequency-modulated system with variations in wavenumber linked to the conservation of an Eulerian phase function. The Fokker–Plank equation derived in the ray-tracing approach additionally represents the average drift of wave packets towards high wavenumbers. The role of resonances and off-resonant interactions in the mean drift and dispersion about that mean drift are also different (Polzin & Lvov 2023), as are the concepts of resonance broadening (Polzin & Lvov 2017) and bandwidth (Cohen & Lee 1990) that are important metrics of finite-amplitude effects in weakly nonlinear systems.

Our efforts have direct parallels with Kraichnan's 1959 and 1965 studies (Kraichnan 1959, 1965) of isotropic homogeneous turbulence using field theoretic techniques. Kraichnan's 1959 study was an Eulerian-based approach that yielded a $k^{-3/2}$ spectrum at high Reynolds number, distinct from Kolmogorov's $k^{-5/3}$ inertial subrange based upon dimensional analysis. (Here the k represents the magnitude of the wavenumber.) This Eulerian description was labelled the 'direct interaction approximation' (DIA) and extant data were not sufficient to assess the theoretical prediction. Kraichnan (1965) subsequently understood that the quasi-uniform translation associated with coherent advection at the largest scales (aka sweeping). This sweeping effect was creating an artefact wherein the correlation time scale was proportional to the root-mean-square Doppler shift rather than a more intuitive notion that energy transfers between scales depended upon the rate of strain. In 1965, Kraichnan subsequently presented a Lagrangian description (the abridged Lagrangian history DIA) that isolated the pressure and viscous terms responsible for fluid parcel deformation. The Lagrangian picture resulted in a $-5/3$ power law and Kolmogorov constant (1.77) quite close to that provided by a summary of atmospheric field data (1.56) (Högström 1996). The analogy to Kraichnan is that the plane wave formulation corresponds to an Eulerian coordinate system and the Lagrangian coordinate system corresponds to a wave packet formulation in which statistics are accumulated along ray paths.

Similar issues about Doppler shifting arise for internal waves (Holloway 1980, 1982), Rossby waves (Holloway & Hendershott 1977; Nazarenko 2011) and in magneto-hydrodynamics (Nazarenko *et al.* 2001). Wave problems are potentially more complicated, in part because the Doppler shifting can be intrinsically related to extreme scale-separated interactions, and due to a multiplicity of time scales introduced through resonant interactions absent in three-dimensional (3-D) turbulence. In wave turbulence one assumes an expansion in terms of small nonlinearity and an assumption about multiple time scales to assess the evolution of amplitude-modulated plane waves. Implicit is a long interaction time scale for third order and a short decorrelation time scale with regards to the higher orders (Newell 1968). Reduction of the DIA to the resonant manifold happens as the correlation time scale is small relative to an interaction time scale, and triple correlations associated with nonlinear coupling can be related to the product of two double correlations. Extreme scale-separated wave problems can also be treated with ray-tracing methods, in which the statistics of frequency-modulated wave packets are accumulated along ray characteristics rather than Lagrangian trajectories. Ray tracing is an extremely attractive route to deal with sweeping as the dynamics of ray tracing is grounded in the explicit representation of variations in Doppler shifting. It is understood that there are ray method

parallels to the interaction and correlation time scales of wave turbulence and the DIA, (McComas & Bretherton 1977; Nazarenko *et al.* 2001). However, the time scale definitions for ray methods have not been sufficiently developed for a detailed comparison of the two strategies for assessing the effects of Doppler shifting. In particular, what has been missing is the identification of the interaction time scale. Here, we provide a derivation of a generalized transport equation for the evolution of an ensemble of wave packets. This generalized transport equation contains a term representing the ensemble mean drift of wave packets in the spectral domain. This mean drift relates to the interaction time scale and can be directly compared with a correlation time scale relating to dispersion about that mean drift. Having accomplished this, we arrive at the understanding that the resonant bandwidths of weakly nonlinear interactions in the two systems, the DIA kinetic equation and from ray methods, are different; that resonant and non-resonant interactions express themselves differently in the correlation time scale than previously understood; and that spectral transports can be significantly altered by the mean-drift term.

We demonstrate here that it is this simple difference in coordinate systems that leads to the celebrated Garrett and Munk (GM) spectrum of the oceanic internal wave field supporting a net downscale transport. Details of this spectral model are presented in [Appendix A](#). The 3-D action spectrum for the GM spectrum is independent of vertical wavenumber, so that in an Eulerian description there is no vertical wavenumber action gradient to support the diffusion of action regardless of how the vertical component of the diffusivity tensor is defined. In a ray description, the mean drift of wave packets to a high wavenumber can be explicitly represented in an ensemble transport equation and an estimate of the action (energy) available for mixing can be obtained by the counting of wave packets past a sufficiently high-wavenumber gate (e.g. Henyey *et al.* 1986).

This paper is organized as follows. Hamiltonian structures and the derivation of the transport equations from them are the focal points of §§ 2 and 3. We review the Hamiltonian structure in § 2.1. In § 2.2 we present a derivation for internal waves that leads to a Fokker–Planck equation. In § 3 we refine the Hamiltonian structure; extracting only those extreme scale-separated interactions in order to derive the Liouville equation (§ 3.1) and its subsequent Fokker–Planck equation (§ 3.2.4). Subsequent to these theoretical developments, we present estimates of energy transport to mixing scales and demonstrate the mismatch between theory and observations in § 4. We end in § 5 by discussing this contradiction in light of our derivations. The reader who is primarily interested in the disparity between observations and theory is advised to read § 4 and use the equation references to navigate §§ 2 and 3.

2. Background

2.1. Hamiltonian structure and field variables

The equations of motion satisfied by an incompressible stratified rotating flow in hydrostatic balance are

$$\left. \begin{aligned} \frac{\partial}{\partial t} \frac{\partial z}{\partial \rho} + \nabla \cdot \left(\frac{\partial z}{\partial \rho} \mathbf{u} \right) &= 0, \\ \frac{\partial \mathbf{u}}{\partial t} + f \mathbf{u}^\perp + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{\nabla M}{\rho} &= 0, \\ \frac{\partial M}{\partial \rho} - g z &= 0. \end{aligned} \right\} \quad (2.1)$$

These equations result from mass conservation, horizontal momentum conservation and hydrostatic balance. The equations are written in isopycnal coordinates with the density ρ replacing the height z in its role as an independent vertical variable. Here, $\mathbf{u} = (u, v)$ is the horizontal component of the velocity field, $\mathbf{u}^\perp = (-v, u)$, $\nabla = (\partial/\partial x, \partial/\partial y)$ is the gradient operator along isopycnals, M is the Montgomery potential

$$M = P + g\rho z, \quad (2.2)$$

with pressure P , gravity g and Coriolis parameter f . The hydrostatic approximation is used since it is an accurate approximation for long internal waves. Investigation of the non-hydrostatic effects is outside of the scope of present paper.

Here, we follow Lvov & Tabak (2004) and take (2.1) and decompose the flow into a potential and a divergence-free part

$$\mathbf{u} = \nabla\phi + \nabla^\perp\psi, \quad (2.3)$$

where

$$\nabla^\perp = \begin{pmatrix} -\partial_y \\ \partial_x \end{pmatrix}. \quad (2.4)$$

The expression for potential vorticity in these coordinates is (Haynes & McIntyre 1987)

$$Q = \frac{f + \partial v/\partial x - \partial u/\partial y}{\Pi}, \quad (2.5)$$

where $\Pi = (\rho/g)\partial^2 M/\partial\rho^2 = \rho\partial z/\partial\rho$ is a normalized differential layer thickness. Since potential vorticity is conserved along particle trajectories

$$\frac{DQ}{Dt} = 0. \quad (2.6)$$

The advection of potential vorticity in (2.6) takes place exclusively along isopycnal surfaces. Therefore, an initial distribution of potential vorticity which is constant on isopycnals, although varying across them, will remain constant. Hence, we shall utilize the following assumption:

$$Q(x, y, \rho, t) = f/\Pi_0, \quad (2.7)$$

where we redefined $\Pi \rightarrow \Pi_0 + \Pi$ to split the potential Π into its equilibrium value $\Pi_0 \equiv -g/N^2$ and deviation from it. Stratification N^2 is permitted to vary with density ρ , but is constant along isopycnals. This effectively decouples the internal wave field from lower-frequency flows such as fronts and mesoscale eddies, which are the subject of their own wave turbulence literature (e.g. Müller 1976; Kafiabad *et al.* 2019). Such internal wave–mean flow interactions can be a significant regional source of internal wave energy (Polzin 2010) and may be a key issue in determining the regional character of the internal wave field (Polzin & Lvov 2011).

The primitive equations of motion (2.1) under the assumption (2.7) can be written as a pair of canonical Hamiltonian equations

$$\frac{\partial \Pi}{\partial t} = \frac{\delta \mathcal{H}}{\delta \phi}, \quad \frac{\partial \phi}{\partial t} = -\frac{\delta \mathcal{H}}{\delta \Pi}, \quad (2.8a,b)$$

where ϕ is the isopycnal velocity potential, and the Hamiltonian is the sum of kinetic and potential energies

$$\mathcal{H} = \frac{1}{2} \int dx d\rho \left[(\Pi_0(\rho) + \Pi(x, \rho)) \left| \nabla \phi(x, \rho) + \frac{f}{\Pi_0} \nabla^\perp \Delta^{-1} \Pi(x, \rho) \right|^2 - g \left| \int^\rho d\hat{\rho} \frac{\Pi(x, \hat{\rho})}{\hat{\rho}} \right|^2 \right], \quad (2.9)$$

with $\nabla^\perp = (-\partial/\partial y, \partial/\partial x)$, where Δ^{-1} is the inverse Laplacian and $\hat{\rho}$ represents a variable of integration.

Our intent is to build a perturbation theory around analytical solutions to the linearized primitive equations as plane waves proportional to $\exp(i[\mathbf{r} \cdot \mathbf{p} - \sigma t])$. We therefore transition to the Fourier space

$$\left. \begin{aligned} \Pi(x, y, \rho) &= \frac{1}{(2\pi)^{3/2}} \int \Pi_p e^{i\mathbf{r} \cdot \mathbf{p}} d\mathbf{p}, & \phi(x, y, \rho) &= \frac{1}{(2\pi)^{3/2}} \int \phi_p e^{i\mathbf{r} \cdot \mathbf{p}} d\mathbf{p}, \\ \mathbf{p} &= (\mathbf{k}, m), & \mathbf{k} &= (k, l), & \mathbf{r} &= (x, y, \rho), \end{aligned} \right\} \quad (2.10)$$

and introduce a complex field variable a_p through the canonical transformation

$$\phi_p = \frac{iN\sqrt{\sigma_p}}{\sqrt{2g|k|}} (a_p - a_{-p}^*), \quad \Pi_p = \frac{\sqrt{g|k|}}{\sqrt{2\sigma_p}N} (a_p + a_{-p}^*). \quad (2.11a,b)$$

Wave frequency σ_p is restricted to be positive. We ignore variations in density as they multiply horizontal momentum, replacing ρ by a reference density ρ_0 (the Boussinesq approximation) and arrive at a linear dispersion frequency σ given by

$$\sigma_p = \sqrt{f^2 + \frac{g^2}{\rho_0^2 N^2} \frac{|k|^2}{m^2}}. \quad (2.12)$$

The equations of motion (2.1) adopt the canonical form

$$i \frac{\partial}{\partial t} a_p = \frac{\delta \mathcal{H}}{\delta a_p^*}, \quad (2.13)$$

with Hamiltonian

$$\begin{aligned} \mathcal{H} &= \int d\mathbf{p} \sigma_p |a_p|^2 + \int d\mathbf{p} d\mathbf{p}_1 d\mathbf{p}_2 (\delta_{\mathbf{p}+\mathbf{p}_1+\mathbf{p}_2} (U_{\mathbf{p},\mathbf{p}_1,\mathbf{p}_2} a_{\mathbf{p}}^* a_{\mathbf{p}_1}^* a_{\mathbf{p}_2}^* + \text{c.c.}) \\ &\quad + \delta_{-\mathbf{p}+\mathbf{p}_1+\mathbf{p}_2} (V_{\mathbf{p}_1,\mathbf{p}_2}^p a_{\mathbf{p}}^* a_{\mathbf{p}_1} a_{\mathbf{p}_2} + \text{c.c.})). \end{aligned} \quad (2.14)$$

Here, $V_{\mathbf{p}_1,\mathbf{p}_2}$ and $U_{\mathbf{p},\mathbf{p}_1,\mathbf{p}_2}$ are the interaction cross-sections that define the strength of nonlinear interactions between wavenumbers $\mathbf{p}$, $\mathbf{p}_1$ and $\mathbf{p}_2$ Lvov & Tabak (2001);

c.c. denotes the complex conjugate. Implicit in the canonical transformation (2.11a,b), Hamilton's equation (2.13) and Hamiltonian (2.14) is a time dependence of $e^{-i\sigma t}$. The U elements have a time dependence of $\exp(i(\sigma_{p_1} + \sigma_{p_2} + \sigma_{p_3})t)$ with $\sigma_p > 0$. They describe the creation of three waves out of nothing and therefore will not appear in the kinetic equation (2.16). Indeed, the U terms are therefore non-resonant and can be excluded from the Hamiltonian by appropriate near-identity canonical transformation (Zakharov, Lvov & Falkovich 1992).

This is the standard form of the Hamiltonian of a system dominated by three-wave interactions (Zakharov *et al.* 1992). Calculations of interaction coefficients are tedious but straightforward tasks, completed in Lvov & Tabak (2004) and Lvov *et al.* (2010). We stress that the field equation (2.13) with the three-wave Hamiltonian ((2.12), (2.14)) is equivalent to the primitive equations of motion for internal waves (2.1) under the following assumptions: potential vorticity is constant along isopycnals and is given by (2.7) and the Boussinesq approximation is made.

2.2. Wave turbulence theory

In wave turbulence theory, one proposes a perturbation expansion in the amplitude of the nonlinearity, yielding linear waves at the leading order. Wave amplitudes are modulated by the nonlinear interactions, and the modulation is statistically described by a kinetic equation (Zakharov *et al.* 1992; Nazarenko 2011) for the wave-action spectral density n_p defined by

$$n_p \delta(\mathbf{p} - \mathbf{p}') = \langle a_p^* a_{p'} \rangle. \quad (2.15)$$

Here, $\langle \dots \rangle$ denotes an ensemble averaging, i.e. averaging over many realizations of the random wave field. Application to the internal wave problem is presented in § 2b of Lvov *et al.* (2010).

2.2.1. Generalized (broadened) kinetic equation

In the limit of small nonlinearity, one develops a perturbation expansion in the nonlinearity strength, which leads, under certain assumptions, to the wave turbulence kinetic equation. The derivation of the resonant kinetic equation is well understood and studied, see Zakharov *et al.* (1992) and Nazarenko (2011). Taking non-resonant interactions leads to a different version of the kinetic equation with the frequency delta functions being replaced by a Lorentzian, see Lvov *et al.* (1997) and Lvov, Polzin & Yokoyama (2012). This derivation also hinges on the assessment that fourth-order cumulants are a subleading term compared with the product of two double correlators (Deng & Hani 2023). For the three-wave Hamiltonian (2.14), the kinetic equation is (2.16), describing general internal waves interacting in both rotating and non-rotating environments

$$\begin{aligned} \frac{\partial}{\partial t} n_p = & \int \int d\mathbf{p}_1 d\mathbf{p}_2 (|V_{p_1, p_2}^p|^2 \delta(\mathbf{p} - \mathbf{p}_1 - \mathbf{p}_2) \mathcal{L}(\Delta\sigma_{p_{12}}, \Gamma_{p_{12}}) [n_{p_1} n_{p_2} - n_p [n_{p_1} + n_{p_2}]] \\ & - |V_{p_2, p}^{p_1}|^2 \delta(\mathbf{p} - \mathbf{p}_1 + \mathbf{p}_2) \mathcal{L}(\Delta\sigma_{12p}, \Gamma_{p_{12}}) [n_{p_2} n_p - n_{p_1} [n_{p_2} + n_p]] \\ & - |V_{p, p_1}^{p_2}|^2 \delta(\mathbf{p} + \mathbf{p}_1 - \mathbf{p}_2) \mathcal{L}(\Delta\sigma_{2p1}, \Gamma_{p_{12}}) [n_p n_{p_1} - n_{p_2} [n_p + n_{p_1}]]), \end{aligned} \quad (2.16)$$

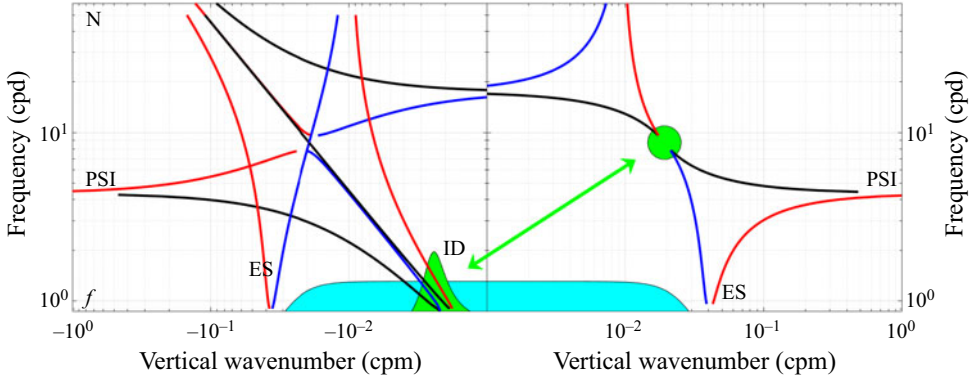


Figure 1. The resonant manifold (2.17) in the situation where the three horizontal wavevectors are either parallel or anti-parallel, plotted in a vertical wavenumber–frequency space, for a wave at the centre of the green circle. With rotation, extreme scale separations in horizontal wavenumber lead to the extreme scale-separated triads mentioned in the introduction. These triads are Bragg scattering (also called elastic scattering, or ES, and a phase velocity equals ground velocity resonance condition, called induced diffusion, or ID, being located at the Coriolis frequency f). This study focuses on the latter class, with scale separation in both horizontal and vertical wavenumbers. Near-resonant ID conditions are depicted in green, bandwidth-limited non-resonant ID forcing in cyan. The third type of extreme scale separated triads, called parametric subharmonic instability, or PSI, does not play a role in this manuscript.

where Lorentzian $\mathcal{L}$ is given by $\mathcal{L} = \Gamma_{p12}/((\Delta\sigma)^2 + \Gamma_{p12}^2)$ and $\Delta\sigma_{p12} = \sigma_p - \sigma_{p_1} - \sigma_{p_2}$ represents the distance from the resonant surface. The resonant manifold is defined by

$$\left. \begin{aligned} \sigma &= \sigma_1 + \sigma_2; & \mathbf{p} &= \mathbf{p}_1 + \mathbf{p}_2, \\ \sigma &= \sigma_1 - \sigma_2; & \mathbf{p} &= \mathbf{p}_1 - \mathbf{p}_2, \\ \sigma &= \sigma_2 - \sigma_1; & \mathbf{p} &= \mathbf{p}_2 - \mathbf{p}_1, \end{aligned} \right\} \quad (2.17)$$

and appears in figure 1. In wave turbulence theory of internal waves the importance of special extreme scale-separated triads was recognized in McComas & Bretherton (1977). These extreme scale-separated limits are called induced diffusion(ID), elastic scattering (ES) and parametric subharmonic instability (PSI). For an explanation of these triads see also McComas & Müller (1981a).

The total resonance width associated with a specific triad is given by $\Gamma_{p12} = \gamma_p + \gamma_1 + \gamma_2$ and the equation for the individual resonance widths is given by

$$\begin{aligned} \gamma_p &= \int \int d\mathbf{p}_1 d\mathbf{p}_2 (|V_{p_1, p_2}^p|^2 \delta(\mathbf{p} - \mathbf{p}_1 - \mathbf{p}_2) \mathcal{L}(\sigma - \sigma_1 - \sigma_2) [n_{p_1} + n_{p_2}] \\ &\quad + |V_{p_2, p}^{p_1}|^2 \delta(\mathbf{p} - \mathbf{p}_1 + \mathbf{p}_2) \mathcal{L}(\sigma - \sigma_1 + \sigma_2) [n_{p_2} - n_{p_1}] \\ &\quad + |V_{p, p_1}^{p_2}|^2 \delta(\mathbf{p} + \mathbf{p}_1 - \mathbf{p}_2) \mathcal{L}(\sigma + \sigma_1 - \sigma_2) [n_{p_1} - n_{p_2}]). \end{aligned} \quad (2.18)$$

Physically, this γ_p represents the fast time scale of decay of a narrow perturbation to the otherwise stationary spectrum (Lvov *et al.* 1997; Polzin & Lvov 2017). It coincides with Langevin rates estimated by Pomphrey, Meiss & Watson (1980) and the decay rate of the McComas (1977) spike experiments. The replacement of the frequency-conserving delta function by the Lorentzian takes into account not only resonant but also near-resonant and non-resonant interactions. Non-resonant interactions appear as a result of the Lorentzian decaying slowly. The role of the non-resonant interactions have to be investigated separately for each particular problem.

2.2.2. Bandwidth estimates

The broadened kinetic equation provides access to a time scale as the inverse of the bandwidth γ_p (2.18).

Considered in isolation from other interactions and in the resonant limit, the extreme scale-separated ID mechanism gives rise to (Polzin & Lvov 2017)

$$\gamma_p^{ID} = \frac{2}{\pi} \frac{ke_0 m_*}{N} \frac{\sigma^2}{f^2}, \quad (2.19)$$

where k is horizontal wavenumber magnitude, e_0 is the total energy density in the GM model and m_* is the vertical wavenumber bandwidth parameter; see Appendix A. In the context of a broadened kinetic equation, γ_p^{ID} represents the decay of a spike introduced into an otherwise smooth spectrum (Lvov *et al.* 1997) and is identified as such by McComas & Müller (1981a) in the numerical experiments of McComas (1977). The associated time scale is far smaller than a wave period and at loggerheads with the intent of having (2.16) describe the slow time evolution of the spectral density.

At finite, but yet small-amplitude, analysis of (2.18) reveals that the bandwidth is composed of both resonant γ_{res} and non-resonant terms

$$\left. \begin{array}{l} \text{resonant: } \gamma_{res} = \frac{2}{\pi} \frac{ke_0 m_*}{N} \left(\frac{\sigma}{f}\right)^2 \\ \text{non-resonant: } \frac{2}{\pi} \frac{k^2 e_0}{f^2} \gamma_{res} \end{array} \right\}. \quad (2.20)$$

The non-resonant contributions come from vertical scales that are larger than those of the high-frequency wave and come from the tails of the Lorentzian admitting non-resonant interactions with the highly energetic near-inertial wave field.

The non-resonant contributions increase at a greater rate with respect to increasing amplitude than resonant contributions. At realistic oceanic amplitudes, numerical evaluation (Polzin & Lvov 2017) demonstrates that γ_p^{ID} is proportional to the root-mean-square (r.m.s.) Doppler shift. The difficulty of placing a physical interpretation on such a short time decay time scale (2.19) and suspicion that the broadened kinetic equation would suffer from the Galilean invariance that plagued Kraichnan's 1959 DIA model for 3-D turbulence led to scepticism and extensive commentary in the literature (Holloway 1980, 1982). Addressing these issues required development of a broadened kinetic equation, which in turn required waiting for the derivation in canonical coordinates presented at the beginning of this section.

2.2.3. Fokker–Planck diffusion limit

Following McComas & Bretherton (1977) for the resonant kinetic equation, we start from (2.16) and pick off the interactions having $\mathbf{p}$ nearly parallel to $\mathbf{p}_1$ with $\mathbf{p}_2$ small, or nearly parallel to $\mathbf{p}_2$ with $\mathbf{p}_1$ small, which selects the ID class triads. For a sufficiently red spectrum, this permits discarding of the small $n_p n_{p1}$ ($n_p n_{p2}$, respectively) terms. We then rewrite (2.16) as

$$\frac{\partial n_p}{\partial t} = \int d\mathbf{q} (\mathcal{B}(\mathbf{p}) - \mathcal{B}(\mathbf{p} + \mathbf{q})) \simeq - \int d\mathbf{q} \left(\mathbf{q} \cdot \frac{\partial}{\partial \mathbf{p}} \right) \mathcal{B}(\mathbf{p}), \quad (2.21)$$

where we introduced

$$\mathcal{B}(\mathbf{p}) = 8\pi \int |V_{p_1, q}^p|^2 n_q (n_{p1} - n_p) \delta_{\mathbf{p}-\mathbf{p}_1-\mathbf{q}} \mathcal{L}(\sigma_p - \sigma_{p1} - \sigma_q) d\mathbf{p}_1, \quad (2.22)$$

and expanded the difference ($\mathcal{B}(\mathbf{p}) - \mathcal{B}(\mathbf{p} + \mathbf{q})$) in a Taylor series using $\mathbf{q}$ to represent the small difference in wavenumber between the two high-frequency waves. Expanding the difference $n_{p_1} - n_p$ for small $\mathbf{q}$ gives

$$\mathcal{B}(\mathbf{p}) \simeq -8\pi \left(\mathbf{q} \cdot \frac{\partial n_p}{\partial \mathbf{p}} \right) \int |V_{p_1, q}^p|^2 n_q \delta_{p-p_1-q} \mathcal{L}(\sigma_p - \sigma_{p_1} - \sigma_q) d\mathbf{p}_1. \quad (2.23)$$

Combining (2.21) with (2.23) we obtain

$$\frac{\partial n_p}{\partial t} = 8\pi \int d\mathbf{q} \left(\mathbf{q} \cdot \frac{\partial}{\partial \mathbf{p}} \right) \left(\mathbf{q} \cdot \frac{\partial n_p}{\partial \mathbf{p}} \right) \left[\int |V_{p_1, q}^p|^2 n_q \delta_{p-p_1-q} \mathcal{L}(\sigma_p - \sigma_{p_1} - \sigma_q) d\mathbf{p}_1 \right], \quad (2.24)$$

or

$$\left. \begin{aligned} \frac{\partial n_p}{\partial t} &= \frac{\partial}{\partial p_i} D_{ij}(\mathbf{p}, \mathbf{q}) \frac{\partial}{\partial p_j} n_p, \\ D_{ij}(\mathbf{p}, \mathbf{q}) &= 8\pi \int d\mathbf{q} (q_i q_j) |V_{p_1, q}^p|^2 n_q \delta_{p-p_1-q} \mathcal{L}(\sigma_p - \sigma_{p_1} - \sigma_q) d\mathbf{p}_1. \end{aligned} \right\} \quad (2.25)$$

This is a Fokker–Planck diffusion equation describing the diffusion of wave action in the system dominated by non-local-in-wavenumber interactions. Comparison between the Fokker–Planck equation (2.25) obtained here, and the similar Fokker–Planck equation (obtained using Wentzel–Kramers–Brillouin (WKB) theory (§ 3.2.4 below) will lead to critical insights into the spectral energy transfers in internal wave systems and ultimately to the parametrization of the energy supply to internal wave breaking processes.

Extended versions of the Garrett and Munk model (Appendix A) can then be inserted into (2.25) to arrive at families of stationary states (e.g. McComas & Müller 1981*b*; Polzin & Lvov 2011; Dematteis *et al.* 2022) and estimate downscale transport. We will require transport estimates the vertical–vertical component of the diffusivity tensor, D_{33} . For the Garrett and Munk model (GM76),

$$D_{33} = \frac{2}{\pi} \frac{km^2 e_0 m_*}{N}. \quad (2.26)$$

Here, k is horizontal wavenumber magnitude, e_0 is the total energy, m_* is a bandwidth parameter and N is buoyancy frequency.

3. Wave–wave interactions in the scale-separated limit

In our previous studies Lvov *et al.* (2010) we have seen that, under a scale-invariant assumption, the integrals in the kinetic equation tend to diverge for small or large wavenumbers or both. Therefore the interactions via extreme scale separations play an important role in energy exchanges in internal waves. In this section, we are going to develop a rigorous formalism based on WKB techniques to study such interactions.

3.1. The primitive equations and Hamiltonian structure

3.1.1. Reynolds decomposition and Hamiltonian structure

To study interactions between long and short waves we start at (2.8a,b) and make a Reynolds decomposition in wave amplitude

$$\Pi \rightarrow \Pi_0 + \Pi + \pi', \quad \phi = \Phi + \phi', \quad \psi = \Psi + \psi'. \quad (3.1a-c)$$

Here, the large-amplitude waves are represented with Π , Φ , Ψ and small-amplitude waves are given by ϕ' , π' and ψ' . Given the potentials Φ and ϕ , the corresponding velocities are

$$\mathcal{U} = \nabla \Phi + \nabla^\perp \Psi, \quad u' = \nabla \phi' + \nabla^\perp \psi'. \quad (3.2a,b)$$

To simplify the presentation, we will utilize the non-rotating approximation ($f = 0$) in which $(\nabla^\perp \Psi, \nabla^\perp \psi) \rightarrow 0$. The case of rotating ocean $f \neq 0$ is presented in Appendix.

We substitute the Reynolds decomposition (3.1a-c) into the equations of motion (2.1). We then make an assumption that the large scales Π and Φ exactly satisfy the primitive equations of motion. In other words, gradient operator applied to the inertial waves with no horizontal structure will return zero, thus these inertial waves are exact analytical solutions of the primitive Boussinesq equations (2.1). For the application of oceanic internal waves, this assumption is realized if the large-scale waves are a collection of inertial waves having frequency $\sigma = f$: inertial waves do not have horizontal structure and thus a superposition of inertial waves exactly satisfies the primitive equations of motion. We then subtract equations for the large-amplitude waves. The result is given by

$$\left. \begin{aligned} \dot{\pi}' + \nabla \cdot ((\Pi_0 + \Pi + \pi') \nabla \Phi) + \nabla(\pi' \Phi) &= 0, \\ \dot{\phi}' + \frac{|\nabla \phi'|^2}{2} + \nabla \phi' \cdot \nabla \Phi + \frac{g}{\rho_0^2} \int \int d\rho' d\rho'' \pi' &= 0. \end{aligned} \right\} \quad (3.3)$$

In these equations, Π and Φ are given time-space-dependent functions representing the large-amplitude waves.

These equations are also Hamilton's equations,

$$\frac{\partial \pi'}{\partial t} = \frac{\delta \mathcal{H}}{\delta \phi'}, \quad \frac{\partial \phi'}{\partial t} = -\frac{\delta \mathcal{H}}{\delta \pi'}, \quad (3.4a,b)$$

with the time-dependent Hamiltonian given by (this Hamiltonian may be obtained by substituting (3.1a-c) to (2.9))

$$\mathcal{H} = \frac{1}{2} \int d\mathbf{r} \left((\Pi_0 + \Pi + \pi') |\nabla \phi'|^2 + 2\pi' \nabla \phi' \cdot \nabla \Phi - g \left| \int^\rho d\hat{\rho} \frac{\pi'}{\hat{\rho}} \right|^2 \right). \quad (3.5)$$

There are two types of terms here – those that will ultimately describe a sea of interacting small-scale waves and those that will describe the influence of large-amplitude large-scale waves on the small-scale waves. In our previous efforts (Lvov & Tabak 2001, 2004; Lvov *et al.* 2010; Polzin & Lvov 2011, 2017) these terms are comingled. Comparing (2.9) and (3.5), we see that (3.5) contains additional terms $\Pi |\nabla \phi'|^2$ and $\pi' \nabla \phi' \cdot \nabla \Phi$ that are explicit representations of what will be scale-separated interactions. The term $\pi' \nabla \phi' \cdot \nabla \Phi$ describes the advection of the small-scale internal field by the given large-scale large-amplitude field. The term $\Pi |\nabla \phi'|^2$ represents a coupling of small scales to large through changes in the stratification by the large-scale wave. The term $\pi' |\nabla \phi'|^2$ will represent interactions local in wavenumber.

We now express the space-dependent variables $\pi'(\mathbf{r})$, $\Pi(\mathbf{r})$, $\phi'(\mathbf{r})$ and $\Phi(\mathbf{r})$ in terms of their Fourier images $\pi'(\mathbf{p})$, $\Pi(\mathbf{p})$, $\phi'(\mathbf{p})$ via (2.10), make the Boussinesq approximation $\Pi/\rho \simeq \Pi/\rho_0$ and use $\int d\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{r}} = (2\pi)^3 \delta(\mathbf{p})$ to obtain

$$\left. \begin{aligned} \mathcal{H} &= \mathcal{H}_{\text{linear}} + \mathcal{H}_{\text{nonlinear}}, \\ \mathcal{H}_{\text{linear}} &= \frac{1}{2} \int d\mathbf{p} \left(\Pi_0 |\mathbf{k}|^2 |\phi'_p|^2 - \frac{g}{\rho_0^2} \frac{|\phi'_p|^2}{m^2} \right), \\ \mathcal{H}_{\text{nonlinear}} &= \mathcal{H}_{\text{local}} + \mathcal{H}_{\text{sweeping}} + \mathcal{H}_{\text{density}}, \end{aligned} \right\} \quad (3.6)$$

$$\left. \begin{aligned} \mathcal{H}_{\text{local}} &= -\frac{1}{2(2\pi)^{3/2}} \int d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{p}_3 \delta(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \mathbf{k}_2 \cdot \mathbf{k}_3 \pi'_{p_1} \phi'_{p_2} \phi'_{p_3}, \\ \mathcal{H}_{\text{sweeping}} &= -\frac{1}{2(2\pi)^{3/2}} \int d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{p}_3 \delta(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) 2\mathbf{k}_1 \cdot \mathbf{k}_2 \Phi_{p_1} \phi'_{p_2} \pi'_{p_3}, \\ \mathcal{H}_{\text{density}} &= -\frac{1}{2(2\pi)^{3/2}} \int d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{p}_3 \delta(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \mathbf{k}_2 \cdot \mathbf{k}_3 \Pi_{p_1} \phi'_{p_2} \phi'_{p_3}. \end{aligned} \right\} \quad (3.7)$$

The Hamiltonian $\mathcal{H}_{\text{nonlinear}}$ is the sum of three terms:

- (i) $\mathcal{H}_{\text{local}}$ represents small amplitudes interacting with small amplitudes. We will call this term ‘local’ interactions in anticipation of making a scale separation between large-amplitude large-scale and small-amplitude small-scale waves in §§ 3.1.2 and 3.1.3;
- (ii) $\mathcal{H}_{\text{density}}$ is the term that describes the variations of stratification that small-amplitude waves experience due to the compression and rarefaction of isopycnals associated with the large-amplitude waves. We will refer to this term as a density term;
- (iii) $\mathcal{H}_{\text{sweeping}}$ is the term that describes the advection (sweeping) of small-amplitude waves by large-amplitude waves. In the future, we refer to this term as a sweeping term.

The main focus of this manuscript is to investigate how the density and sweeping terms affect the overall spectral energy density.

3.1.2. Sweeping Hamiltonian

Following the traditional wave turbulence approach, we make a transformation to the wave-action variables that represent wave amplitude and phase

$$\phi'_p = \frac{iN\sqrt{\sigma_p}}{\sqrt{2g|\mathbf{k}|}} (a_p - a_{-p}^*), \quad \pi'_p = \frac{\sqrt{g}|\mathbf{k}|}{\sqrt{2\sigma_p}N} (a_p + a_{-p}^*). \quad (3.8a,b)$$

We substitute (3.8a,b) into $\mathcal{H}_{\text{sweeping}}$ of (3.7), and obtain

$$\begin{aligned} \mathcal{H}_{\text{sweeping}} &= -\frac{1}{2(2\pi)^{3/2}} \int d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{p}_3 \delta(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) 2\mathbf{k}_1 \cdot \mathbf{k}_2 \Phi_{p_1} \frac{iN\sqrt{\sigma_{p_2}}}{\sqrt{2g|\mathbf{k}_2|}} \frac{\sqrt{g}|\mathbf{k}_3|}{\sqrt{2\sigma_{p_3}}N} \\ &\quad \times (a_{p_2} - a_{-p_2}^*)(a_{p_3} + a_{-p_3}^*). \end{aligned} \quad (3.9)$$

The next step is algebraically trivial but conceptually fundamental. We invoke an extreme scale-separated limit in which the two small-amplitude waves have similar frequency and horizontal wavenumber magnitude. No condition is required on the

vertical wavenumber. This conditioning retains both the ID and Bragg scattering (ES) branches of the resonant manifold; see figure 1.

In this scale-separated limit of the internal wave problem, $\sigma_{p_2} \cong \sigma_{p_3}$ and $|k_2| \cong |k_3|$. Beyond the obvious algebraic simplifications, upon expanding the brackets, we find terms of the type $a_{p_2} a_{p_3}$, $a_{-p_2}^* a_{-p_3}^*$ and $a_{p_2} a_{-p_3}^*$, $a_{-p_2}^* a_{-p_3}$. In what follows, we neglect the $a_{p_2} a_{p_3}$ and $a_{-p_2}^* a_{-p_3}^*$ terms and retain the $a_{p_2} a_{-p_3}^*$, $a_{-p_2}^* a_{-p_3}$ terms, since the former terms are non-resonant, while the latter may be in the resonance for some wavenumbers. The discarded terms lead to a process when one lower-frequency wave decays into two high-frequency waves and thus the frequencies do not sum to zero. Such decay is a non-resonant process, so we can remove these terms at the onset.

After relabelling subscripts, in which $2 \rightarrow 1$ and $3 \rightarrow 2$, we obtain

$$\left. \begin{aligned} \mathcal{H}_{\text{sweeping}} &= \int d\mathbf{p}_1 d\mathbf{p}_2 A_{\text{sweeping}}(\mathbf{p}_1, \mathbf{p}_2) a_{p_1} a_{p_2}^*, \\ &\quad \text{with} \\ A_{\text{sweeping}}(\mathbf{p}_1, \mathbf{p}_2) &= -\frac{1}{2}i (\mathbf{k}_1 - \mathbf{k}_2) \cdot (\mathbf{k}_1 + \mathbf{k}_2) \Phi_{p_1 - p_2}. \end{aligned} \right\} \quad (3.10)$$

In the limit of large vertical background scales, (3.10) describes a quasi-coherent translation of small-scale small-amplitude waves by the large-scale background. At $1/2$ the vertical scale of $\mathbf{p}_1$ and $\mathbf{p}_2$, (3.10) describes a Bragg scattering process. This is distinct from ‘local’ interactions as the large-amplitude wave has a much larger horizontal scale.

3.1.3. Density Hamiltonian

We now can repeat the same steps for the density Hamiltonian. We substitute (3.8a,b) into the $\mathcal{H}_{\text{density}}$ of (3.7), use $\sigma_{p_2} \cong \sigma_{p_3}$ and $|k_2| \cong |k_3|$, relabel subscripts and obtain

$$\left. \begin{aligned} \mathcal{H}_{\text{density}} &= \int d\mathbf{p}_1 d\mathbf{p}_2 A_{\text{density}}(\mathbf{p}_1, \mathbf{p}_2) a_{p_1} a_{p_2}^*, \\ &\quad \text{with} \\ A_{\text{density}}(\mathbf{p}_1, \mathbf{p}_2) &= \frac{1}{2\Pi_0} \Pi_{p_2 - p_1} \sqrt{\sigma_{p_1} \sigma_{p_2}}. \end{aligned} \right\} \quad (3.11)$$

In the limit of large vertical background scales, (3.11) describes the modulation of the background stratification. At $1/2$ the vertical scale of $\mathbf{p}_1$ and $\mathbf{p}_2$, (3.11) describes a Bragg scattering process. This is distinct from ‘local’ interactions as the large-amplitude wave has a much larger horizontal scale.

3.1.4. Quadratic Hamiltonian for inhomogeneous wave turbulence

Let us neglect the local interaction term $\mathcal{H}_{\text{local}}$ in Hamiltonian (3.7). Then the Hamiltonian (3.7) of small-amplitude small horizontal scale internal waves a_p superimposed into a field of large-amplitude large horizontal scale internal waves given by space–time-dependent Π

and Φ are given by a form that is quadratic in the small-scale variables a_p

$$\left. \begin{aligned} \mathcal{H} &= \int d\mathbf{p}_1 d\mathbf{p}_2 A(\mathbf{p}_1, \mathbf{p}_2) a_{\mathbf{p}_1} a_{\mathbf{p}_2}^*, \\ &\text{with} \\ A(\mathbf{p}_1, \mathbf{p}_2) &= \sigma_p \delta(\mathbf{p}_1 - \mathbf{p}_2) + A_{\text{sweeping}}(\mathbf{p}_1, \mathbf{p}_2) + A_{\text{density}}(\mathbf{p}_1, \mathbf{p}_2), \end{aligned} \right\} \quad (3.12)$$

where $A_{\text{sweeping}}(\mathbf{p}_1, \mathbf{p}_2)$ and $A_{\text{density}}(\mathbf{p}_1, \mathbf{p}_2)$ are given in (3.10) and (3.11). The $A(\mathbf{p}_1, \mathbf{p}_2)$ are time dependent and depend upon the phases of the external field.

3.2. Wentzel–Kramers–Brillouin approach

At this point, we have a three-wave Hamiltonian (3.12) with one large-amplitude large horizontal scale wave interacting with two smaller-amplitude smaller horizontal scale waves that have similar frequencies. Below, we assume a nearly resonant paradigm and perform algebraic manipulations that take into account the ID portion of the resonant manifold.

3.2.1. The wave packet transport equation

In the spatially homogeneous wave turbulence of § 2.2 we used wave-action density n_p and a linear dispersion relation σ_p , both being a function of a wavenumber $\mathbf{p}$. In the case when there is a slowly varying large-scale background, i.e. a system where spatial inhomogeneity is present, the properties of wave action and the dispersion relation will depend on the position in space. Then it makes sense to introduce an additional parameter, a position vector $\mathbf{r}$ in wave action and linear dispersion relation. The theory for this spatial dependence is developed in Gershgorin, Lvov & Nazarenko (2009) using a Gabor transform to represent the envelope structure describing the spatial localization and carrier frequency. The leading-order balance in Gershgorin *et al.* (2009) leads to action and phase conservation along ray paths. The balance on the envelope scale implies phase modulation, for which we direct the reader to the appendix of Cohen & Lee (1990) for clarity. Associated with the envelope structure is a residual circulation (Bühler & McIntyre 2005). The potential for the wave packet to interact with its envelope structure is possible (Bühler & McIntyre 2005; Dosser & Sutherland 2011) but would require a modification of the uniform potential vorticity statement (2.7). It is at this stage that one might also want to consider the potential for nonlinear wave steepening effects associated with $\mathcal{H}_{\text{local}}$ to counter the dispersion of ray characteristics as a precursor to the description of the solitary wave dynamics.

The familiar statement of action conservation that we are after is obtained in Gershgorin *et al.* (2009) by assuming the scale of the envelope structure is large in comparison with the inverse wavenumber of the small-scale wave, i.e. the wave packet contains many oscillations. The result required here can be obtained more simply by using a Wigner transform alone to define the space–time-dependent wave-action spectral density

$$n_{p,r} \equiv \int e^{i\mathbf{q} \cdot \mathbf{r}} \langle a_{\mathbf{p}+\mathbf{q}/2} a_{\mathbf{p}-\mathbf{q}/2}^* \rangle d\mathbf{q}, \quad (3.13)$$

in which the transform variable $\mathbf{q}$ is the difference between two large wavenumbers $\mathbf{p}_1$ and $\mathbf{p}_2$, $\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_2$ and the field variables a are evaluated at $\mathbf{p} \pm \mathbf{q}/2$. We similarly introduce

the space-dependent intrinsic frequency $\omega_{p,r}$

$$\omega_{p,r} = \int e^{ir \cdot q} A(\mathbf{p} + \mathbf{q}/2, \mathbf{p} - \mathbf{q}/2) d\mathbf{q}. \quad (3.14)$$

This is a generalization of our ‘traditional’ wave action (2.15) which allows for slow variations of the background. It will incorporate the large vertical scale contributions of $\mathcal{H}_{sweeping}$ and $\mathcal{H}_{density}$.

To derive the transport equation, we take definition (3.13), differentiate it with respect to time, use Hamilton’s equation of motion (2.13), with the Hamiltonian (3.12). The spatial dependence is then made explicit by representing $A(\mathbf{p}_1, \mathbf{p}_2)$ and $a(\mathbf{p}_1)a^*(\mathbf{p}_2)$ with their corresponding Fourier transforms. After an inspired change of variables (Lvov & Rubenchik 1977; Gershgorin *et al.* 2009) that maps the ID portion of the resonant manifold, one obtains an intermediary expression

$$\begin{aligned} i \frac{\partial n_{p,r}}{\partial t} &= \int \frac{d\mathbf{r}' d\mathbf{r}''}{(2\pi)^6} d\mathbf{q}' d\mathbf{q}'' \exp(i\mathbf{q}' \cdot (\mathbf{r} - \mathbf{r}') + i\mathbf{q}'' \cdot (\mathbf{r} - \mathbf{r}'')) \\ &\times [\omega_{p+q''/2}(\mathbf{r}') n_{p-q'/2}(\mathbf{r}'') - \omega_{p-q''/2}(\mathbf{r}') n_{p+q'/2}(\mathbf{r}'')]. \end{aligned} \quad (3.15)$$

After expanding ω and n in Taylor series with respect to $\mathbf{p}$ and truncating higher-order terms, one ultimately arrives at the action balance

$$\frac{\partial n_{p,r}}{\partial t} + \nabla_p \sigma_{p,r} \cdot \nabla_r n_{p,r} - \nabla_r \sigma_{p,r} \cdot \nabla_p n_{p,r} = 0, \quad (3.16)$$

or alternately

$$\frac{\partial n_{p,r}}{\partial t} + \nabla_r \cdot [n_{p,r} \nabla_p \sigma_{p,r}] - \nabla_p \cdot [n_{p,r} \nabla_r \sigma_{p,r}] = 0. \quad (3.17)$$

Integration of (3.17) over wavenumber provides a connection to the space–time variational formulations found in Witham (1974, Chapters 11.7 and 14). The Bragg scattering process residing in the scale-separated Hamiltonian (3.10) has been eliminated by using the Wigner transform and Taylor series expansion that serendipitously exploits the symmetries associated with the ID resonance.

3.2.2. Application for internal waves

We now take the expression for A_{p_1, p_2} from (3.12) and substitute it into (3.14) for the space-dependent linear dispersion relation. The result is given by

$$\omega_{p,r} = \sigma_p - \mathbf{k} \cdot \mathcal{U}(\mathbf{r}) + \sigma_p \frac{\Pi(\mathbf{r})}{2\Pi_0}, \quad (3.18)$$

where $\mathbf{p} = (\mathbf{k}, m)$ is the wavevector, $\mathcal{U}(\mathbf{r})$ is the time-dependent horizontal velocity of the external large-scale wave field and $\Pi(\mathbf{r})$ is the time-dependent stratification of the external wave field. Here, the σ_p term was produced by the term proportional to $\delta(\mathbf{p}_1 - \mathbf{p}_2)$, $\mathbf{k} \cdot \mathcal{U}$ comes from the sweeping term in the Hamiltonian and $\sigma_p \Pi / 2\Pi_0$ comes from density term in the Hamiltonian. The frequency σ_p is given by (2.12) using $f = 0$. For a high-frequency wave in a rotating ocean, $\mathcal{U}$ is replaced by $\nabla \Phi + \nabla^\perp \Psi$. The derivation for $f \neq 0$ appears in the Appendix.

3.2.3. The ray-path wave packet transport equation

The transport equation (3.16) or its alternative formulation (3.17) with the space–time-dependent linear dispersion relationship (3.18) is a fundamental result expressing wave-action conservation that provides the basis for the analyses to follow.

The action balance (3.16) can be simply solved by the method of characteristics. The characteristics of (3.16), also called rays, are defined by

$$\dot{\mathbf{r}}(t) \equiv \nabla_{\mathbf{p}} \sigma_{\mathbf{p},r}, \quad \dot{\mathbf{p}}(t) \equiv -\nabla_{\mathbf{r}} \sigma_{\mathbf{p},r}. \quad (3.19a,b)$$

Equations (3.19a,b) imply that wave-action spectral density is conserved along these characteristics

$$\frac{\partial n_{\mathbf{p},r}}{\partial t} + \dot{\mathbf{r}} \cdot \nabla_{\mathbf{r}} n_{\mathbf{p},r} + \dot{\mathbf{p}} \cdot \nabla_{\mathbf{p}} n_{\mathbf{p},r} = 0. \quad (3.20)$$

This is the classical representation for the conservation of action spectral density along ray trajectories. A derivation for a general Hamiltonian set is presented in Gershgorin *et al.* (2009), here, the derivation is specifically for internal waves in the scale-separated limit. Integration over wavenumber provides the space–time result presented in Witham (1974). This result often appears as an analogy to Liouville’s theorem for the conservation of the phase volume of particles without justification. Note that it is the balance of first-order terms in a Taylor series expansion of (3.15).

3.2.4. An ensemble path transport equation

In what follows we derive a combined advective–diffusive transport equation for the action balance (3.16) by generalizing an approach found in Nazarenko *et al.* (2001). There are strong parallels here to the discussion of Taylor (1921) that appears in the appendix of McComas & Bretherton (1977): the illuminating analogy with particle dispersion in Taylor (1921) is to substitute wavenumber $\mathbf{p}$ for the Lagrangian particle position $\mathbf{r}$ and obtain a quantitative approach for discussing the migration and dispersion of wave packets in the spectral domain following ray trajectories. Having said this, it should be intuitively obvious that, since particle dispersion in $\mathbf{r}$ has little to do with resonance, neither should the issue of dispersion of $\mathbf{p}$ in phase space be intrinsically tied to resonance! Yet, an emphasis on the resonant paradigm is the interpretive context pursued in McComas & Bretherton (1977) and Nazarenko *et al.* (2001). A second key departure is that our motivation stems from the fact that the GM76 spectrum is a no-flux solution to the Fokker–Planck equation (2.25). We acknowledge the oceanographic literature in this regard and so are focused upon the issue of a mean drift of wave packets to high wavenumber that is accessible in ray-tracing simulations (e.g. Henyey *et al.* 1986) but not explicitly represented in (2.25). To underscore the distinction, the existence of a mean drift implies relative dispersion (e.g. Bennett 1984) rather than the issue of absolute dispersion addressed in Taylor (1921).

The very first step in our derivation is to specify a wave packet ensemble mean drift and departures from the mean drift. This decomposition is an improvement on the arguments presented in McComas & Bretherton (1977), Nazarenko *et al.* (2001) and Lanchon & Cortet (2023). Our limited knowledge of the ray literature does not permit us to comment on the originality of our interpretation. It is, however, crucial in understanding intrinsic differences between the kinetic equation and ray theory. Given the nature of our understanding, the issue assuredly carries over to other physical problems such as the interaction of near-inertial waves with lower-frequency flows, (e.g. Young & Jelloul 1997; Kafiabad *et al.* 2019; Dong, Bühler & Smith 2020) surface gravity wave interaction

with lower-frequency flows (Villas Bôas & Young 2020) and Rossby wave–Rossby wave interactions (Nazarenko 2011).

We represent the wave action of a single wave packet as a sum of a spatially homogeneous part $\bar{n}$ and small ‘wiggles’ $\tilde{n}$

$$n_{p,r} = \bar{n}_p + \tilde{n}_{p,r}; \quad n_p = \int d\mathbf{r} n_{p,r}; \quad \int d\mathbf{r} \tilde{n}_{p,r} = 0, \quad (3.21a-c)$$

and acknowledge the presence of a mean drift in the spectral domain by adding zero

$$\dot{\mathbf{p}} = \dot{\mathbf{p}} - \langle \dot{\mathbf{p}} \rangle + \langle \dot{\mathbf{p}} \rangle, \quad (3.22)$$

in which $\langle \dots \rangle$ is an ensemble average for a system with spatially homogenous statistics, so that $\langle \dots \rangle$ is independent of $\mathbf{r}$. The mean drift arises due to inhomogeneities of the ray-path statistics in the spectral domain. Please note that the dimensions of $\bar{n}_p$ and n_p are different.

Starting from the flux form of the action balance (3.17), we substitute (3.21a–c), (3.22) and invoke an ensemble average and integrate over $\mathbf{r}$ to obtain

$$\frac{\partial \langle n_p \rangle}{\partial t} = - \int d\mathbf{r} \langle \nabla_p \cdot [\dot{\mathbf{p}} - \langle \dot{\mathbf{p}} \rangle] n_{p,r} \rangle - \nabla_p \cdot [\langle \dot{\mathbf{p}} \rangle \langle n_p \rangle]. \quad (3.23)$$

Closure of this equation depends upon writing $n_{p,r}$ in terms of n_p . At this juncture, we invoke that property that wave-action spectral density does not change along trajectories

$$\begin{aligned} n_{p,r} &\equiv n(\mathbf{p}(t), \mathbf{r}(t), t) = n(\mathbf{p}(t-T), \mathbf{r}(t-T), t-T) \\ &= \bar{n} \left[\mathbf{p}(t) - \int_{t-T}^t \dot{\mathbf{p}}(t') dt'; t-T \right] \\ &\quad + \tilde{n} \left[\mathbf{p}(t) - \int_{t-T}^t \dot{\mathbf{p}}(t') dt'; \mathbf{r}(t) - \int_{t-T}^t \dot{\mathbf{r}}(t') dt'; t-T \right]. \end{aligned} \quad (3.24)$$

We then execute a Taylor series expansion of $\bar{n}$

$$n(\mathbf{p}(t-T), \mathbf{r}(t-T), t-T) \simeq \bar{n}(\mathbf{p}(t); t-T) - \nabla_p \bar{n}(\mathbf{p}(t); t-T) \cdot \int_{t-T}^t \dot{\mathbf{p}}(t') dt' + \tilde{n}(\dots), \quad (3.25)$$

substitute, add zero once again and, using the definition of ensemble averaging $\langle \dots \rangle$,

$$\langle \nabla_p [\dot{\mathbf{p}} - \langle \dot{\mathbf{p}} \rangle] \bar{n}(\mathbf{p}(t), t-T) \rangle \cong \nabla_p \langle \dot{\mathbf{p}} - \langle \dot{\mathbf{p}} \rangle \rangle \langle \bar{n}(\mathbf{p}(t), t-T) \rangle = 0, \quad (3.26)$$

and

$$\begin{aligned} &\left\langle \nabla_p [\dot{\mathbf{p}} - \langle \dot{\mathbf{p}} \rangle] \int_{t-T}^t \langle \dot{\mathbf{p}} \rangle dt' \cdot \nabla_p \bar{n}(\mathbf{p}(t), t-T) \right\rangle \\ &\cong \nabla_p \langle \dot{\mathbf{p}} - \langle \dot{\mathbf{p}} \rangle \rangle \int_{t-T}^t \langle \dot{\mathbf{p}} \rangle dt' \cdot \nabla_p \langle \bar{n}(\mathbf{p}(t), t-T) \rangle \\ &= 0, \end{aligned} \quad (3.27)$$

as $\langle \dot{\mathbf{p}} - \langle \dot{\mathbf{p}} \rangle \rangle \equiv 0$ and neglecting an initial transient term

$$\langle \nabla_p [\dot{\mathbf{p}}(t) - \langle \dot{\mathbf{p}} \rangle] \tilde{n}(\mathbf{p}(t-T), \mathbf{r}(t-T), t-T) \rangle, \quad (3.28)$$

so that (3.23) becomes

$$\frac{\partial \langle n_p \rangle}{\partial t} = -\nabla_p \cdot \int_{t-\tau}^t \langle [\dot{\mathbf{p}}(t) - \langle \dot{\mathbf{p}} \rangle][\dot{\mathbf{p}}(t' - \tau) - \langle \dot{\mathbf{p}} \rangle] \rangle dt' \nabla_p \langle n_p \rangle - \nabla_p \langle \dot{\mathbf{p}} \rangle \langle n_p \rangle. \quad (3.29)$$

We introduce the auto-lag covariance matrix

$$C_{ij}(\mathbf{p}, t, t') = \langle [\dot{\mathbf{p}}(\mathbf{r}(t)) - \langle \dot{\mathbf{p}}(\mathbf{r}(t)) \rangle]_i [\dot{\mathbf{p}}(\mathbf{r}(t')) - \langle \dot{\mathbf{p}}(\mathbf{r}(t')) \rangle]_j \rangle. \quad (3.30)$$

The final result is then

$$\frac{\partial \langle n_p \rangle}{\partial t} = -\nabla_{p_i} \cdot \int_{t-\tau}^t C_{ij}(\mathbf{p}, t, t') dt' \cdot \nabla_{p_j} \langle n_p \rangle - \nabla_{p_i} \langle \dot{\mathbf{p}}_i \rangle \langle n_p \rangle. \quad (3.31)$$

Convergence of the time integral, in which one can replace the lower limit of integration by $-\infty$, is the hallmark of a Markov approximation which we investigate further in Polzin & Lvov (2023). This equation encapsulates our fundamental theoretical result: the transport equation changes from diffusion (2.25) to an expression that involves both advection and diffusion in the spectral domain. Instead of being a no-flux stationary state, the GM76 spectrum, for which $\langle n_p \rangle \propto m^0$, now supports a downscale action flux.

Our decomposition of wavenumber tendency into mean drift and dispersion about that mean drift provides a concrete mathematical interpretation for an interaction time scale τ_i

$$\tau_i^{-1} = |\langle \dot{\mathbf{p}} \rangle| / |\mathbf{p}| \quad (3.32)$$

and correlation time scale τ_c :

$$\tau_c = \int_{-\infty}^t C_{ii}(\mathbf{p}, t, t') dt' / \langle |\dot{\mathbf{p}}(\mathbf{r}(t)) - \langle \dot{\mathbf{p}}(\mathbf{r}(t)) \rangle|^2 \rangle. \quad (3.33)$$

Intuitive notions of the interplay between τ_i and τ_c are discussed in Müller *et al.* (1986) and in Nazarenko *et al.* (2001) using expressions for C_{ij} (3.30) in which the ensemble mean drift has not been subtracted. The sentiment in Müller *et al.* (1986) is that a separation between τ_c and τ_i is problematic for the oceanic internal wave field. Numerical ray-tracing results (Henyey & Pomphrey 1983; Henyey, Pomphrey & Meiss 1984) do not elucidate why this might be. Our decomposition of wavenumber tendency into mean drift and dispersion about that mean drift, the revised Fokker–Planck (3.31) and revised covariance matrix (3.30) provide a concrete mathematical interpretation for such judgements about τ_i and τ_c .

Having summoned the analogy between particle dispersion Taylor (1921) and dispersion of wave packets in wavenumber, we are led to a degree of scepticism concerning the McComas & Bretherton (1977) and Nazarenko *et al.* (2001) interpretation that ray tracing should collapse onto the Fokker–Planck equation and diffusivity derived from the resonant kinetic equation (2.25). Convergence of the time-lagged auto-covariance (3.33) relates a finite diffusivity to the product of a covariance and correlation time scale. If one casts this as a resonant process, the covariance will be infinitely small and the correlation time scale infinitely long, thus leading to an inconsistency between interaction and correlation time scales in the resonant limit. Introducing a broadened kinetic equation (2.2.1) with finite bandwidth (2.19) might naively be considered progress. However, that bandwidth represents a spike decay rate rather than a correlation time scale and is not a resolution. As one runs to the finite amplitude of the weakly nonlinear problem, the resonant bandwidth becomes the r.m.s. Doppler shift (Polzin & Lvov 2017). This is aphysical.

We find through numerical experimentation in Polzin & Lvov (2023) that the mean drift $\langle \dot{\mathbf{p}} \rangle$ is a resonant process and dispersion about the mean drift C_{ij} is non-resonant. The latter

should not come as a surprise once one appreciates the direct analogy between particle dispersion and ray tracing originally suggested in McComas & Bretherton (1977): particle dispersion in turbulence has nothing to do with the concept of resonance. The moments $\langle \dot{\mathbf{p}} \rangle$ and C_{ij} , and changes in the structure and scaling of the resonant bandwidth, are the signature differences of ray theory vs the kinetic equation, paralleling differences between Eulerian (Kraichnan 1959) and Lagrangian (Kraichnan 1965) representations of 3-D turbulence. These differences are rooted in the distinctions between amplitude-modulated and frequency-modulated signals.

4. Energy transport in oceanic internal waves

In Polzin & Lvov (2011, 2017) we note the tension between an apparent pattern match between observed spectral power laws being in apparent agreement with stationary states of the Fokker–Planck equation derived from the kinetic equation (2.16)

$$\frac{\partial n(\mathbf{p})}{\partial m} + \frac{\partial}{\partial m} D_{33} \frac{\partial}{\partial m} n(\mathbf{p}) = 0, \quad (4.1)$$

and this result being inconsistent with what is observationally understood about the energy sources and sinks. Those stationary states come from asserting a balance only in vertical wavenumber, for which there are two families: no-flux states for which $n(\mathbf{p}) \propto k^{-x} m^{-y}$ with $y = 0$ and constant-flux states for which a linear relationship between x and y is attained. Both families are oceanographically relevant Polzin & Lvov (2011) and, more to the point, the Garrett and Munk model (GM76) is a member of the no-flux family. This no-flux result is attained simply because that spectrum $((x, y) = (4, 0))$ has no gradients in action in vertical wavenumber.

The ray-path perspective (3.31) moves away from this interpretation so that the downscale transport contains an advection contribution of $\langle \dot{m} \rangle \langle n_p \rangle$ in addition to the diffusive transport $D_{33} \partial_m n_p$. An extensive oceanographic literature on passive tracer dispersion provides a basis for further insight. Here, it is understood that spatially inhomogeneous turbulence will result in a mean drift of either particles (Davis 1991) or tracer centre of mass (Armi 1979; Ledwell, Watson & Law 1998) at a rate proportional to the spatial gradient of the diffusivity. Here, we accept this identification, so that from (2.26) we have

$$\langle \dot{m} \rangle = \frac{\partial D_{33}}{\partial m} = \frac{4}{\pi} \frac{k m e_0 m_*}{N}. \quad (4.2)$$

Including a factor of two to account for the two-sided spectral representation of $e(\sigma, m)$, the downscale energy transport is

$$\mathcal{P} = 2 \int_f^N \langle \dot{m} \rangle e(m, \sigma) d\sigma = 2 \left(\frac{2}{\pi} \right)^2 \left(\frac{e_0 m_*}{N} \right)^2 f \log \left(\frac{N}{f} \right) \cong 1.0 \times 10^{-8} [\text{W kg}^{-1}], \quad (4.3)$$

which, apart from the prefactor of 1.0×10^{-8} being an order of magnitude too large, is virtually identical to the fine-scale parameterization; see Polzin *et al.* (2014), their (27) and (40). In Polzin & Lvov (2023) we demonstrate through a path integral closure that, indeed, $\langle \dot{m} \rangle = \partial_m D$ for the 1-D representation of high-frequency internal waves interacting with inertial waves. We further demonstrate that both are consistent with simple scale-invariant ray-tracing numerical simulations that treat the interaction as a 1-D problem.

We believe that this 1-D treatment is a reasonable representation of extreme scale-separated interactions. The 1-D version of (2.25), (4.1), dates to the dawn of modern oceanography and is supported by basic scale analysis (McComas & Bretherton 1977; Sun & Kunze 1999a). It is underpinned by the integrable singularity of the inertial peak in the internal wave frequency spectrum and the lack of horizontal velocity gradients in that peak that is encoded in the dispersion relation. While near-inertial motions are very energetic, they are also the most linear part of the internal wave field. In the absence of horizontal wave gradients, nonlinearity vanishes, and thus a superposition of inertial motions is an exact solution of the primitive equations describing the evolution of the large-scale background assumed in § 3.1.1. Such an inertial wave representation is used in Polzin & Lvov (2023) to numerically validate (2.26) and (4.2).

A modern analysis of local and extreme scale-separated interactions in a non-rotating context (Dematteis *et al.* 2022) assigns an energy transport associated with horizontal and diagonal terms of the diffusivity tensor that is an order of magnitude smaller than the fine-scale parameterization, two orders of magnitude smaller than (4.3). These non-vertical transports are likely overestimates in a rotating paradigm due to the vanishing of horizontal velocity gradients at the inertial peak.

Thus, our interpretation is that the advective transport arises in association with variations in the dispersion of wave packets in phase space in a manner that parallels the behaviour of passive tracers. The kinetic equation and ray tracing differ in that the trajectory of a wave packet in phase space returns significantly different information about the statistical inhomogeneity of phase space than is represented in the kinetic equation.

The result (4.3) stands in dramatic contrast with numerical results concerning the more general problem of high-frequency internal wave packets refracting in a background sea of internal waves (e.g. Henyey *et al.* 1986; Sun & Kunze 1999b; Ijichi & Hibiya 2017). This contention has required a systematic examination and physical interpretation of the assumptions within both the kinetic equation (§ 2.2) and ray-path (§ 3.2) approaches to pinpoint the multiple junctures which might underpin a systematic difference between observation and theory concerning extreme scale-separated interactions that are presented above in (4.3).

To date, we have identified three potential soft spots that could resolve the contradiction with general ray-tracing metrics.

The first is that both the ray tracing and kinetic equation discard a coupling between leading-order processes that leads to a subtractive cancellation of these leading orders. These leading orders are provided by the ID mechanism and the Bragg scattering mechanism, in which the phase locking introduced by ID is damped by Bragg scattering. Both are part of an extreme scale-separated Hamiltonian (3.10); Bragg scattering is discarded in a Taylor series expansion about the ID resonance (3.15) that produces the action conservation statement (3.16). Representation of this damping process can be defined by bundling eight distinct triads from the resonant manifold; the standard three-wave kinetic equation is a perturbation expansion in wave amplitude limited to one triad through a ‘random phase’ approximation to obtain (2.16). This coupling is pursued in the companion manuscript (Polzin & Lvov 2023), where we argue that these new physics are the route through which the ocean resolves the contradiction.

The second is that there is the potential for finite-amplitude effects in which interactions, both resonant and non-resonant, represent a stochastic forcing in phase space on a short time scale that disrupts the phase velocity–group velocity resonance on a long time scale. This requires assessment of the resonant bandwidth, which differs between Eulerian (kinetic equation, (2.18)) and ray coordinates, in combination with the amplitude and

decorrelation time scales of that forcing. This is investigated more fully in Polzin & Lvov (2023). Our opinion is that (4.3) and associated scaling is a fundamental metric that should be recoverable by kitchen sink efforts as a small-amplitude limit of wave turbulence and using a scale separation that aligns with the assumptions underpinning ray tracing. We offer the opinion that this is how the kitchen sink numerics of ray tracing resolves the contradiction.

The third is that extant efforts at ray tracing of the high-frequency internal wave packets refracting in a background sea of internal waves (Henyey *et al.* 1986; Sun & Kunze 1999b; Ijichi & Hibiya 2017) cannot be relied upon as a robust arbiter of this contradiction. Ray tracing is an asymptotic method requiring a scale separation in horizontal wavenumber (3.1.2) in addition to the spatial averaging implied in the envelope structure (3.2.1 Gershgorin *et al.* 2009). The ray-tracing numerical studies acknowledge none of this and regard scale separation as a tuneable parameter. They consistently document sensitivity to the specification of the scale separation and consistently find that the observed fine-scale metric of energy sourced to turbulent dissipation (Polzin *et al.* 2014) requires a scale equivalence, i.e. requires the small parameter of an asymptotic expansion to be $\sim O(1)$. This is the hallmark of interactions represented as $\mathcal{H}_{local}$ (3.7) that are spectrally local in wavenumber and need to be treated by other methods (Dematteis & Lvov 2021; Dematteis *et al.* 2022). The kitchen sink numerics will additionally suffer from the issue that the background is not a solution of the equations of motion.

5. Conclusions

We have presented two distinct derivations of transport equations for the refraction of high-frequency internal waves in inertial wave shear. One derivation results from ‘standard’ wave turbulence techniques with the addition of near-resonant interactions and describes the wave field as a system of amplitude-modulated waves. This kinetic equation-based derivation results in a Fokker–Planck equation which returns an estimate of no net downscale transport in vertical wavenumber for the canonical spectrum of oceanic internal waves referred to as the Garrett–Munk spectrum and small transports associated with horizontal and off-diagonal elements (Dematteis *et al.* 2022). The second derivation is based on ray-tracing techniques in the WKB limit. Here the ensemble-averaged transport equation (3.31) contains a mean-drift term that is absent from the Fokker–Planck equation derived from the kinetic equation (2.25). This term leads to a prediction of ocean dissipation (4.3) an order of magnitude greater than supported by ocean observations.

The disparity of these results can be interpreted from an analogy between ray characteristics and Lagrangian paths. Recognizing this distinction improves earlier derivations of the ray-path transport equation in McComas & Bretherton (1977) and Nazarenko *et al.* (2001): it provides a rigorous basis for the intuitive characterization of the transport of energy to dissipation invoked in Henyey *et al.* (1986), Sun & Kunze (1999b) and Ijichi & Hibiya (2017). We arrived at this rigorous result by adding zero and explicitly invoking an ensemble average.

Holloway (1980, 1982) argue that internal wave interactions might not be sufficiently weak for wave turbulence theory to be valid. That commentary is directed at inferences about the decay rate of narrowband perturbations to the spectrum being much larger than the internal wave frequency (Müller *et al.* 1986). This is inconsistent with the express intent that the kinetic equation describes the slow evolution of the wave spectrum. In Polzin & Lvov (2017) we identify this decay rate as the small-amplitude limit of the resonant bandwidth Γ (2.18). However, the differences in our two Fokker–Planck equations do not hinge upon this issue. We have demonstrated in Polzin & Lvov (2017)

that, at finite amplitude, the bandwidth is proportional to the r.m.s. Doppler shift. This denotes a degenerate state in which the bandwidth describes the quasi-coherent translation of small-scale waves, i.e. sweeping, rather than their interaction. In the WKB-based derivation, we operate upon the Hamiltonian with a Wigner transform that integrates over that narrow band perturbation and its nearly resonant decay partners in an extreme scale-separated limit. This removes the apparent discrepancy and arrives at different notions of bandwidth and the role of off-resonant interactions (Polzin & Lvov 2017). We investigate these issues in greater detail in Polzin & Lvov (2023).

As we look back over the landscape of this endeavour, what we have is a well-established metric for ocean mixing known as fine-scale parameterization (Polzin *et al.* 2014). At best, the fine-scale parameterization is underpinned by a heuristic description as an advective spectral closure (Polzin 2004) in the context of an energy transport equation that eschews action conservation in which energy transport in horizontal wavenumber keeps pace with that in vertical wavenumber. This interpretation contrasts with the pivotal role that ID was perceived to play in determining downscale transports in vertical wavenumber only. A possible resolution can be found in recent characterizations of the internal wave kinetic equation (Dematteis & Lvov 2021; Dematteis *et al.* 2022) that coincide in magnitude and scaling with the description of transports encapsulated within the fine-scale parameterization. That work emphasizes the importance of local interactions.

The heart of our manuscript is a novel derivation that is an inspired addition of zero to an action balance that permits us to deal with inhomogeneities of dispersive tendencies in phase space. These inhomogeneities give rise to the leading-order transport terms in our system. Adding zero is simplistic, but yet is grounded in deep physical insight into this problem. Our work exposes a major lack of conceptual understanding of extreme scale-separated interactions in the existing literature.

To conclude, here we have derived a transport equation (3.29) based upon a packet ensemble that contains an advective transport term. When the advective transport term is evaluated in a rotating context, it gives rise to a transport estimate (4.3) that is an order of magnitude greater than the fine-scale parameterization. This inconsistency will be analysed in Polzin & Lvov (2023), where we suggest that the fourth-order cumulants, which are subleading to the product of two two-point correlators of inhomogeneous wave turbulence, are playing a significant role in the spatially inhomogeneous ray coordinate system.

Funding. Y.L. gratefully acknowledges the support of the National Science Foundation Division of Mathematical Sciences award 2009418. K.P. gratefully acknowledges support from the Office of Naval Research (Grant N000141712852) and Woods Hole Oceanographic Institution's Investment in Science Program for this project.

Declaration of interests. The authors report no conflict of interest.

Author ORCIDs.

 Yuri V. Lvov <https://orcid.org/0000-0003-2153-4047>;

 Kurt L. Polzin <https://orcid.org/0000-0002-9991-1941>.

Appendix A. The Garrett and Munk spectrum

We utilize what is referred to as GM76, the 1976 version of the Garrett and Munk model. We refer the reader to Garrett & Munk (1972) and Garrett & Munk (1979) for historical perspectives and to Müller *et al.* (1986); Polzin & Lvov (2011) for reviews. The Garrett and Munk models treat the energy spectrum $e(\sigma, m)$ as separable in vertical wavenumber m and frequency σ with the potential for variable high-wavenumber/high-frequency power-law

regimes and vertical wavenumber bandwidth. The two-sided Garrett and Munk 1976 model is

$$e(\sigma, m) = \frac{1}{2} e_0 \frac{2m_*}{\pi} \frac{1}{m_*^2 + m^2} \frac{2f}{\pi} \frac{1}{\sigma(\sigma^2 - f^2)^{1/2}}, \quad (\text{A1})$$

in which vertical wavenumber m can be both positive and negative, frequency σ is positive, f is the Coriolis parameter and $m_* = 4\pi/b$ with thermocline scale height $b = 1300$ m the vertical wavenumber bandwidth. The vertical wavenumber bandwidth m_* implicitly scales with buoyancy as $m_* \rightarrow m_* N/N_0$, in which N is the buoyancy frequency and $N_0 = 3$ cph is a thermocline reference value. The total energy e_0 is obtained by integrating over vertical wavenumber and frequency and then summing both sides of the vertical wavenumber spectrum. One uses linear kinematics and the dispersion relation to convert from energy density to wave-action density and spectral density in frequency to spectral density in horizontal wavenumber.

Appendix B. Rotations included

We now repeat all the calculations with rotations included. We start from the primitive equations (2.1), decompose the velocity using (2.3) and then use the expression for potential vorticity (2.5) to obtain

$$\left. \begin{aligned} \Pi_t + \nabla \cdot \left(\Pi \left(\nabla \phi + \nabla^\perp \Delta^{-1} \left(\frac{f}{\Pi_0} \Pi - f \right) \right) \right) &= 0, \\ \phi_t + \frac{1}{2} \left| \nabla \phi + \nabla^\perp \Delta^{-1} \left(\frac{f}{\Pi_0} \Pi - f \right) \right|^2 \\ &+ \Delta^{-1} \nabla \cdot \left[\frac{f}{\Pi_0} \Pi \left(\nabla^\perp \phi - \nabla \Delta^{-1} \left(\frac{f}{\Pi_0} \Pi - f \right) \right) \right] \\ &+ \frac{g}{\rho} \int^\rho \int^{\rho_2} \frac{(\Pi - \Pi_0)}{\rho_1} d\rho_1 d\rho_2 = 0. \end{aligned} \right\} \quad (\text{B1})$$

We now substitute the Reynolds decomposition (3.1a–c) into (B1). In doing so, we use the fact that potential vorticity is assumed to be conserved on isopycnals

$$\frac{f}{\Pi_0} = \frac{f + \Delta \Psi + \Delta \psi'}{\Pi_0 + \Pi + \pi'}. \quad (\text{B2})$$

Here, we denote by $\Psi + \psi'$ the divergence-free part of the velocity field

$$\Psi + \psi' = \Delta^{-1} \left(\frac{f}{\Pi_0} (\Pi_0 + \Pi + \pi') \right). \quad (\text{B3})$$

Then

$$\left. \begin{aligned} & \dot{\pi}' + \nabla \cdot \left((\Pi_0 + \Pi + \pi') \left(\nabla \phi' + \nabla^\perp \left(\Delta^{-1} \left(\frac{f\pi'}{\Pi_0} \right) \right) \right) \right) \\ & + \nabla \cdot \left(\pi' \left(\nabla \Phi + \nabla^\perp \left(\Delta^{-1} \left(\frac{f\Pi}{\Pi_0} \right) \right) \right) \right) = 0, \\ & \dot{\phi}' + \frac{1}{2} \left| \nabla \phi' + \nabla^\perp \left(\Delta^{-1} \left(\frac{f\pi'}{\Pi_0} \right) \right) \right|^2 \\ & + \Delta^{-1} \nabla \cdot \left(\frac{f}{\Pi_0} (\Pi_0 + \Pi + \pi') \left(\nabla^\perp \phi' - \nabla \left(\Delta^{-1} \left(\frac{f\pi'}{\Pi_0} \right) \right) \right) \right) \\ & + \Delta^{-1} \cdot \nabla \left(\frac{f\pi'}{\Pi_0} \left(\nabla^\perp \Phi - \nabla \left(\Delta^{-1} \left(\frac{f\Pi}{\Pi_0} \right) \right) \right) \right) \\ & + \left(\nabla \phi' + \nabla^\perp \left(\Delta^{-1} \left(\frac{f\pi'}{\Pi_0} \right) \right) \right) \cdot \left(\nabla \Phi + \nabla^\perp \left(\Delta^{-1} \left(\frac{f\Pi}{\Pi_0} \right) \right) \right) \\ & + \frac{g}{\rho} \int^\rho \int^{\rho_2} \frac{\pi'}{\rho} d\rho_1 d\rho_2 = 0. \end{aligned} \right\} \quad (\text{B4})$$

Remarkably, these equations are indeed Hamiltonian, with the Hamiltonian given by

$$\begin{aligned} \mathcal{H} = & \frac{1}{2} \int \left[(\Pi_0 + \Pi + \pi') \left| \nabla \phi' + \nabla^\perp \Delta^{-1} \left(\frac{f\pi'}{\Pi_0} \right) \right|^2 - g \left| \int^\rho \frac{\pi'}{\rho_1} d\rho_1 \right|^2 \right. \\ & \left. + 2\pi' \left(\nabla \phi' + \nabla^\perp \Delta^{-1} \left(\frac{f\pi'}{\Pi_0} \right) \right) \cdot \left(\nabla \Phi + \nabla^\perp \Delta^{-1} \left(\frac{f\Pi}{\Pi_0} \right) \right) \right] d\mathbf{r}. \end{aligned} \quad (\text{B5})$$

Please compare this with (3.5).

We now can rewrite the Hamiltonian (B5) in the following form:

$$\left. \begin{aligned} \mathcal{H} &= \mathcal{H}_{\text{linear}} + \mathcal{H}_{\text{nonlinear}}, \\ \mathcal{H}_{\text{linear}} &= \frac{1}{2} \int \left[\Pi_0 \left| \nabla \phi' + \nabla^\perp \Delta^{-1} \left(\frac{f\pi'}{\Pi_0} \right) \right|^2 - g \left| \int^\rho \frac{\pi'}{\rho} d\rho_1 \right|^2 \right] d\mathbf{r}, \\ \mathcal{H}_{\text{nonlinear}} &= \mathcal{H}_{\text{local}} + \mathcal{H}_{\text{sweeping}} + \mathcal{H}_{\text{density}}, \\ \mathcal{H}_{\text{local}} &= \frac{1}{2} \int \pi' \left| \nabla \phi' + \nabla^\perp \left(\Delta^{-1} \frac{f\pi'}{\Pi_0} \right) \right|^2 d\mathbf{r}, \\ \mathcal{H}_{\text{density}} &= \frac{1}{2} \int \Pi \left| \nabla \phi' + \nabla^\perp \left(\Delta^{-1} \frac{f\pi'}{\Pi_0} \right) \right|^2 d\mathbf{r}, \\ \mathcal{H}_{\text{sweeping}} &= \int \pi' \left(\nabla \phi' + \nabla^\perp \left(\Delta^{-1} \frac{f\pi'}{\Pi_0} \right) \right) \cdot \left(\nabla \Phi + \nabla^\perp \left(\Delta^{-1} \frac{f\Pi}{\Pi_0} \right) \right) d\mathbf{r}. \end{aligned} \right\} \quad (\text{B6})$$

Making the Fourier transformation and making the Boussinesq approximation allows us to rewrite this in the form that generalizes (3.7)

$$\begin{aligned}
 \mathcal{H} &= \mathcal{H}_{\text{linear}} + \mathcal{H}_{\text{nonlinear}}, \\
 \mathcal{H}_{\text{linear}} &= \frac{1}{2} \int d\mathbf{p} \left(\Pi_0 k^2 |\phi'_{\mathbf{p}}|^2 + \left(\frac{f^2}{k^2 \Pi_0} - \frac{g}{\rho_0^2 m^2} \right) |\Pi_{\mathbf{p}}|^2 \right), \\
 \mathcal{H}_{\text{nonlinear}} &= \mathcal{H}_{\text{local}} + \mathcal{H}_{\text{sweeping}} + \mathcal{H}_{\text{density}}, \\
 \mathcal{H}_{\text{local}} &= \frac{1}{2} \frac{1}{(2\pi)^{3/2}} \int d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{p}_3 \delta(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \\
 &\quad \times \left(-\mathbf{k}_2 \cdot \mathbf{k}_3 \pi'_{\mathbf{p}_1} \phi'_{\mathbf{p}_2} \phi'_{\mathbf{p}_3} - \left(\frac{f}{\Pi_0} \right)^2 \frac{\mathbf{k}_2 \cdot \mathbf{k}_3}{k_2^2 k_3^2} \pi'_{\mathbf{p}_1} \pi'_{\mathbf{p}_2} \pi'_{\mathbf{p}_3} - 2 \frac{f}{\Pi_0} \frac{\mathbf{k}_2 \cdot \mathbf{k}_3^\perp}{k_3^2} \pi'_{\mathbf{p}_1} \phi'_{\mathbf{p}_2} \pi'_{\mathbf{p}_3} \right), \\
 \mathcal{H}_{\text{density}} &= \frac{1}{2} \frac{1}{(2\pi)^{3/2}} \int d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{p}_3 \delta(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \left(-\mathbf{k}_2 \cdot \mathbf{k}_3 \Pi_{\mathbf{p}_1} \phi'_{\mathbf{p}_2} \phi'_{\mathbf{p}_3} \right. \\
 &\quad \left. - \left(\frac{f}{\Pi_0} \right)^2 \frac{\mathbf{k}_2 \cdot \mathbf{k}_3}{k_2^2 k_3^2} \Pi_{\mathbf{p}_1} \pi'_{\mathbf{p}_2} \pi'_{\mathbf{p}_3} + 2 \frac{f}{\Pi_0} \frac{\mathbf{k}_2 \cdot \mathbf{k}_3^\perp}{k_3^2} \Pi_{\mathbf{p}_1} \phi'_{\mathbf{p}_2} \pi'_{\mathbf{p}_3} \right), \\
 \mathcal{H}_{\text{sweeping}} &= \frac{1}{(2\pi)^{3/2}} \int d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{p}_3 \delta(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \times \left(-\mathbf{k}_2 \cdot \mathbf{k}_3 \pi'_{\mathbf{p}_1} \phi'_{\mathbf{p}_2} \Phi_{\mathbf{p}_3} \right. \\
 &\quad \left. + \frac{f}{\Pi_0} \frac{\mathbf{k}_2 \cdot \mathbf{k}_3^\perp}{k_3^2} \pi'_{\mathbf{p}_1} \phi'_{\mathbf{p}_2} \Pi_{\mathbf{p}_3} + \frac{f}{\Pi_0} \frac{\mathbf{k}_2^\perp \cdot \mathbf{k}_3}{k_2^2} \pi'_{\mathbf{p}_1} \pi'_{\mathbf{p}_2} \Phi_{\mathbf{p}_3} - \left(\frac{f}{\Pi_0} \right)^2 \frac{\mathbf{k}_2 \cdot \mathbf{k}_3}{k_2^2 k_3^2} \pi'_{\mathbf{p}_1} \pi'_{\mathbf{p}_2} \Pi_{\mathbf{p}_3} \right).
 \end{aligned} \tag{B7}$$

The next step is to substitute formulas for Π_q and Φ_q from (2.11a,b), and to perform the steps and approximations that are developed in §§ 3.1, 3.2. Results are the generalization of (3.12) and (3.10) and (3.11)

$$\begin{aligned}
 \mathcal{H} &= \int d\mathbf{p}_1 d\mathbf{p}_2 A^f(\mathbf{p}_1, \mathbf{p}_2) a_{\mathbf{p}_1} a_{\mathbf{p}_2}^*, \quad \text{with} \\
 A^f(\mathbf{p}_1, \mathbf{p}_2) &= \omega_{\mathbf{p}_1} \delta(\mathbf{p}_1 - \mathbf{p}_2) + A_{\text{sweeping}}^f(\mathbf{p}_1, \mathbf{p}_2) + A_{\text{density}}^f(\mathbf{p}_1, \mathbf{p}_2), \\
 A_{\text{sweeping}}^f(\mathbf{p}_1, \mathbf{p}_2) &= \frac{1}{(2\pi)^{3/2}} \left(\frac{1}{2} i(\mathbf{k}_1 - \mathbf{k}_2) \cdot (\mathbf{k}_1 + \mathbf{k}_2) \right. \\
 &\quad \left. + \frac{f}{\Pi_0} \mathbf{k}_1 \cdot \mathbf{k}_2^\perp \left(\frac{1}{k_2^2} - \frac{1}{k_1^2} \right) \frac{g k_1 k_2}{2 \sqrt{\omega_1 \omega_2} N^2} \right) \Phi_{\mathbf{p}_2 - \mathbf{p}_1} \\
 &\quad + \left(\frac{-i \frac{f}{\Pi_0} \mathbf{k}_1^\perp \cdot \mathbf{k}_2}{|\mathbf{k}_2 - \mathbf{k}_1|^2} - \left(\frac{f}{\Pi_0} \right)^2 \frac{(\mathbf{k}_2 - \mathbf{k}_1)}{(\mathbf{k}_1 - \mathbf{k}_2)^2} \cdot \left(\frac{\mathbf{k}_2}{k_2^2} + \frac{\mathbf{k}_1}{k_1^2} \right) \frac{g k_1 k_2}{2 \sqrt{\omega_1 \omega_2} N^2} \right) \Pi_{\mathbf{p}_2 - \mathbf{p}_1}, \\
 A_{\text{density}}^f(\mathbf{p}_1, \mathbf{p}_2) &= \frac{1}{(2\pi)^{3/2}} \left(\frac{N^2}{2 g \sqrt{\omega_{\mathbf{p}_1} \omega_{\mathbf{p}_2}}} \frac{\mathbf{k}_2 \cdot \mathbf{k}_3}{k_2 k_3} + \frac{i}{2 N^2} \frac{g}{\Pi_0^2 \sqrt{\omega_{\mathbf{p}_1} \omega_{\mathbf{p}_2}}} \frac{f^2}{\sqrt{\omega_{\mathbf{p}_1} \omega_{\mathbf{p}_2}}} \right) \Pi_{\mathbf{p}_1 - \mathbf{p}_2}.
 \end{aligned} \tag{B8}$$

Using these equations and repeating steps above and discarding terms proportional to q^2 leads to

$$\begin{aligned}\omega(\mathbf{p}, \mathbf{r}) = & \sigma_p - \mathbf{k} \cdot \nabla \Phi(\mathbf{r}, t) - \frac{f}{\Pi_0} \mathbf{k} \cdot \nabla^\perp \Delta^{-1} \Pi(\mathbf{r}, t) \\ & + \frac{N^2}{2g} \Pi(\mathbf{r}, t) \sigma_p + \frac{f^2 g}{\Pi_0^2 \omega_p N^2} \Pi(\mathbf{r}, t) \\ & + \frac{f^2 g}{2\omega_p \Pi_0^2 N^2} \int \cos(2\theta_{qp}) \Pi(\mathbf{q}, t) e^{i\mathbf{r} \cdot \mathbf{q}} d\mathbf{q}.\end{aligned}\quad (\text{B9})$$

Here, θ_{qp} is an angle between the horizontal part of wave vector $\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_2$ (3.13) and the horizontal part of $\mathbf{p}$, that is $\mathbf{k}$, with the sign defined using the right-hand rule going from $\mathbf{q}$ TO $\mathbf{p}$. The Eulerian frequency σ_p is given by (2.12). Note that the first line is simply $\sigma_p - \mathbf{k} \cdot \mathcal{U}$, the second line is the density term that we have considered above and the third line is the contribution from a rotating ocean. Equation (3.18), which is lines one and two of (B9) with $\mathcal{U} = \nabla \Phi$, applies to a non-rotating ocean. In a rotating ocean, a wave whose frequency is high enough not to be impacted by rotation is described using lines one and two.

REFERENCES

- ARBIC, B.K., *et al.* 2018 A primer on global internal tide and internal gravity wave continuum modeling in HYCOM and MITgcm. In *New Frontiers in Operational Oceanography* (ed. E.P. Chassignet, A. Pascual, J. Tintoré & J. Verron), chap. 13. GODAE OceanView.
- ARMI, L. 1979 Effects of variations in eddy diffusivity on property distributions in the oceans. *J. Mar. Res.* **37** (3), 515–530.
- BENNETT, A.F. 1984 Relative dispersion: local and nonlocal dynamics. *J. Atmos. Sci.* **41**, 1881–1886.
- BÜHLER, O. & MCINTYRE, M.E. 2005 Wave capture and wave-vortex duality. *J. Fluid Mech.* **534**, 67–95.
- COHEN, L. & LEE, C. 1990 Instantaneous bandwidth for signals and spectrogram. In *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing*, pp. 2451–2454.
- DAVIS, R. 1991 Observing the general circulation with floats. *Deep-Sea Res.* **38**, S531–S571.
- DEMATTEIS, G. & LVOV, Y. 2021 Downscale energy fluxes in scale invariant oceanic internal wave turbulence. *J. Fluid Mech.* **915**, A129.
- DEMATTEIS, G., POLZIN, K.L. & LVOV, Y.V. 2022 On the origins of the oceanic ultraviolet catastrophe. *J. Phys. Oceanogr.* **51**, 597–616.
- DENG, Y. & HANI, Z. 2023 Full derivation of the kinetic equation. *Invent. Math.* **233**, 543–724.
- DONG, W., BÜHLER, O. & SMITH, K.S. 2020 Frequency diffusion of waves by unsteady flows. *J. Fluid Mech.* **905**, R3.
- DOSSER, H. & SUTHERLAND, B. 2011 Anelastic internal wave packet evolution and stability. *J. Atmos. Sci.* **68**, 2844–2859.
- GARRETT, C.J.R. & MUNK, W.H. 1972 Space-time scales of internal waves. *Geophys. Fluid Dyn.* **3**, 225–264.
- GARRETT, C.J.R. & MUNK, W.H. 1979 Internal waves in the ocean. *Annu. Rev. Fluid Mech.* **11**, 339–369.
- GERSHGORIN, B., LVOV, Y.V. & NAZARENKO, S. 2009 Canonical Hamiltonians for waves in inhomogeneous media. *J. Math. Phys.* **50**, 013527.
- GREGG, M.C. 1989 Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.* **94**, 9686–9698.
- HAYNES, P.H. & MCINTYRE, M.E. 1987 On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces. *J. Atmos. Sci.* **44**, 828–841.
- HENYEY, F.S., WRIGHT, J. & FLATTÉ, S.M. 1986 Energy and action flow through the internal wave field. An eikonal approach. *J. Geophys. Res.* **91**, 8487–8495.
- HENYEY, F.S. & POMPHREY, N. 1983 Eikonal description of internal wave interactions: a non-diffusive picture of “induced diffusion”. *Dyn. Atmos. Oceans* **7**, 189–208.
- HENYEY, F.S., POMPHREY, N. & MEISS, J.D. 1984 Comparison of short-wavelength internal wave transport theories. LJI-R-84-248, La Jolla Institute.
- HÖGSTRÖM, U.L.F. 1996 Review of some basic characteristics of the atmospheric surface layer. *Boundary-Layer Meteorol.* **78**, 215–246.

- HOLLOWAY, G. 1980 Oceanic internal waves are not weak waves. *J. Phys. Oceanogr.* **10**, 906–914.
- HOLLOWAY, G. 1982 On interaction timescales of oceanic internal waves. *J. Phys. Oceanogr.* **12**, 293–296.
- HOLLOWAY, G. & HENDERSHOTT, M.C. 1977 Stochastic closure for nonlinear Rossby waves. *J. Fluid Mech.* **82**, 747–765.
- IJICHI, T. & HIBIYA, T. 2017 Eikonal calculations for energy transfer in the deep-ocean internal wave field near mixing hotspots. *J. Phys. Oceanogr.* **47**, 199–210.
- KAFIABAD, H.A., SAVVA, A., MILES, A.C. & VANNESTE, J. 2019 Diffusion of inertia-gravity waves by geostrophic turbulence. *J. Fluid Mech.* **869**, R7.
- KRAICHNAN, R.H. 1959 The structure of isotropic turbulence at very high Reynolds numbers. *J. Fluid Mech.* **5**, 497–543.
- KRAICHNAN, R.H. 1965 Lagrangian-history closure approximation for turbulence. *Phys. Fluids* **8**, 575–598.
- LANCHON, N. & CORTET, P.-P. 2023 Energy spectra of non-local internal gravity wave turbulence. *Phys. Rev. Lett.* **131**, 264001.
- LEDWELL, J.R., WATSON, A.J. & LAW, C.S. 1998 Mixing of a tracer in the pycnocline. *J. Geophys. Res.: Oceans* **103**, 21499–21529.
- LVOV, V.S., LVOV, Y.V., NEWELL, A.C. & ZAKHAROV, V.E. 1997 Statistical description of acoustic turbulence. *Phys. Rev. E* **56**, 390–405.
- LVOV, Y.V. & TABAK, E.G. 2001 Hamiltonian formalism and the Garrett and Munk spectrum of internal waves in the ocean. *Phys. Rev. Lett.* **87**, 169501.
- LVOV, Y.V. & TABAK, E.G. 2004 A Hamiltonian formulation for long internal waves. *Physica D* **195**, 106–122.
- LVOV, Y.V., POLZIN, K.L., TABAK, E.G. & YOKOYAMA, N. 2010 Oceanic internal wavefield: theory of scale-invariant spectra. *J. Phys. Oceanogr.* **40**, 2605–2623.
- LVOV, Y.V., POLZIN, K.L. & YOKOYAMA, N. 2012 Resonant and near-resonant internal wave interactions. *J. Phys. Oceanogr.* **42**, 669–691.
- LVOV, V.S. & RUBENCHIK, A.M. 1977 Spatially non-uniform singular weak turbulence spectra. *Sov. Phys. JETP* **45**, 67. Also at V.S. Lvov, Fundamentals of Nonlinear Physics, Lecture Notes, Eqs. (7.7)–(7.8), www.lvov.weizmann.ac.il.
- MCCOMAS, C.H. 1977 Equilibrium mechanisms within the oceanic internal wavefield. *J. Phys. Oceanogr.* **7**, 836–845.
- MCCOMAS, C.H. & BRETHERTON, F.P. 1977 Resonant interaction of oceanic internal waves. *J. Geophys. Res.* **83**, 1397–1412.
- MCCOMAS, C.H. & MÜLLER, P. 1981a Timescales of resonant interactions among oceanic internal waves. *J. Phys. Oceanogr.* **11**, 139–147.
- MCCOMAS, C.H. & MÜLLER, P. 1981b The dynamic balance of internal waves. *J. Phys. Oceanogr.* **11**, 970–986.
- MÜLLER, P. 1976 On the diffusion of momentum and mass by internal gravity waves. *J. Fluid Mech.* **77**, 789–823.
- MÜLLER, P., HOLLOWAY, G., HENYEY, F. & POMPHREY, N. 1986 Nonlinear interactions among internal gravity waves. *Rev. Geophys.* **24**, 493–536.
- NAZARENKO, S.V., NEWELL, A.C. & GALTIER, S.V. 2001 Nonlocal MHD turbulence. *Physica D* **152–153**, 646–652.
- NAZARENKO, S. 2011 *Wave Turbulence*. Springer Science & Business Media.
- NEWELL, A.C. 1968 The closure problem in a system of random gravity waves. *Rev. Geophys.* **6**, 1–31.
- PAN, Y., ARBIC, B.K., NELSON, A.D., MENEMENLIS, D., PELTIER, W.R., XU, W. & LI, Y. 2020 Numerical investigation of mechanisms underlying oceanic internal gravity wave power-law spectra. *J. Phys. Oceanogr.* **50**, 2713–2733.
- POLZIN, K.L. 2004 A heuristic description of internal wave dynamics. *J. Phys. Oceanogr.* **34** (1), 214–230.
- POLZIN, K.L., TOOLE, J.M. & SCHMITT, R.W. 1995 Finescale parameterizations of turbulent dissipation. *J. Phys. Oceanogr.* **25**, 306–328.
- POLZIN, K.L. 2010 Mesoscale-eddy internal wave coupling. II. Energetics and results from polymode. *J. Phys. Oceanogr.* **40**, 789–801.
- POLZIN, K.L. & LVOV, Y.S. 2011 Toward regional characterizations of the oceanic internal wave spectrum. *Rev. Geophys.* **49**, RG4003.
- POLZIN, K.L., NAVEIRA GARABATO, A.C., SLOYAN, B.M., HUUSSEN, T. & WATERMAN, S.N. 2014 Finescale parameterizations of turbulent dissipation. *J. Geophys. Res.* **119**, 1383–1419.
- POLZIN, K.L. & LVOV, Y.S. 2017 An oceanic ultra-violet catastrophe, wave-particle duality and a strongly nonlinear concept for geophysical turbulence. *Fluids* **2**, 36.

- POLZIN, K. & LVOV, Y. 2023 Scale separated interactions of oceanic internal waves: a one-dimensional model. *J. Phys. Oceanogr.* (submitted). [arXiv:2204.00069](https://arxiv.org/abs/2204.00069).
- POMPHREY, N., MEISS, J.D. & WATSON, K.M. 1980 Description of nonlinear internal wave interactions using Langevin methods. *J. Geophys. Res.* **85**, 1085–1094.
- SUN, H. & KUNZE, E. 1999*a* Internal wave–wave interactions. Part I: the role of internal wave vertical divergence. *J. Phys. Oceanogr.* **29**, 2886–2904.
- SUN, H. & KUNZE, E. 1999*b* Internal wave–wave interactions. Part II: spectral energy transfer and turbulence production. *J. Phys. Oceanogr.* **29**, 2905–2919.
- TAYLOR, G.I. 1921 Diffusion by continuous movements.
- VILLAS BÔAS, A.B. & YOUNG, W.R. 2020 Directional diffusion of surface gravity wave action by ocean macroturbulence. *J. Fluid Mech.* **890**, R3.
- WITHAM, G.B. 1974 *Linear and Nonlinear Waves*, p. 636. Wiley-Interscience.
- YOUNG, W.R. & JELLOUL, M.B. 1997 Propagation of near-inertial oscillations through a geostrophic flow. *J. Mar. Res.* **55**, 735–766.
- ZAKHAROV, V.E., LVOV, V.S. & FALKOVICH, G. 1992 *Kolmogorov Spectra of Turbulence*. Springer.