JM PAPERS

Regimes of near-inertial wave dynamics

Scott Conn[1,†]**, Jörn Callies**[1] **and Albion Lawrence**[1,2]

[1]California Institute of Technology, Pasadena, CA 91125, USA

[2]Brandeis University, Waltham, MA 02453, USA

When atmospheric storms pass over the ocean, they resonantly force near-inertial waves (NIWs), internal waves with a frequency close to the local Coriolis frequency f. It has long been recognised that the evolution of NIWs is modulated by the ocean's mesoscale eddy field. This can result in NIWs being concentrated into anticyclones which provide an efficient pathway for NIW propagation to depth. Here we analyse the eigenmodes of NIWs in the presence of mesoscale eddies and heavily draw on parallels with quantum mechanics. Whether the eddies are effective at modulating the behaviour of NIWs depends on the wave dispersiveness $\varepsilon^2 = f\lambda^2/\Psi$, where λ is the deformation radius and Ψ is a scaling for the eddy streamfunction. If $\varepsilon \gg 1$, NIWs are strongly dispersive, and the waves are only weakly affected by the eddies. We calculate the perturbations away from a uniform wave field and the frequency shift away from f. If $\varepsilon \ll 1$, NIWs are weakly dispersive, and the wave evolution is strongly modulated by the eddy field. In this weakly dispersive limit, the Wentzel–Kramers–Brillouin approximation, from which ray tracing emerges, is a valid description of the NIW evolution even if the large-scale atmospheric forcing apparently violates the requisite assumption of a scale separation between the waves and the eddies. The large-scale forcing excites many wave modes, each of which varies on a short spatial scale and is amenable to asymptotic analysis analogous to the semi-classical analysis of quantum systems. The strong modulation of weakly dispersive NIWs by eddies has the potential to modulate the energy input into NIWs from the wind, but we find that this effect should be small under oceanic conditions.

Key words: internal waves, ocean processes, quasi-geostrophic flows

† Email address for correspondence: sconn@caltech.edu

1. Introduction

Near-inertial waves (NIWs) play an important role in the global climate system. Being associated with strong vertical shears, they are prone to shear instabilities, which are an important driver of upper ocean mixing (for a review, see Alford *et al.* 2016). As such, the generation of NIWs is one of the primary mechanisms by which atmospheric storms induce a deepening of the surface mixed layer. This deepening requires mixing with water from below, implicating NIWs in the surface ocean heat budget (Jochum *et al.* 2013). In the interior of the ocean, NIWs make up a major fraction of the internal wave kinetic energy (Ferrari & Wunsch 2009; Alford *et al.* 2016), and it has been hypothesised that NIW kinetic energy may provide a source of mixing in the deep ocean (Munk & Wunsch 1998). Near-inertial waves might also extract energy from mesoscale eddies (Xie & Vanneste 2015; Rocha, Wagner & Young 2018) and, hence, play a role in the mesoscale energy budget.

In-situ observations of NIWs usually lack significant spatial resolution. The spatial structure of NIWs can generally only be resolved through dedicated field campaigns, for example, the Ocean Storms experiment (D'Asaro 1985) or the NISKINe field campaign (Voet *et al.* 2024). Despite this, it has become clear that NIW evolution can be strongly modulated by the presence of mesoscale eddies (e.g. Thomas *et al.* 2020; Conn, Fitzgerald & Callies 2024). Given the sparsity of NIW observations, theoretical progress has been important in understanding the dynamics of NIWs in the upper ocean.

Early work on NIW–eddy interactions was based on ray tracing theory. Kunze (1985) derived a dispersion relation for NIWs in the presence of a geostrophic background flow. Throughout this paper, we make the assumption of a barotropic (depth-independent) background flow. The ray tracing equations for a single (flat-bottom) baroclinic mode propagating through such a background flow are

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau} = \frac{\partial\omega}{\partial\boldsymbol{k}}, \quad \frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}\tau} = -\frac{\partial\omega}{\partial\boldsymbol{x}}, \quad \omega = \frac{f\lambda^2|\boldsymbol{k}|^2}{2} + \boldsymbol{u}\cdot\boldsymbol{k} + \frac{\zeta}{2}, \qquad (1.1a\text{--}c)$$

where $\boldsymbol{x} = (x, y)$ is the ray position, τ is time, $\boldsymbol{k}$ is the horizontal wave vector, $\boldsymbol{u}$ is the background velocity, $\zeta = \partial_x v - \partial_y u$ is the background vorticity and λ is the deformation radius. Here, and throughout the rest of this paper, ω refers to the frequency shift of a NIW away from the local inertial frequency f such that the true frequency is $f + \omega$. Based on these equations, Kunze (1985) argued that NIWs would be trapped in regions of anticyclonic vorticity where the effective frequency is less than the local f. This trapping arises from the refraction of rays by the background vorticity, i.e. from changes in the wavenumber vector due to spatial gradients of the $\zeta/2$ term in the dispersion relation. Concentration of NIW energy into anticyclones has indeed been observed in the ocean (e.g. Perkins 1976; Kunze & Sanford 1984; Thomas *et al.* 2020; Yu *et al.* 2022).

Ray tracing is based on the assumption that the NIWs are propagating through a slowly varying medium. This means that the horizontal scale of the waves has to be much smaller than the scale of the background mesoscale eddy field. Young & Ben Jelloul (1997, from hereon YBJ) criticised this spatial scale assumption based on the argument that NIWs are forced by large-scale storms and so, at least initially, the waves have a much larger scale than mesoscale eddies. As a remedy, YBJ developed a theory of NIW–eddy interactions that does not rely on the assumption of a spatial scale separation. This was also partly motivated by a desire to explain observations from the Ocean Storms experiment (D'Asaro *et al.* 1995), a field campaign that studied the evolution of NIWs in the wake of a large storm in the North Pacific. A key result of this campaign was that the effect of the

mesoscale vorticity on the wave evolution was in clear contradiction with predictions from ray tracing (D'Asaro 1995).

The YBJ equation describes the evolution of NIWs in the presence of a prescribed geostrophic eddy field while only assuming a temporal scale separation between the inertial period and the characteristic time scale of the eddies. For the barotropic background flow considered throughout this paper, the wave evolution can be split into baroclinic modes that do not interact, so we consider a single baroclinic mode with NIW velocity $[u_w(x, y, t), v_w(x, y, t)]g(z)$, where $g(z)$ is the baroclinic mode structure. The YBJ equation is cast in terms of the variable $\phi = (u_w + \mathrm{i}v_w)\mathrm{e}^{\mathrm{i}ft}$, where the factor $\mathrm{e}^{\mathrm{i}ft}$ removes oscillations at the inertial frequency and leaves ϕ to describe the slow evolution of the envelope that modulates the NIWs. For a single mode propagating through a barotropic background flow, the equation becomes

$$\frac{\partial \phi}{\partial t} + J(\psi, \phi) + \frac{\mathrm{i}\zeta}{2}\phi - \frac{\mathrm{i}f\lambda^2}{2}\nabla^2\phi = 0, \tag{1.2}$$

where ψ is the background streamfunction, $\zeta = \nabla^2\psi$ is the background vorticity and $J(a, b) = \partial_x a\,\partial_y b - \partial_y a\,\partial_x b$ is the Jacobian operator. The second term describes advection of the NIW field by the background flow. The third term is known as the ζ-refraction term and describes refraction of the NIW field by the background vorticity. This term is necessary to obtain concentration of NIWs into regions of anticyclonic vorticity. The last term is responsible for wave dispersion. Here and throughout this paper, we set the meridional gradient of planetary vorticity $\beta = 0$. The YBJ equation can be modified to include β by replacing $\zeta/2$ with $\zeta/2 + \beta y$ in the refraction term. The β effect has been proposed to explain the observed equatorward propagation of NIWs in the ocean (Anderson & Gill 1979; Garrett 2001; Yu *et al.* 2022), and it can dominate the overall NIW evolution in regions with weak mesoscale eddies (e.g. D'Asaro *et al.* 1995). Because mesoscale vorticity gradients typically dominate over β however, we here restrict ourselves to $\beta = 0$ for simplicity.

Despite both ray tracing and the YBJ equation being used in the NIW literature, it remains unclear how they relate to each other. Ray tracing has been one of the most widely used tools to interpret observations of NIWs. Ray tracing has had qualitative success in describing observed features of NIW evolution, however, we are not aware of any rigorous comparisons between ray tracing predictions and observations. Ray tracing has revealed aspects of NIW dynamics such as trapping in anticyclones along with an associated propagation to depth (Jaimes & Shay 2010), stalling in cyclones (Oey *et al.* 2008) and the interplay between NIWs and turbulent dissipation (Kunze, Schmitt & Toole 1995; Essink *et al.* 2022). Non-standard propagation patterns of NIWs in observations have also been explained using ray tracing (e.g. Byun *et al.* 2010; Chen *et al.* 2013). The YBJ equation has been used primarily as a tool in theoretical and numerical studies, although there has been some attempt to make connections with observations. Asselin & Young (2020) used simulations of the YBJ equation coupled to a quasi-geostrophic mesoscale eddy field to investigate the sequence of events that lead to the downward propagation of wind-forced NIWs. Thomas *et al.* (2020) calculated the NIW wave vector using an expression based on the YBJ equation. The predictions from YBJ were broadly in agreement with observations. Conn *et al.* (2024) directly used the YBJ equation to interpret NIW observations on a mooring array, showing that it successfully captured the amplitude and phase evolution, including differences across the mooring array caused by mesoscale vorticity gradients. Any comparison of the results of these disparate studies is complicated by the different methods used. A better understanding of the relationship between ray tracing and YBJ would clarify the physical similarities and differences.

Further complicating the picture, observations reveal a varied picture of the importance of the mesoscale vorticity on NIW evolution. During the Ocean Storms experiment, mesoscale eddies had a muted impact on the NIW field (D'Asaro 1995), whereas other observational studies found a strong imprint of mesoscale eddies onto the NIW field. For example, Thomas *et al.* (2020) demonstrated that the evolution of the NIW wave vector was driven by gradients in the mesoscale vorticity during the NISKINe experiment in the North Atlantic. Extending the original argument by YBJ, Thomas *et al.* (2024*a*) argued that these differences in the impact of mesoscale vorticity could be explained primarily by differences in the strength of wave dispersion. The stronger dispersion in the Ocean Storms experiment, they argued, was the result of the forcing projecting onto lower baroclinic modes, a stronger stratification and weaker eddies. As a result, the effect of refraction by mesoscale vorticity was suppressed in the Ocean Storms experiment, whereas it was more pronounced in NISKINe.

In this paper we aim to clarify how ray tracing relates to YBJ dynamics. Given the widespread use of ray tracing in the literature, we aim to understand the conditions under which results from ray tracing are accurate. To this end, we consider the YBJ equation in both a strong- and weak-dispersion regime. We begin by providing a simplified treatment of the strong-dispersion regime. Next, we show that the ray tracing equations emerge asymptotically from the YBJ equation in the limit of weak dispersion. Our analysis shows that the Wentzel–Kramers–Brillouin (WKB) approximation and, thus, ray tracing can be valid even in the presence of a large-scale forcing, despite the YBJ critique. The forcing decomposes into several modes that themselves exhibit small-scale structure. We find the existence of isotropic and anisotropic modes. The isotropic modes are characterised by fast variations along streamlines, while the anisotropic modes have weak variations along streamlines. We discuss the physical processes important in both classes. Finally, we consider how these regimes might modulate the energy injection into the NIW band by the winds, finding that such a modulation is likely weak under oceanic conditions.

2. The YBJ equation

2.1. *Decomposition into horizontal modes*

We begin by non-dimensionalising the YBJ equation. Given the scalings $x, y \sim L$, $\psi \sim \Psi$ and $t \sim L^2/\Psi$, we obtain the following non-dimensional form of the YBJ equation:

$$\frac{\partial \phi}{\partial t} + J(\psi, \phi) + \frac{\mathrm{i}\zeta}{2}\phi - \frac{\mathrm{i}\varepsilon^2}{2}\nabla^2\phi = 0. \tag{2.1}$$

Here $\varepsilon^2 = f\lambda^2/\Psi$ is the wave dispersiveness (assuming $f > 0$). For readers familiar with Young & Ben Jelloul (1997), our ε^2 is equivalent to their Υ^{-1}. We remind the reader that we have assumed a single baroclinic mode, but ε does vary among baroclinic modes through λ. The wave dispersiveness also varies spatially throughout the ocean (figure 1). We calculate ε for the first four baroclinic modes from observations as described in Appendix A. Except for the high latitudes, the first and second baroclinic modes are almost entirely in the strongly dispersive regime ($\varepsilon \gg 1$). Higher baroclinic modes are to be more weakly dispersive, with $\varepsilon < 1$ almost everywhere for mode 4. For a given baroclinic mode, low-latitude regions are more strongly dispersive, while higher latitudes and western boundary currents are more weakly dispersive.

Note that (2.1) is a Schrödinger equation. This parallel is made clear if we write (2.1) as

$$\mathrm{i}\frac{\partial \phi}{\partial t} = H\phi, \quad H = -\frac{\varepsilon^2}{2}\nabla^2 - \mathrm{i}J(\psi, \cdot) + \frac{\zeta}{2}. \tag{2.2a,b}$$
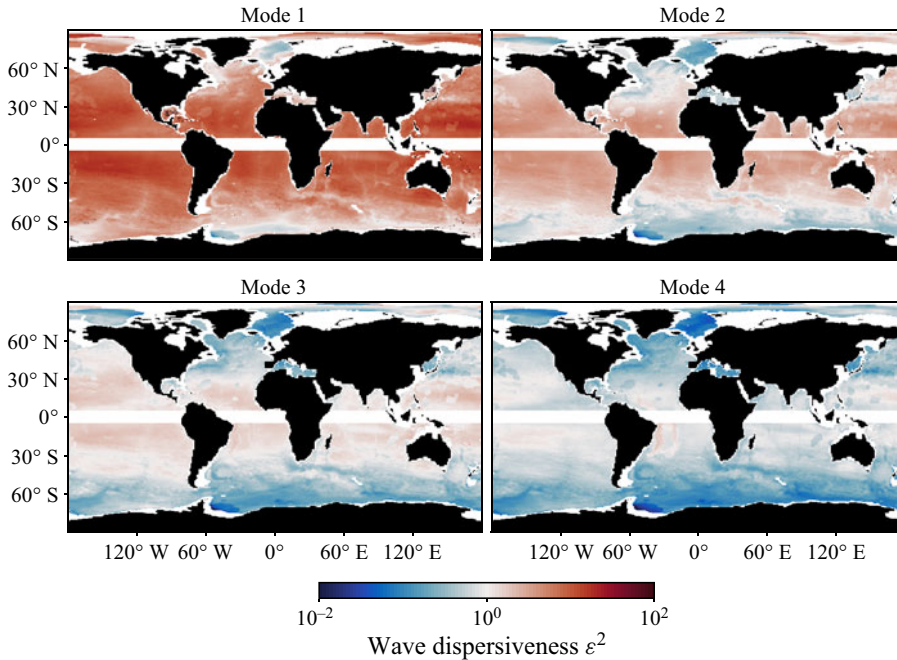
Figure 1. Wave dispersiveness $\varepsilon^2 = |f|\lambda^2/\Psi$ plotted throughout the ocean for the first four baroclinic modes, with the deformation radius λ estimated from hydrography and the streamfunction magnitude Ψ from altimetry. The equatorial band is blocked out because the mean flow amplitude cannot be estimated with confidence there.

The operator H is known as the Hamiltonian operator. While the presence of first derivatives in the Hamiltonian stemming from advection may be unfamiliar to some, such terms arise in quantum mechanics when describing a charged particle in a magnetic field. This analogy to quantum mechanics was pointed out by Balmforth, Llewellyn Smith & Young (1998) and we will exploit it here extensively. Rocha *et al.* (2018) also used this analogy to derive the equivalent of Ehrenfest's theorem for NIWs, while Danioux, Vanneste & Bühler (2015) explained the concentration of NIWs into anticyclones via the analogue of quantum conservation laws. While we are setting $\beta = 0$ in this paper, we note that the quantum analogue to $\beta \neq 0$ is known as the 'Wannier–Stark ladder', where the potential due to the mesoscale vorticity modulates a linear ramp due to β (Balmforth & Young 1999).

The operator H is Hermitian, i.e.

$$\int \varphi^* H\phi \, \mathrm{d}^2\boldsymbol{x} = \int (H\varphi)^*\phi \, \mathrm{d}^2\boldsymbol{x} \qquad (2.3)$$

for sufficiently regular functions φ and ϕ so it has real eigenvalues. We also assume H is compact so that the eigenmodes form a complete orthonormal basis. Let $\boldsymbol{\mu}$ label the eigenmodes $\hat{\phi}_{\boldsymbol{\mu}}(x, y)$ and associated eigenvalues $\omega_{\boldsymbol{\mu}}$ of the operator H,

$$H\hat{\phi}_{\boldsymbol{\mu}} = \omega_{\boldsymbol{\mu}}\hat{\phi}_{\boldsymbol{\mu}}. \qquad (2.4)$$

The field ϕ can then be expanded in the eigenmode basis as

$$\phi(x, y, t) = \sum_{\boldsymbol{\mu}} a_{\boldsymbol{\mu}}(t)\hat{\phi}_{\boldsymbol{\mu}}(x, y), \qquad (2.5)$$
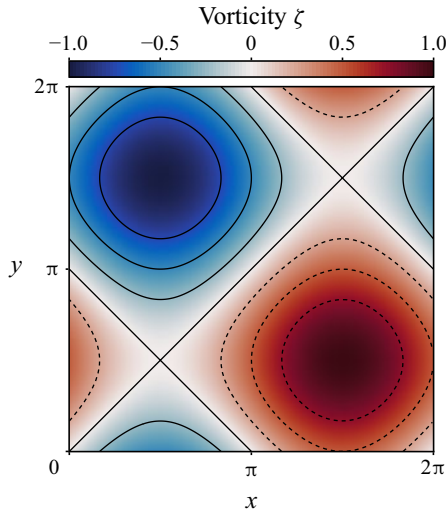
1002 A22-5

Figure 2. Dipole vorticity with an anticyclone in the upper left corner and a cyclone in the lower right corner. The contours depict the streamfunction with positive values denoted by solid lines and negative values denoted by dashed lines.

where $a_\mu(t)$ is the projection of ϕ onto the eigenmode $\hat{\phi}_\mu$. The coefficients $a_\mu(t)$ then evolve according to

$$\frac{\mathrm{d}a_\mu}{\mathrm{d}t} = -\mathrm{i}\omega_\mu a_\mu, \quad \text{so} \quad a_\mu(t) = a_\mu(0)\exp(-\mathrm{i}\omega_\mu t). \tag{2.6}$$

Therefore, the eigenvalue ω_μ represents the frequency shift of the mode away from f. The total dimensional NIW frequency is hence given by $f(1 + Ro\,\omega_\mu)$, where $Ro = \Psi/fL^2$ is the Rossby number. Furthermore, because the eigenvalues are real and the modes are orthogonal, the kinetic energy of the waves is conserved.

We consider this problem on a doubly periodic domain with size $2\pi \times 2\pi$. This is intended to represent a local view of an ocean that is filled with a random sea of eddies. The solutions we calculate are perfectly periodic and extend across all eddies. In reality, of course, the background field is not perfectly periodic and this causes the solutions to become localised in certain regions. Therefore, the solutions we calculate on the $2\pi \times 2\pi$ should be thought of similarly.

As a key example in this paper, we consider a $2\pi \times 2\pi$ domain that contains a dipole vortex given by (figure 2; cf. Asselin *et al.* 2020)

$$\psi = \tfrac{1}{2}\left(\sin x - \sin y\right). \tag{2.7}$$

The analysis below is general, however, and can be applied to more general background flows.

2.2. *Numerical calculation of eigenvalues and eigenmodes*

For most choices of the background flow ψ, analytical solutions for the eigenfunctions of H do not exist and numerical solutions are required. Solving the eigenvalue equation numerically requires us to discretise the operator H. The discrete eigenfunction is expressed as a vector, and the problem reduces to finding the eigenvalues of a finite matrix.

The operator H is Hermitian and so it is desirable for any discrete representation of H to also be Hermitian. A fourth-order central finite difference scheme for the Laplacian term preserves this property. More care is required for the advection operator, for which we use the enstrophy-conserving scheme from Arakawa (1966) to preserve the Hermitian nature of the operator and guarantee that the eigenvalues of the matrix are real. Having real eigenvalues ensures that the conservation of NIW kinetic energy is respected in the discrete system. The exact method of numerically solving the eigenvalue problem is detailed in Appendix B.

3. The strong-dispersion limit

The limit $\varepsilon \gg 1$ is known as the strong-dispersion limit. Young & Ben Jelloul (1997) showed that in this limit, the solution to the YBJ equation becomes proportional to the streamfunction ψ. They additionally showed that frequency shifts away from f are proportional to the domain-averaged kinetic energy of the mesoscale flow. These same results can be derived by considering the eigenvalue problem posed above; see (3.8) and (3.11) for the result. In our framework, we can additionally derive information about the next-order perturbations to the NIW field; see (3.15) and (3.16) below.

When ε is large, we split the operator H into two parts, i.e.

$$H = \varepsilon^2 H^{(0)} + H^{(1)}, \tag{3.1}$$

where $H^{(0)} = -\frac{1}{2}\nabla^2$ and $H^{(1)} = \frac{1}{2}\zeta - \mathrm{i}J(\psi, \cdot)$. Because $\varepsilon^2 \gg 1$, this implies that $H^{(1)}$ is a small correction to $\varepsilon^2 H^{(0)}$, and perturbation theory can be used to solve this system. We expand both $\hat{\phi}_{\boldsymbol{\mu}}$ and $\omega_{\boldsymbol{\mu}}$ in powers of ε^{-2}:

$$\hat{\phi}_{\boldsymbol{\mu}} = \sum_{n=0}^{\infty} \varepsilon^{-2n} \hat{\phi}_{\boldsymbol{\mu}}^{(n)}, \quad \omega_{\boldsymbol{\mu}} = \varepsilon^2 \sum_{n=0}^{\infty} \varepsilon^{-2n} \omega_{\boldsymbol{\mu}}^{(n)}. \tag{3.2a,b}$$

At $O(\varepsilon^2)$ the eigenvalue problem is

$$H^{(0)}\hat{\phi}_{\boldsymbol{\mu}}^{(0)} = \omega_{\boldsymbol{\mu}}^{(0)} \hat{\phi}_{\boldsymbol{\mu}}^{(0)}, \tag{3.3}$$

where $\hat{\phi}_{\boldsymbol{\mu}}^{(0)}$ is the eigenfunction of the unperturbed problem with eigenvalue $\omega_{\boldsymbol{\mu}}^{(0)}$. We assume the domain is doubly periodic and goes from 0 to 2π in x and y. The solution is

$$\hat{\phi}_{\boldsymbol{\mu}}^{(0)} = \mathrm{e}^{\mathrm{i}\boldsymbol{\mu}\cdot\boldsymbol{x}}, \quad \omega_{\boldsymbol{\mu}}^{(0)} = \frac{|\boldsymbol{\mu}|^2}{2}, \tag{3.4a,b}$$

where $\boldsymbol{\mu}$ is a two-dimensional vector with integer components, such that the eigenfunctions are plane waves in x and y.

3.1. *The leading-order mode*

Near-inertial waves are forced by atmospheric storms, which have a much larger horizontal scale than mesoscale eddies and can be idealised as a uniform forcing. We assume that the result of this forcing is the excitation of a constant non-zero ϕ. The projection of this initial condition onto a given mode can thus be found by integrating that mode across the domain. For plane waves, a domain integral will vanish unless $\boldsymbol{\mu} = 0$, such that a uniform forcing will only project onto the $\boldsymbol{\mu} = 0$ mode in the unperturbed case. We begin by focusing on that case to obtain expressions for the perturbations to its spatial structure as well as its

frequency shift. A small part of the forcing, however, projects onto modes with $\mu \neq 0$ and we will return to these higher modes below.

The leading-order solution for $\mu = 0$ is $\hat{\phi}_0^{(0)} = 1$ and $\omega_0^{(0)} = 0$, and there is no modulation of the waves by the mesoscale eddy field. To obtain this modulation, we must go to higher order. At $O(\varepsilon^0)$, the eigenvalue problem is

$$H^{(0)}\hat{\phi}_0^{(1)} + H^{(1)}\hat{\phi}_0^{(0)} = \omega_0^{(0)}\hat{\phi}_0^{(1)} + \omega_0^{(1)}\hat{\phi}_0^{(0)}. \tag{3.5}$$

With $\omega_0^{(0)} = 0$ and the advection term in $H^{(1)}$ vanishing when acting on $\hat{\phi}_0^{(0)} = 1$, this reduces to

$$-\frac{1}{2}\nabla^2\hat{\phi}_0^{(1)} + \frac{\zeta}{2} = \omega_0^{(1)}. \tag{3.6}$$

The two terms on the left vanish when integrated over the doubly periodic domain, so we conclude that $\omega_0^{(1)} = 0$.

There is, however, a correction to the eigenfunction at this order, determined by

$$\nabla^2\hat{\phi}_0^{(1)} = \nabla^2\psi. \tag{3.7}$$

With periodic boundary conditions, the solution to this is

$$\hat{\phi}_0^{(1)} = \psi, \tag{3.8}$$

where we have assumed that ψ is defined such that it has zero domain average. This recovers the expression for $\hat{\phi}$ from YBJ. The structure of the mesoscale eddy field is imprinted onto the waves by the $\varepsilon^{-2}\hat{\phi}_0^{(1)}$ term. Because the modulation is by the real streamfunction ψ, only the NIW amplitude is modulated by mesoscale eddies. The NIW field remains in phase across the domain.

We now also seek the leading non-zero correction to the eigenvalue, for which we go up another order. The eigenvalue equation at $O(\varepsilon^{-2})$ is

$$H^{(0)}\hat{\phi}_0^{(2)} + H^{(1)}\hat{\phi}_0^{(1)} = \omega_0^{(0)}\hat{\phi}_0^{(2)} + \omega_0^{(1)}\hat{\phi}_0^{(1)} + \omega_0^{(2)}\hat{\phi}_0^{(0)}. \tag{3.9}$$

With $\omega_0^{(0)} = \omega_0^{(1)} = 0$ and $J(\psi, \psi) = 0$, this simplifies to

$$-\frac{1}{2}\nabla^2\hat{\phi}_0^{(2)} + \frac{1}{2}\psi\nabla^2\psi = \omega_0^{(2)}. \tag{3.10}$$

The first term on the left vanishes under domain integration. Integrating the second term on the left by parts yields

$$\omega_0^{(2)} = -\frac{1}{2}\frac{\displaystyle\int |\boldsymbol{\nabla}\psi|^2 \, d^2\boldsymbol{x}}{\displaystyle\int d^2\boldsymbol{x}}. \tag{3.11}$$

The leading-order frequency shift is $\varepsilon^{-2}\omega_0^{(2)}$. Given that $\varepsilon^{-2} \ll 1$, the frequency shift away from f is suppressed substantially, even compared with the small frequency shift assumed from the outset. Re-dimensionalising the expression results in

$$\omega_0^{(2)} = -\frac{1}{2f_0\lambda^2}\frac{\displaystyle\int |\boldsymbol{\nabla}\psi|^2 \, d^2\boldsymbol{x}}{\displaystyle\int d^2\boldsymbol{x}}. \tag{3.12}$$

This agrees with the YBJ result for the dispersion relation in the strong-dispersion regime, indicating that the frequency shift is proportional to the average kinetic energy of the eddy field.

1002 A22-8

3.2. *Higher-order modes*

We now return to the higher modes with $\mu \neq 0$. These modes are degenerate to leading order. For example, the modes $(1, 0)$, $(-1, 0)$, $(0, 1)$ and $(0, -1)$ all have $\omega_{\mu}^{(0)} = \frac{1}{2}$. Degenerate perturbation theory is necessary to calculate the first-order corrections to the eigenvalues and eigenfunctions (e.g. Sakurai & Napolitano 2020). To obtain these corrections, we proceed naively with the calculation. We will run into a contradiction that motivates us to choose a different basis set than was chosen in (3.4*a*,*b*). To those familiar with degenerate perturbation theory, this may seem unnecessary, but we believe it to be more pedagogical.

We again start from the $O(\varepsilon^0)$ equation, which now reads

$$H_0\hat{\phi}_{\mu}^{(1)} + H_1\hat{\phi}_{\mu}^{(0)} = \omega_{\mu}^{(0)}\hat{\phi}_{\mu}^{(1)} + \omega_{\mu}^{(1)}\hat{\phi}_{\mu}^{(0)}. \tag{3.13}$$

Multiplying this equation by $\hat{\phi}_{\nu}^{(0)*}$, with both ν and μ labelling one of the modes in the degenerate group, and integrating over the domain results in

$$\int \hat{\phi}_{\nu}^{(0)*}\left(H_0 - \omega_{\mu}^{(0)}\right)\hat{\phi}_{\mu}^{(1)}\,\mathrm{d}^2x = \omega_{\mu}^{(1)}\int \hat{\phi}_{\nu}^{(0)*}\hat{\phi}_{\mu}^{(0)}\,\mathrm{d}^2x - \int \hat{\phi}_{\nu}^{(0)*}H_1\hat{\phi}_{\mu}^{(0)}\,\mathrm{d}^2x. \tag{3.14}$$

Using integration by parts, the H_0 on the left can be swapped for $\omega_{\nu}^{(0)}$. Because the modes are degenerate to this order, the left-hand side vanishes. Furthermore, using the orthonormality of the eigenfunctions, the corrections to the eigenvalues are determined by

$$\omega_{\mu}^{(1)}\delta_{\nu\mu} = \frac{1}{4\pi^2}\int \hat{\phi}_{\nu}^{(0)*}H_1\hat{\phi}_{\mu}^{(0)}\,\mathrm{d}^2x. \tag{3.15}$$

We have now arrived at our contradiction. The left-hand side of this equation is diagonal, whereas the right-hand side is not necessarily so. In (3.4*a*,*b*) we chose a basis for the unperturbed eigenfunctions: $\{e^{ix}, e^{-ix}, e^{iy}, e^{-iy}\}$ for $|\mu| = 1$. The key to degenerate perturbation theory is to choose a basis of the degenerate space to avoid this contradiction. It is clear that the correct basis must diagonalise H_1, which our original choice does not.

To proceed, we calculate the right-hand side of (3.15) in the original basis. This results in a 4×4 matrix. We diagonalise this matrix and find the corresponding linear combination of the original basis functions that diagonalises H_1. The corresponding eigenfunction corrections can be found by solving the screened Poisson equation obtained from the first-order equation (3.13). Rearranging, we have

$$\left(H_0 - \omega_{\mu}^{(0)}\right)\hat{\phi}_{\mu}^{(1)} = \left(\omega_{\mu}^{(1)} - H_1\right)\hat{\phi}_{\mu}^{(0)}, \tag{3.16}$$

where the $\hat{\phi}_{\mu}^{(0)}$ should be in the basis diagonalising H_1. If H_1 identically vanishes in this subspace, the degeneracy must be lifted at the next order, as in the example below.

3.3. *Dipole flow solutions*

We now consider the specific example of the dipole flow (2.7). Numerical solutions for $\varepsilon = 2$ show that a uniform initial condition projects strongly (98.5 % of the energy) onto the $\hat{\phi}_0$ mode (figure 3). There is a small but negative frequency shift of $\omega_0 = -0.03104$. This agrees excellently with the predicted frequency shift from (3.11) of $\varepsilon^{-2}\omega_0^{(2)} = \frac{1}{32} = -0.03125$. Additionally, there is weak horizontal structure that aligns with the streamfunction as expected. The root-mean-squared error between the
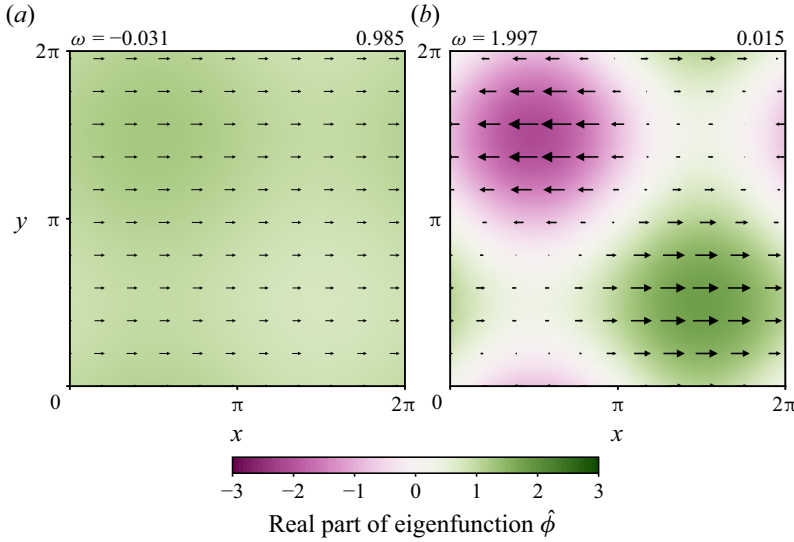
Figure 3. Numerical solution to the eigenvalue problem (2.4) with $\varepsilon = 2$ for the dipole flow. A uniform forcing primarily projects onto the mode shown on the left, with the mode with the second-highest projection shown on the right only making up less than 2 % of the energy. The eigenvalues ω and projection fractions (of energy) are shown in the panel titles. Vectors show the corresponding NIW velocities.

numerical eigenmode and the analytical eigenmode $\hat{\phi}_0^{(0)} + \varepsilon^{-2}\hat{\phi}_0^{(1)} = 1 + \varepsilon^{-2}\psi$ is 1 %. The agreement is excellent despite ε not being particularly large.

For the dipole flow, the right-hand side of (3.15) is zero for all combinations of basis functions of the $\omega_\mu^{(0)} = \frac{1}{2}$ subspace. Therefore, there are no first-order frequency shifts, $\omega_\mu^{(1)} = 0$, and the degeneracy is not lifted at this order. Performing the same procedure that led to (3.15) on the second-order equation yields

$$\omega_\mu^{(2)}\delta_{\nu\mu} = \frac{1}{4\pi^2}\int \hat{\phi}_\nu^{(0)*}H_1\hat{\phi}_\mu^{(1)}\,\mathrm{d}^2\mathbf{x}. \tag{3.17}$$

For our trial basis consisting of the four plane waves, we solve the screened Poisson equation (3.16) for the corresponding $\hat{\phi}_\mu^{(1)}$. This is tedious but doable because the right-hand side is just a sum of sines and cosines. The equation for the second-order frequency shift can be diagonalised, and this time the eigenvalues are not zero and the degeneracy is lifted. We find for $\omega_\mu^{(2)}$ the values $-\frac{1}{96}$, $-\frac{7}{96}$, $-\frac{49}{96}$ and $-\frac{55}{96}$, only the first of which corresponds to an eigenfunction that the forcing projects onto at this order. The leading-order eigenfunction of that mode is $\hat{\phi}_\mu^{(0)} = -\psi$ (figure 3). The eigenvalue $\varepsilon^2\omega_\mu^{(0)} + \varepsilon^{-2}\omega_\mu^{(2)} = 1.99739$ is again in excellent agreement with the numerical eigenvalue of 1.99729.

In this regime, the horizontal structure in the waves primarily arises due to $\hat{\phi}_0^{(1)}$, which is suppressed by $O(\varepsilon^{-2})$. There is also horizontal structure due to modes with $\mu \neq 0$, but these are projected onto weakly; the fraction of the variance accounted for by such a mode is $O(\varepsilon^{-4})$ (Sakurai & Napolitano 2020). As such, the wave potential energy, which depends on horizontal gradients in the wave field, is also suppressed. Xie & Vanneste (2015) associated the generation of wave potential energy with a sink of the background eddy kinetic energy in a process known as stimulated generation.

Given the weak generation of horizontal structure, stimulated generation is weak in the strong-dispersion regime.

4. The weak-dispersion limit

The limit $\varepsilon \ll 1$ is known as the weak-dispersion limit. Because ε^2 multiplies the highest-order derivative in the eigenvalue equation, the limit $\varepsilon \to 0$ is a singular perturbation problem. Before addressing the general problem, we build intuition with two simple examples. These examples suggest that there are two classes of modes. One class is characterised by waves that vary slowly along the streamlines of the background flow and more rapidly across streamlines; they are captured by an anisotropic scaling of the wavenumber with ε. The other class has even faster variations in both directions and requires an isotropic scaling. We develop a uniformly valid approximation that captures both of these classes.

4.1. *Parallel shear flow*

We begin with an example of a parallel shear flow in which the streamfunction ψ is a function of x only. The symmetry in y means the problem reduces to a one-dimensional eigenvalue problem. Balmforth *et al.* (1998) considered this problem for a specific example of a shear flow that can be solved in closed form. Zhang & Xie (2023) considered the limits of strong and weak dispersion for the same mean flow. Here, we address how the weak-dispersion limit can be analysed for a general parallel shear flow and apply the procedure to the example flow from Balmforth *et al.* (1998). Our goal is to calculate the structure of the eigenmodes and their corresponding eigenvalues. We begin by introducing the WKB method from which the two scalings arise. For the anisotropic scaling, the key result is (4.8); for the isotropic scaling, the equivalent result is (4.12).

We assume that the streamfunction $\psi(x)$ is periodic on the domain $[-\pi, \pi]$. The eigenvalue problem (2.4) reduces to

$$-\frac{\varepsilon^2}{2}\nabla^2\hat{\phi} - \mathrm{i}v\frac{\partial\hat{\phi}}{\partial y} + \frac{\zeta}{2}\hat{\phi} = \omega\hat{\phi}, \tag{4.1}$$

where $\zeta = \nabla^2\psi$ and $v = \partial_x\psi$ are both functions of x only, and we have suppressed the label on the eigenmode. The coefficients are independent of y, which motivates the ansatz $\hat{\phi} = \Phi(x)\mathrm{e}^{\mathrm{i}my}$. Given that the domain has width 2π in y, the wavenumber m must be an integer. With this ansatz, we are left with the one-dimensional eigenvalue problem

$$-\frac{\varepsilon^2}{2}\frac{\mathrm{d}^2\Phi}{\mathrm{d}x^2} + \left(\frac{\varepsilon^2 m^2}{2} + mv + \frac{\zeta}{2}\right)\Phi = \omega\Phi. \tag{4.2}$$

This is the Schrödinger equation of a particle in one-dimensional potential, with the bracketed term playing the role of the potential (Balmforth *et al.* 1998).

As ε is small, WKB analysis can be used to find approximations to the eigenvalues and eigenfunctions (e.g. Bender & Orszag 1999). In WKB theory the field Φ is expanded as

$$\Phi(x) = \exp\frac{1}{\delta}\sum_{j=0}^{\infty}\delta^j S_j(x), \tag{4.3}$$

S. Conn, J. Callies and A. Lawrence

where $\delta \ll 1$ is a scaling parameter that we are yet to determine. Substituting this into (4.2) yields

$$-\frac{\varepsilon^2}{2}\left[\frac{1}{\delta^2}\left(\sum_{j=0}^{\infty}\delta^j\frac{\mathrm{d}S_j}{\mathrm{d}x}\right)^2 + \frac{1}{\delta}\sum_{j=0}^{\infty}\delta^j\frac{\mathrm{d}^2S_j}{\mathrm{d}x^2}\right] + \frac{\varepsilon^2 m^2}{2} + mv + \frac{\zeta}{2} = \omega. \qquad (4.4)$$

If we assume that $m \sim O(1)$, both the refraction term and the advection terms are $O(1)$, and they must be balanced by a dispersion term of the same order. Requiring the lowest-order dispersion term to be $O(1)$ implies that $\delta = \varepsilon$, and the $O(1)$ equation becomes

$$-\frac{1}{2}\left(\frac{\mathrm{d}S_0}{\mathrm{d}x}\right)^2 + mv + \frac{\zeta}{2} = \omega. \qquad (4.5)$$

By writing $\varepsilon^{-1}\,\mathrm{d}S_0/\mathrm{d}x = \mathrm{i}k$, this equation is analogous to the dispersion relation (1.1a–c) specialised to this parallel shear flow. The function S_0 is found to be

$$S_0(x) = \pm\sqrt{2}\mathrm{i}\int^x \sqrt{\omega - mv(x') - \frac{\zeta(x')}{2}}\,\mathrm{d}x' \qquad (4.6)$$

and determines the leading-order phase variations of the solution. One can additionally show (Bender & Orszag 1999, (10.1.12)) that the next-order solution is

$$S_1(x) = -\frac{1}{4}\ln\left(\omega - mv - \frac{\zeta}{2}\right), \qquad (4.7)$$

which determines the leading-order amplitude modulation of the solution.

This asymptotic expansion is valid away from regions where the integrand above is zero. These are known as turning points of the problem and exist if $\omega < \max(mv + \zeta/2)$. The associated eigenfunctions are referred to as bound states. Near turning points, $\omega - mv - \zeta/2$ can be approximated by a linear function of x, and solutions to (4.5) are given by Airy functions. The Airy function solutions must be asymptotically matched to the solutions away from the turning points. This yields an integral constraint from which the eigenvalues ω can be determined. The problem as formulated above is the classic two-turning point problem, and the asymptotic matching procedure is well documented (e.g. Bender & Orszag 1999, (10.5.6)). The resulting condition for ω, often referred to as a quantisation condition, is

$$\frac{\sqrt{2}}{\varepsilon}\int_{x_0}^{x_1}\sqrt{\omega - mv(x) - \frac{\zeta(x)}{2}}\,\mathrm{d}x = \left(n + \frac{1}{2}\right)\pi, \quad \text{with } n = 0, 1, \ldots, \qquad (4.8)$$

where x_0 and x_1 are the turning points. The projection of a uniform forcing onto these modes can also be calculated asymptotically. The domain integral of a mode is dominated by contributions from the turning points (e.g. Bender & Orszag 1999, (10.4.24)).

If $\omega > \max(\zeta/2 + mv)$ then there are no turning points. The corresponding eigenmodes are referred to as free states, and the quantisation condition is replaced by

$$\frac{\sqrt{2}}{\varepsilon}\int_{-\pi}^{\pi}\sqrt{\omega - mv(x) - \frac{\zeta(x)}{2}}\,\mathrm{d}x = 2n\pi, \quad \text{with } n = 0, 1 \ldots. \qquad (4.9)$$

Note the lack of a half-integer shift that for bound states arises from the Airy behaviour near turning points. The lack of turning points in the free states also means (4.6) is valid

1002 A22-12

across the entire domain. Because the eigenfunctions of these free states are oscillatory in the entire domain, a uniform forcing projects only weakly onto them, and we do not discuss them any further. We also note that the discretisation of the free states is due to the periodic domain; they would be replaced by a continuum of modes in an infinite domain, while the bound states would remain discrete.

Under this scaling, the WKB modes are anisotropic. We assumed $m \sim O(1)$, which means that the modes' phase varies in y on a length scale $O(1)$. In contrast, the leading-order phase variations in x come from $\varepsilon^{-1} S_0$ and, therefore, occur on a scale $O(\varepsilon)$. The phase varies slowly along streamlines and rapidly across streamlines. This makes refraction and advection come in at the same order as cross-streamline dispersion.

An alternative would be to choose the scaling $m \sim O(\varepsilon^{-2})$. Repeating the WKB ansatz requires a choice of $\delta = \varepsilon^2$ and $\omega \sim O(\varepsilon^{-2})$ in order to end up with an equation of a similar form to (4.5):

$$-\frac{1}{2}\left(\frac{\mathrm{d}S_0}{\mathrm{d}x}\right)^2 + \frac{\varepsilon^4 m^2}{2} + \varepsilon^2 mv = \varepsilon^2 \omega. \tag{4.10}$$

With the scaling given above, each term is $O(1)$. We can solve for S_0 and the corresponding quantisation condition for bound modes:

$$S_0(x) = \pm\sqrt{2}\mathrm{i}\varepsilon \int^x \sqrt{\omega - \frac{\varepsilon^2 m^2}{2} - mv(x')}\,\mathrm{d}x', \tag{4.11}$$

$$\frac{\sqrt{2}}{\varepsilon} \int_{x_0}^{x_1} \sqrt{\omega - \frac{\varepsilon^2 m^2}{2} - mv(x)}\,\mathrm{d}x = \left(n + \frac{1}{2}\right)\pi. \tag{4.12}$$

These modes are isotropic. The phase variations in y occur on a scale $O(\varepsilon^2)$, which is the same as in x because phase variations in x now come from $\varepsilon^{-2} S_0$. This makes advection and along-streamline dispersion come in at the same order, and it makes refraction negligible.

Despite the different characteristics of the two scalings, they lead to similar quantisation conditions that differ only by what terms are included. We can combine them into a uniformly valid quantisation condition:

$$\frac{\sqrt{2}}{\varepsilon} \int_{x_0}^{x_1} \sqrt{\omega - \frac{\varepsilon^2 m^2}{2} - mv(x) - \frac{\zeta(x)}{2}}\,\mathrm{d}x = \left(n + \frac{1}{2}\right)\pi. \tag{4.13}$$

The 'potential' governing the wave evolution is therefore

$$V(x) = \frac{\varepsilon^2 m^2}{2} + mv(x) + \frac{\zeta(x)}{2}. \tag{4.14}$$

Under the anisotropic scaling $m \sim O(1)$, the along-streamline dispersion term is suppressed by a factor ε^2, leaving the $O(1)$ refraction and advection terms to dominate. Under the isotropic scaling $m \sim O(\varepsilon^{-2})$, the advection and along-streamline dispersion terms are enhanced by a factor ε^{-2} and dominate over a now negligible refraction term. In both cases, the general equation is obtained by retaining a term that is of higher order, which is allowed in an asymptotic theory. A uniform forcing only projects onto modes with $m = 0$, so all the modes projected onto are of the anisotropic variety.
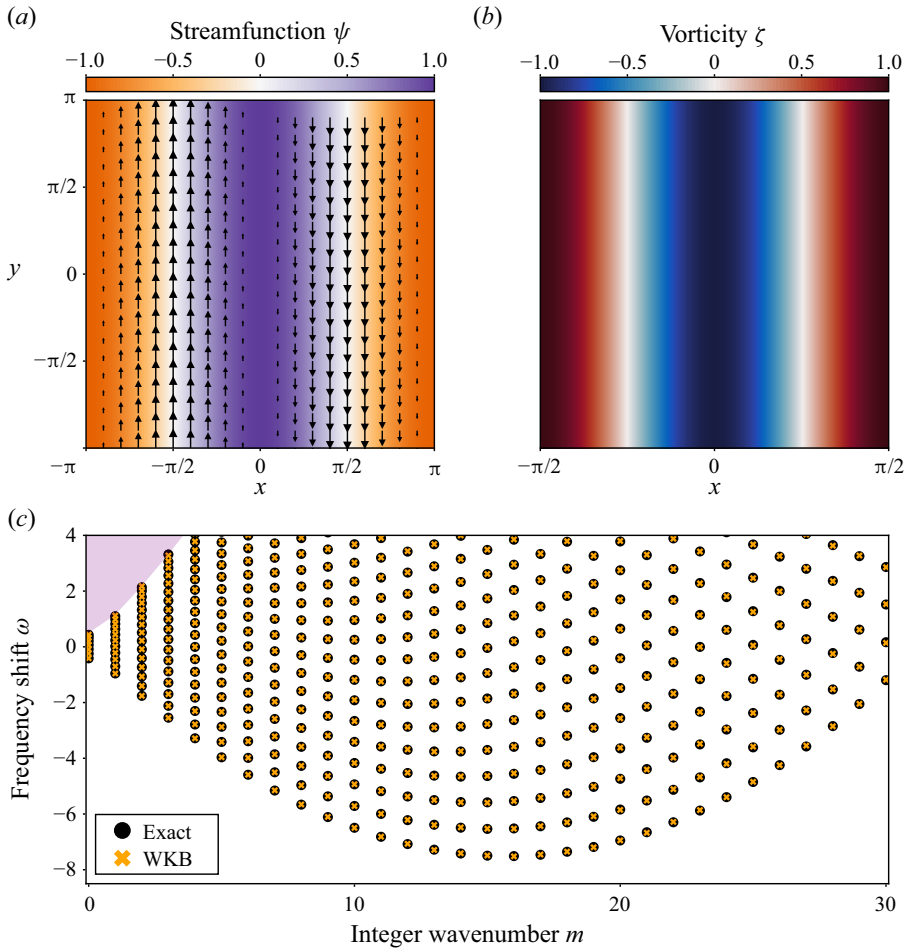
Figure 4. (*a*) Streamfunction and flow vectors for the shear flow example. (*b*) Vorticity, showing the anticyclonic vorticity concentrated in the centre of the domain and cyclonic vorticity on the outside. (*c*) Eigenvalues ω as a function of the integer wavenumber m for $\varepsilon = \frac{1}{4}$. The results from WKB theory (orange crosses) are shown along with the exact eigenvalues found from numerical solutions (black circles). The WKB results agree remarkably well with the numerical results, although there are some spurious eigenvalues near the boundary between free and bound modes. The purple shading indicates the region where free modes exist, which are not shown here.

We now consider a specific example of a parallel shear flow that varies sinusoidally in x:

$$\psi = \cos x. \tag{4.15}$$

This shear flow has a region of anticyclonic vorticity at the centre of the domain and cyclonic vorticity centred on $x = \pm\pi$ (figure 4*a,b*). This is a rare example in which the eigenvalue problem (4.2) can be solved exactly using Mathieu functions (Balmforth *et al.* 1998). The generally applicable WKB theory described above accurately predicts the eigenvalues, even for a modestly small $\varepsilon = \frac{1}{4}$ (figure 4*c*). We provide the analytical solutions to the WKB integrals in Appendix C. We also note that the symmetry of the problem means that a uniform wind forcing only projects onto modes with even n.

For $m = 0$, the eigenmodes are shaped by the potential $V = \zeta/2$ (figure 5*a*). Where $\omega > V$, S_0 is imaginary and the solutions are oscillatory; where $\omega < V$, S_0 is real and the
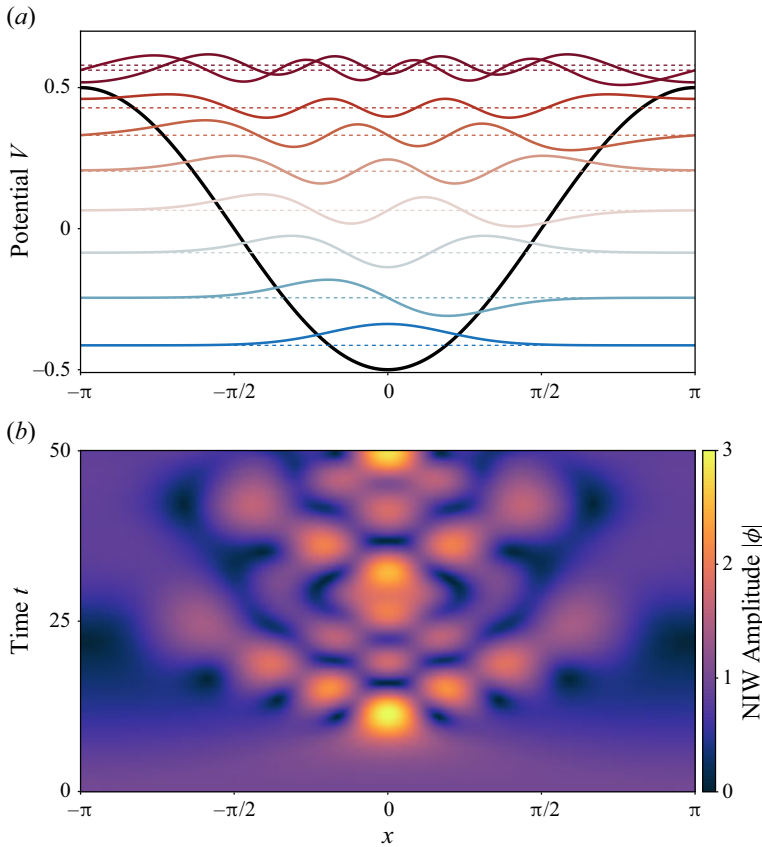
(*a*)

(*b*)

Figure 5. (*a*) The potential $V = \zeta/2$ (black line) of the parallel shear flow for the $m = 0$ mode. The dashed lines show the level of each eigenvalue ω for $\varepsilon = \frac{1}{4}$. The solid coloured lines represent a scaled representation of each eigenfunction corresponding to a given eigenvalue, as identified by the colours. (*b*) Time evolution of the NIW amplitude $|\phi|$ for the parallel shear flow example with $\varepsilon = \frac{1}{4}$, starting by a uniform field.

solutions are decaying (figure 5*a*). Near the anticyclonic centre of the flow, the potential is at its lowest and all the modes are oscillatory. Moving further out into the cyclonic region, more and more of the modes become evanescent.

The dependence of the potential on the vorticity ζ leads to trapping of NIWs in anticyclones. The trapping arises from the dephasing of the modes that make up the initial condition. This is analogous to the argument in Gill (1984) regarding the vertical propagation of NIWs due to the β effect. The evolution occurs in three phases (figure 5*b*). First, refraction imprints the mesoscale vorticity onto the initially uniform wave phase, leaving the amplitude unchanged (cf. Asselin *et al.* 2020). Second, once these phase gradients are sufficiently pronounced, cross-streamline dispersion becomes important and concentrates the wave energy into the centre of the anticyclone, and the wave field assumes a spatial scale $O(\varepsilon)$. Third, the amplitude remains elevated on average within the anticyclonic region but is more spread out than during the initial concentration. It is this long-time behaviour that corresponds to the fully dephased eigenmodes. The time it takes for this dephasing to occur depends inversely on the spacing of the eigenvalues ω. As ε decreases, the eigenvalues become more finely spaced. The $m = 0$ modes have a spacing $O(\varepsilon)$, so it takes $t \sim O(\varepsilon^{-1})$ for them to dephase. Another way to think of this is that as ε

decreases, dispersion becomes weaker and it takes longer for phase gradients to build up to a level where dispersion is important.

Finally, we note that Asselin *et al.* (2020) discussed solutions to the YBJ equation for which phase lines are aligned with streamlines and straining is ineffective in driving a decrease in the spatial scale of the wave. That analysis sets the dispersion term to zero, however, and so only captures the initial phase in which refraction dominates. The WKB theory presented above shows that cross-streamline dispersion is of leading order and should not be dropped if the long-term evolution is of interest (cf. figure 5*b*). Our $m = 0$ anisotropic modes can thus be understood as a generalisation of Asselin *et al.*'s (2020) solution. The ineffectiveness of straining due to the alignment of the wave phase with streamlines remains apparent, but cross-streamline dispersion is now taken into account such that the solution remains valid at late times. We further note that our anisotropic modes also allow for slow variations of the wave field along streamlines, such that advection assumes the same importance as refraction and cross-streamline dispersion. These anisotropic modes with $m > 0$ may be excited by a non-uniform forcing, such as a passing atmospheric front (cf. Thomas, Smith & Bühler 2017).

4.2. *Axisymmetric flow*

We now consider a streamfunction with axial symmetry, such that $\psi = \psi(r)$, where r is the radial distance from the origin. Llewellyn Smith (1999) studied NIWs with azimuthal wavenumber zero in an axisymmetric vortex and provided asymptotic expressions for the frequency of the lowest radial mode. Kafiabad, Vanneste & Young (2021) studied a similar case but also considered the impact of NIWs back on the vortex. Using WKB theory, we consider NIWs with an arbitrary azimuthal wavenumber and provide a transcendental equation that can be solved for their frequency as for the parallel shear flows above.

We make the ansatz $\hat{\phi} = A(r)e^{im\theta}$, where θ is the azimuthal angle, and again we drop the mode label. In polar coordinates, (2.4) then reduces to

$$-\frac{\varepsilon^2}{2}\left(\frac{d^2A}{dr^2} + \frac{1}{r}\frac{dA}{dr}\right) + \left(\frac{\varepsilon^2 m^2}{2r^2} + \frac{mv}{r} + \frac{\zeta}{2}\right)A = \omega A, \qquad (4.16)$$

where $v = \partial_r \psi$ denotes the azimuthal velocity. There are some subtleties involved in applying WKB theory to this equation. For modes with $m > 0$, the potential diverges at the origin. This issue has long been noted in the quantum mechanics literature and can be addressed by performing a so-called Langer transform on the equation. For $m = 0$, there is no divergence of the potential, but there is a phase shift at the origin. As pointed out by Berry & Ozorio de Almeida (1973), both cases turn out to give the same quantisation condition:

$$\frac{\sqrt{2}}{\varepsilon}\int_{r_0}^{r_1}\sqrt{\omega - V(r)}\,dr = \left(n + \frac{1}{2}\right)\pi, \quad \text{with } n = 0, 1, 2, \ldots. \qquad (4.17)$$

Here the potential is

$$V(r) = \frac{\varepsilon^2 m^2}{2r^2} + \frac{mv}{r} + \frac{\zeta}{2}. \qquad (4.18)$$

If $m > 0$, the integration bounds r_0 and r_1 are the two zeros of the integrand; if $m = 0$, $r_0 = 0$ and r_1 is the one zero of the integrand. As in the case of a parallel shear flow, this expression is uniformly valid in the sense that it works for both $m \sim O(1)$ and $m \sim O(\varepsilon^{-2})$. These again correspond to anisotropic and isotropic modes, respectively, with

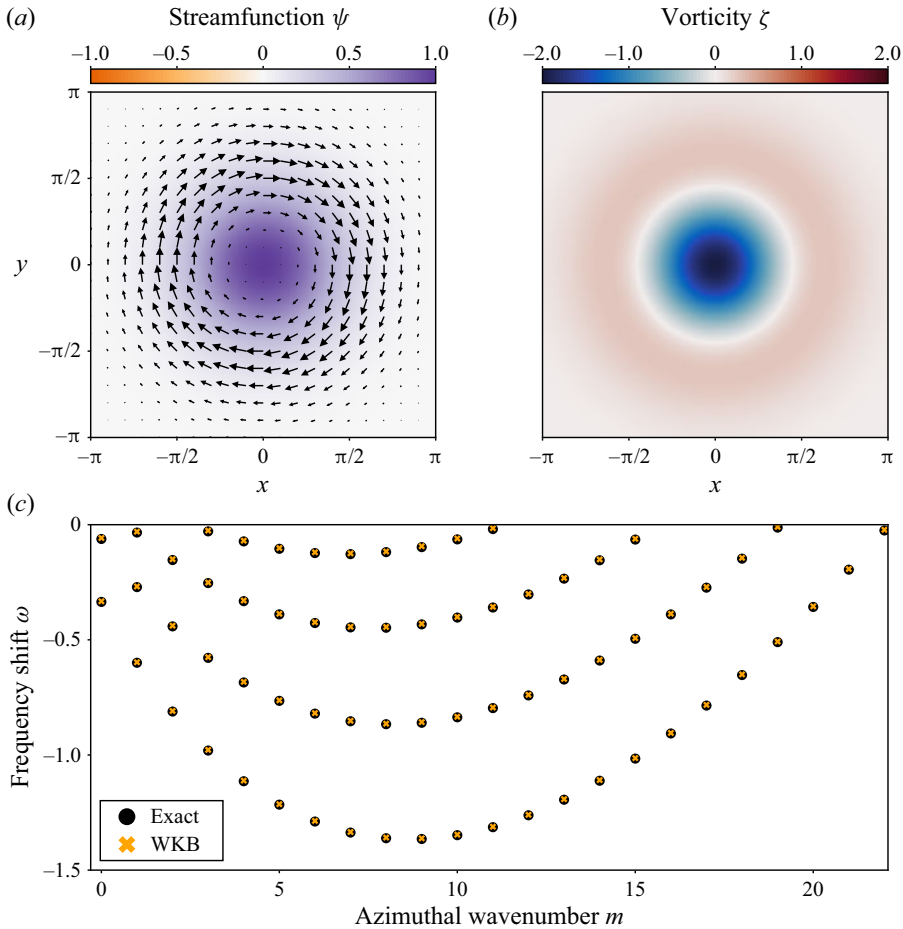(a) Streamfunction ψ (b) Vorticity ζ

Figure 6. (*a*) Streamfunction and flow vectors for the axisymmetric flow example. (*b*) Vorticity field showing the anticyclonic vorticity concentrated in the centre of the domain, which is flanked by a halo of cyclonic vorticity. (*c*) Eigenvalues ω as a function of azimuthal wavenumber m for $\varepsilon = \frac{1}{4}$. The results from WKB theory (orange crosses) are shown along with the exact eigenvalues found from numerical solutions (black circles). The WKB approximation agrees remarkably well with the numerical results.

refraction, advection and dispersion along and across streamlines playing the same roles as before. The only difference is that the streamlines are now circular.

We consider the concrete example of an isolated Gaussian vortex on an infinite domain:

$$\psi(r) = \mathrm{e}^{-r^2/4}. \tag{4.19}$$

This corresponds to an anticyclone in the centre of the domain that is surrounded by a halo of cyclonic vorticity (figure 6*a*,*b*). Again, the WKB calculation for $\varepsilon = \frac{1}{4}$ yields eigenvalues that agree extremely well with the exact eigenvalues (figure 6*c*). The structure of the first few modes is shown in figure 7. For $m = 0$, the modes are concentrated in the anticyclone. For $m > 0$, there is a repulsion from the very centre of the anticyclone due to the advection and dispersion terms in $V(r)$. This repulsion increases with m, but the modes remain primarily concentrated in the region of anticyclonic vorticity. The modes become
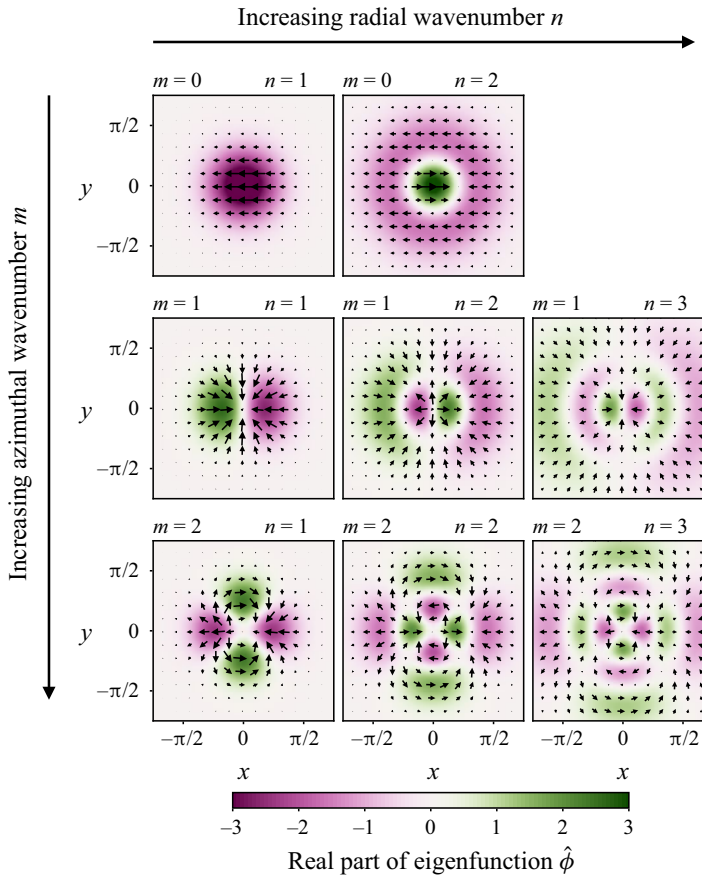
Figure 7. Real part of the eigenfunctions for the axisymmetric flow example with $\varepsilon = \frac{1}{4}$. The radial wavenumber n increases from left to right and corresponds to an increasing number of nodes in the radial direction. The azimuthal wavenumber increases from top to bottom and corresponds to an increasing number of nodes in the azimuthal direction.

more isotropic as m is increased. Note that again, a uniform forcing only projects onto the $m = 0$ mode due to the symmetry of the vorticity field. We also note that only the bound states form a discrete spectrum in an infinite domain, there will also be a continuum of free modes.

4.3. *General case*

Based on the intuition gained above, we wish to construct a uniformly valid asymptotic expansion for a general two-dimensional background flow. We again make a WKB ansatz that leads to (4.23). This equation can be solved by the method of characteristics and recovers Kunze's ray tracing. The quantisation condition for the general case is (4.27*a,b*).

In analogy with the isotropic scaling, we begin by assuming a solution of the form

$$\hat{\phi}(x, y) = \exp\left[\frac{1}{\varepsilon^2} \sum_{j=0}^{\infty} \varepsilon^{2j} S_j(x, y) \right], \tag{4.20}$$

again dropping the mode label. Substituting this into (2.4) yields

$$-\frac{1}{2\varepsilon^2}\left|\sum_{j=0}^{\infty}\varepsilon^{2j}\boldsymbol{\nabla} S_j\right|^2 - \frac{1}{2}\sum_{j=0}^{\infty}\varepsilon^{2j}\nabla^2 S_j - \frac{\mathrm{i}}{\varepsilon^2}\sum_{j=0}^{\infty}\varepsilon^{2j}J(\psi, S_j) + \frac{\zeta}{2} = \omega. \qquad (4.21)$$

Assuming $\omega \sim O(\varepsilon^{-2})$ and collecting leading-order terms, we obtain

$$-\tfrac{1}{2}|\boldsymbol{\nabla} S_0|^2 - \mathrm{i}\,J(\psi, S_0) = \varepsilon^2\omega. \qquad (4.22)$$

In the simple examples discussed above, we obtained a uniformly valid approximation by retaining the higher-order refraction term in the leading-order equation arising from an isotropic scaling. We do so again here:

$$-\frac{1}{2}|\boldsymbol{\nabla} S_0|^2 - \mathrm{i}\,J(\psi, S_0) + \frac{\varepsilon^2\zeta}{2} = \varepsilon^2\omega. \qquad (4.23)$$

We anticipate that the order of these terms again changes for anisotropic modes. If the phase varies slowly along streamlines, the advection term is reduced by a factor $O(\varepsilon^2)$ and cross-streamline dispersion, acting on spatial variations on a scale of $O(\varepsilon)$ rather than $O(\varepsilon^2)$, will attain the same order, whereas along-streamline dispersion becomes negligible. Equation (4.23) can therefore capture both isotropic and anisotropic modes.

We now introduce the wavenumber vector $\boldsymbol{k}$ by writing $\varepsilon^{-2}\partial S_0/\partial\boldsymbol{x} = \mathrm{i}\boldsymbol{k}$. Equation (4.23) can be solved using the method of characteristics:

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau} = \varepsilon^2\boldsymbol{k} + \boldsymbol{u}, \quad \frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}\tau} = -\frac{\partial}{\partial\boldsymbol{x}}\left(\boldsymbol{u}\cdot\boldsymbol{k} + \frac{\zeta}{2}\right), \quad \omega = \frac{\varepsilon^2|\boldsymbol{k}|^2}{2} + \boldsymbol{u}\cdot\boldsymbol{k} + \frac{\zeta}{2}. \quad (4.24a\text{–}c)$$

These are the non-dimensionalised ray tracing equations of Kunze (1985). We further elaborate on this connection between YBJ and Kunze's ray tracing below.

Numerical solutions for the dipole flow show that the majority of a uniform forcing projects onto anisotropic modes that show little structure along streamlines and vary more rapidly across streamlines (figure 8). With $\varepsilon = \frac{1}{4}$ there is also some projection onto modes that show more characteristics of isotropic phase variations. The variations are more rapid, as emerges from the isotropic scaling discussed above.

Finally, we show how approximations to the eigenvalues can be obtained in the weak-dispersion limit when the flow problem is not separable, as it was in the cases of a parallel shear flow or axisymmetric flow. To this end, we utilise results from the quantum mechanics literature. Recall that the YBJ equation is equivalent to the Schrödinger equation, with the YBJ operator

$$H = -\frac{\varepsilon^2}{2}\nabla^2 - \mathrm{i}\boldsymbol{u}\cdot\boldsymbol{\nabla} + \frac{\zeta}{2} \qquad (4.25)$$

playing the role of the Hamiltonian. The weak-dispersion limit corresponds to the classical limit of the equivalent quantum system, and the ray tracing equations are the analogue of the classical Hamiltonian dynamics. The classical Hamiltonian is obtained from H by making the substitution $\boldsymbol{\nabla} \mapsto \mathrm{i}\boldsymbol{k}$, yielding the dispersion relation in (4.24a–c). The Hamiltonian dynamics is then

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau} = \frac{\partial\omega}{\partial\boldsymbol{k}} \quad \text{and} \quad \frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}\tau} = -\frac{\partial\omega}{\partial\boldsymbol{x}}, \qquad (4.26a,b)$$

the ray tracing equations stated in (4.24a–c). The connection with the Schrödinger equation is most easily seen in the Hamilton–Jacobi description of classical mechanics (e.g. Bühler 2006; Sakurai & Napolitano 2020).
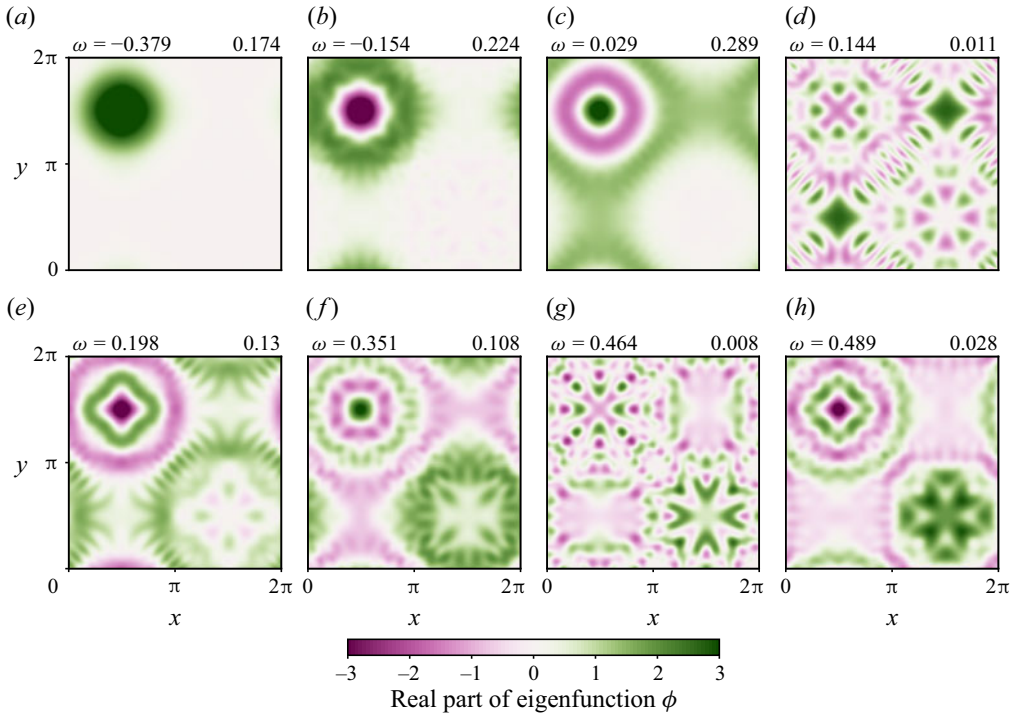
Figure 8. Real part of the eigenfunctions of the dipole flow with $\varepsilon = \frac{1}{4}$. Together, these eight eigenfunctions represent over 97 % of the energy excited by a uniform forcing. They are the eight modes with the strongest projection and are then ordered by eigenvalue ω. The eigenvalues are shown in the top left corner and the projections of a uniform forcing onto the eigenfunction (energy fraction) are shown in the top right corner.

The quantisation conditions derived above for separable problems, from which we obtained good approximations of the frequency shifts ω, can be generalised to some extent to non-separable problems like the dipole flow (figure 2). This semi-classical analysis of a quantum system was developed by Einstein (1917), Brillouin (1926) and Keller (1958), extending the Bohr–Sommerfeld quantum theory. The resulting approach is referred to as the Einstein–Brillouin–Keller (EBK) method (see also Berry & Mount 1972; Percival 1977; Keller 1985). The starting point is that the rays (classical trajectories in the quantum problem), being constrained by the invariant ω (energy in the quantum problem), trace out invariant tori in the phase space spanned by $\boldsymbol{x}$ and $\boldsymbol{k}$. A ray starting on such a torus will remain on it forever. The quantisation condition selects invariant tori that correspond to allowed bound states by insisting that phase increments along closed loops on the invariant torus integrate to multiples of 2π. Recalling that $\varepsilon^{-2} S_0 = i\boldsymbol{k}$, so $\boldsymbol{k}$ is the spatial gradient of the phase and $\boldsymbol{k} \cdot d\boldsymbol{x}$ is a phase increment, the quantisation conditions read

$$\oint_{\mathcal{C}_1} \boldsymbol{k} \cdot d\boldsymbol{x} = 2\pi \left(n + \frac{1}{2} \right), \quad \oint_{\mathcal{C}_2} \boldsymbol{k} \cdot d\boldsymbol{x} = 2\pi m, \qquad (4.27a,b)$$

where n and m are integers. The contours $\mathcal{C}_1$ and $\mathcal{C}_2$ are topologically independent closed curves on the invariant torus (figure 9a). In our example, the curve $\mathcal{C}_1$ passes through the hole of the phase space torus, whereas the curve $\mathcal{C}_2$ goes around the hole. The two curves
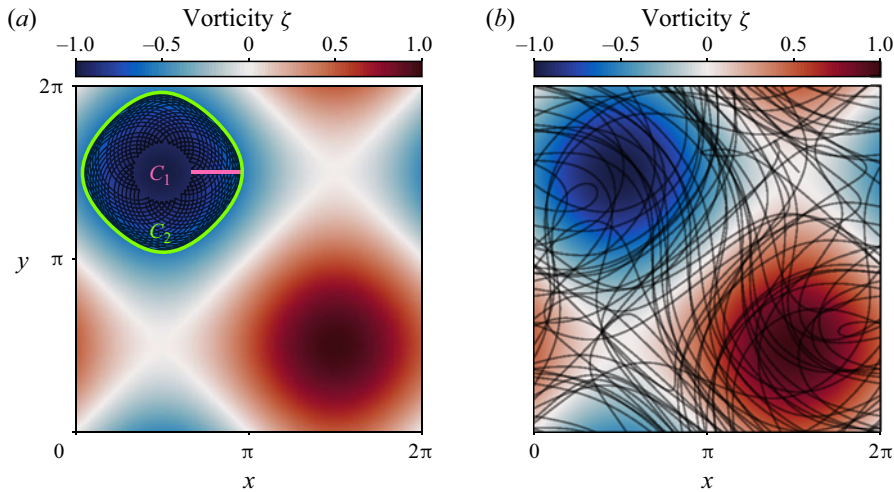
Figure 9. (*a*) Example of a trajectory tracing out an invariant torus for the dipole case. This torus corresponds to $n = 2$, $m = 0$ for $\varepsilon = \frac{1}{4}$. The background colours show the vorticity field. The black line shows a finite-time trajectory on the torus. The green and magenta lines represent a choice for the two invariant curves on the torus. They are independent because no continuous deformation of one can transform it into the other. (*b*) Different initial conditions result in different trajectories. This example is not bound to an invariant torus but is instead an example of a chaotic trajectory.

are independent in the sense that neither one can be continuously deformed into the other. There is a half-integer phase shift in the quantisation condition arising from the integral along the curve $\mathcal{C}_1$ because this curve passes through two caustics, the generalisation of a turning point, where additional phase shifts are incurred (Brillouin 1926; Keller 1958; Maslov 1972). The curve $\mathcal{C}_2$ encounters no caustics. The integer wavenumbers n and m correspond to the cross- and along-streamline variations, respectively. These EBK quantisation conditions are entirely analogous to the WKB quantisation conditions derived above for the separable parallel shear flow and axisymmetric flow.

We apply the EBK quantisation to the dipole flow with $\varepsilon = \frac{1}{4}$. Our procedure closely follows Percival & Pomphrey (1976): we find the invariant tori satisfying the quantisation condition by writing the Hamiltonian equations in action–angle variables and employing Newton's method; see Appendix D for details. All eigenvalues calculated by this EBK method show excellent agreement with the numerical values (figure 10).

As foretold by Einstein (1917), not all modes are accessible by the EBK approach. If the system is non-integrable, trajectories in phase space can become chaotic instead of tracing out an invariant torus (figure 9*b*). States corresponding to such chaotic trajectories are not amenable to the EBK method. This 'quantum chaos' has received much attention in the physics literature and has connections to random matrix theory (e.g. Gutzwiller 1992; Stone 2005). Methods exist to estimate eigenvalues as well as their statistics (e.g. Edelman & Rao 2005; Edelman & Sutton 2007). We do not pursue these issues any further here, in part because a uniform forcing projects most strongly onto the regular modes accessible with the EBK method (figure 8).

5. Relation to the ray tracing equations

The previous section made clear that the ray tracing equations of Kunze (1985) are closely related to the YBJ dynamics. In the same way that Hamiltonian dynamics emerge in
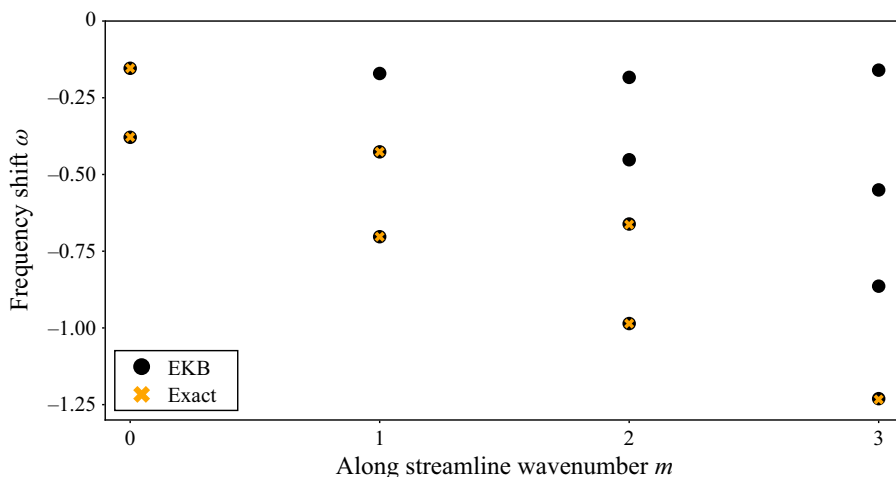
Figure 10. Numerical eigenvalues (black circles) and EBK eigenvalues (orange crosses) calculated for the dipole flow with $\varepsilon = \frac{1}{4}$. Here EBK calculations are only shown for the sufficiently confined modes where the invariant tori are easy to calculate. The EBK values agree with the numerical values to $O(10^{-3})$.

the classical limit of the Schrödinger equation, the ray tracing equations emerge in the weak-dispersion limit of the YBJ equation. Young & Ben Jelloul (1997) criticised Kunze's assumption that the waves have a smaller spatial scale than the background flow, insisting that atmospheric forcing produces NIWs at larger – not smaller – scales than mesoscale eddies, calling into question Kunze's ray-theoretical description in general. The analysis above clarifies that the spatial scale of the forcing is not what determines the applicability of WKB. Instead, the scale on which dynamical modes vary determines whether WKB analysis can be applied, and this spatial scale is set by how strongly dispersive the waves are. An initially uniform wave field can be thought of as consisting of a superposition of several modes, all varying on a small scale but combining into a uniform field. The distinct frequencies ω of these modes make them dephase over time, and the superposition develops the small scales of the modes.

Our analysis also provides some additional insight into the evolution of weakly dispersive NIWs. The isotropic and anisotropic scalings show that refraction is not always of leading-order importance. The refraction term is significant only for the anisotropic modes. For isotropic modes, the refraction term is asymptotically weak and the dispersion relation is dominated by advection and dispersion. A large-scale forcing, however, will project primarily onto the anisotropic modes, as can be seen in the specific solutions for the dipole case (figure 8). More generally, the large values of the along-streamline wavenumber m in the isotropic case produce rapid variations that lead to strong cancellations when calculating the projection of a uniform forcing onto these modes. As such, only a weak projection can remain.

To help interpret observations from the NISKINe study, Thomas *et al.* (2020) performed a simplified ray tracing calculation, which predicted a rapid strain-driven growth in the wavenumber that stood in stark contrast to the data. In this region of the North Atlantic the waves are weakly dispersive (Thomas *et al.* 2024a), so one may worry that this result contradicts our conclusion that ray tracing can be deployed gainfully in the weak-dispersion regime. Thomas *et al.* (2020) approximated the full wave-vector evolution by assuming a uniform and time-independent vorticity gradient, as well as a strain field with strain rate α and its principal axis aligned with the vorticity gradient.

In that set-up, the wavenumber component $k_\perp$ that is aligned with the vorticity gradient, i.e. perpendicular to vorticity contours, evolves according to

$$\frac{\mathrm{d}k_\perp}{\mathrm{d}\tau} = -\frac{|\nabla\zeta|}{2} + \alpha k_\perp, \quad \text{so} \quad k_\perp = -\frac{|\nabla\zeta|}{2\alpha}\left(\mathrm{e}^{\alpha\tau} - 1\right) \tag{5.1}$$

if $k_\perp = 0$ at time $\tau = 0$, approximating large-scale wind forcing. The exponential growth predicted by this equation does not match the data. Our analysis suggests, however, that a large-scale forcing primarily excites modes whose phase is aligned with streamlines. In this configuration, the strain is ineffective and the initial wavenumber evolution is dominated by refraction:

$$\frac{\mathrm{d}k_\perp}{\mathrm{d}\tau} = -\frac{|\nabla\zeta|}{2}, \quad \text{so} \quad k_\perp = -\frac{|\nabla\zeta|}{2}\tau. \tag{5.2}$$

This recovers the Asselin *et al.* (2020) solution that Thomas *et al.* (2020) showed roughly matches the data. Our analysis therefore suggests that it was not ray tracing *per se* that caused the mismatch with the data but the assumptions that went into the simplified solution. (It should be noted that a pure strain field does not produce a compact operator H, so the machinery based on a discrete set of eigenmodes does not apply. Instead, one should view the pure strain field as the local behaviour of some more complicated background flow that is described by a compact operator H. Because ray tracing is inherently local, the behaviour detailed above would still apply.)

Kunze (1985) considered three-dimensional ray tracing, which allows for both baroclinicity in the mean flow and a vertical wavenumber for the NIWs that corresponds to propagation of the waves in the vertical. In this paper we have restricted our attention to a barotropic mean flow and considered the propagation of a single baroclinic mode, such that the problem reduces to two-dimensional ray tracing. Exploring how the three-dimensional ray tracing is related to the full YBJ equation that also allows for baroclinicity in the background flow is left to future work.

6. Near-inertial wind work

One may speculate that the frequency shifts in the weak-dispersion limit could impact the energy input into NIWs by the winds. To study this, we need to consider a forced version of the YBJ equation. So far, we have focused on the problem with a horizontally uniform initial condition. This was to represent the NIW field excited by the passage of a large-scale atmospheric storm, and we studied the evolution of this NIW field in the absence of any further forcing. Real NIWs, in contrast, are continually forced by the winds, which we now represent by including a horizontally uniform forcing term in the modal YBJ equation. If we include sources of NIW energy then we must also include the sinks, such that the wave energy can equilibrate. In the real ocean, NIW energy is primarily dissipated through mixing. Mechanisms of NIW dissipation are complicated and depend, among other factors, on the local stratification (e.g. Kunze *et al.* 1995; Qu, Thomas & Hetland 2021) and the mesoscale eddy field (Sanford, Ma & Alford 2021; Essink *et al.* 2022; Thomas *et al.* 2024*b*). For simplicity, we model these processes as a linear drag. In mixed-layer models, a linear drag is often used to model NIW propagation out of the mixed-layer (see e.g. Pollard & Millard 1970), but in our model vertical propagation is already accounted for by the decomposition into baroclinic modes. The linear drag in our

model therefore represents the irreversible sink of NIW kinetic energy due to mixing. With these alterations, the modal YBJ equation reads

$$\mathrm{d}a_t = \left(-\mathrm{i}\omega a_t - r a_t + F_t \mathrm{e}^{\mathrm{i}ft}\right) \mathrm{d}t, \tag{6.1}$$

where a_t denotes the modal amplitude at time t, r is the linear drag coefficient and F_t the wind forcing projected onto the mode under consideration. We suppress the mode index $\boldsymbol{\mu}$ for now but keep in mind that this equation must be solved for each mode. Note also that we have re-dimensionalised the equation here. The factor of $\mathrm{e}^{\mathrm{i}ft}$ back rotates the forcing to match the back-rotated description of the NIW evolution by the YBJ equation. To proceed, we describe the wind by an Ornstein–Uhlenbeck process that satisfies

$$\mathrm{d}F_t = -c F_t \,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \tag{6.2}$$

where c^{-1} is the decorrelation time scale of the wind forcing, σ is the amplitude of the stochastic excitation and W_t is a Wiener process. The power spectrum of the process F_t is

$$S(\omega) = \frac{2}{\pi}\frac{c}{c^2 + \omega^2}. \tag{6.3}$$

For $\omega \gg c$, the power falls off with frequency as ω^{-2}, i.e. the spectrum is red. We find that this is a good model of the power spectrum of the wind stress from reanalysis, especially over the ocean (see Appendix E for more details).

We consider the system spun up from $t = -\infty$, such that it has statistically equilibrated for all t. This results in the formal solution for the forcing

$$F_t = \sigma \int_{-\infty}^{t} \mathrm{e}^{-c(t-t')} \,\mathrm{d}W_{t'}, \tag{6.4}$$

and the formal solution for the mode amplitude a is given by

$$a_t = \mathrm{e}^{-(\mathrm{i}\omega + r)t} \int_{-\infty}^{t} \mathrm{e}^{[\mathrm{i}(f+\omega)+r]t'} F_{t'} \,\mathrm{d}t'. \tag{6.5}$$

The NIW kinetic energy equation can be obtained in the usual way by multiplying (6.1) with a_t^* and adding the complex conjugate. This is allowed because it is the integral of a Wiener process that appears in (6.1) and not the Wiener process itself. The wind work Γ_t arises as

$$\Gamma_t = \tfrac{1}{2}\left(a_t^* \mathrm{e}^{\mathrm{i}ft} F_t + \text{c.c.}\right). \tag{6.6}$$

We are interested in the average of Γ_t over an ensemble of many realisations of the wind forcing. Let $\langle \cdot \rangle$ denote such an ensemble average. Hence, the ensemble average wind work is

$$\langle \Gamma_t \rangle = \tfrac{1}{2}\left(\mathrm{e}^{[\mathrm{i}(f+\omega)-r]t} \int_{-\infty}^{t} \mathrm{e}^{[-\mathrm{i}(f+\omega)+r]t'} \langle F_{t'}^* F_t \rangle \,\mathrm{d}t' + \text{c.c.}\right). \tag{6.7}$$

The covariance function of the Ornstein–Uhlenbeck process F_t is

$$\langle F_{t'}^* F_t \rangle = \frac{\sigma^2}{2c}\mathrm{e}^{-c|t-t'|}, \tag{6.8}$$

so the ensemble average of Γ_t reduces to

$$\langle \Gamma_t \rangle = \frac{\sigma^2}{2c}\frac{c+r}{(c+r)^2 + (f+\omega)^2}. \tag{6.9}$$

As expected, given the initialisation at $t = -\infty$, the power input is independent of time t. This equilibrated wind work is balanced by the linear drag, such that the ensemble

1002 A22-24

averaged kinetic energy in a given mode is finite. From this expression, we can furthermore see that $\langle \Gamma_t \rangle$ is smaller for $\omega > 0$ than for $\omega < 0$. This is because the wind forcing has more power at low frequencies.

We now define Q as the ratio between the equilibrium wind work in the presence of a mesoscale eddy field to the equivalent wind work in the absence of mesoscale eddies. Without mesoscale eddies, $\psi = 0$ and there are no frequency shifts, so $\omega = 0$ for the uniform mode excited by the wind. The wind work is simply

$$\langle \Gamma_t \rangle = \frac{\sigma^2}{2c} \frac{c + r}{(c + r)^2 + f^2}. \tag{6.10}$$

We calculate Q as a weighted sum of the ratio over individual modes, where the weighting is given by the projection $F_{\boldsymbol{\mu}}$ of the forcing onto a given mode $\boldsymbol{\mu}$:

$$Q = \sum_{\boldsymbol{\mu}} |F_{\boldsymbol{\mu}}|^2 \frac{(c + r)^2 + f^2}{(c + r)^2 + (f + \omega_{\boldsymbol{\mu}})^2}. \tag{6.11}$$

Here we restored the subscripts for the modes. This expression depends on the wave dispersiveness ε^2 through $F_{\boldsymbol{\mu}}$ and $\omega_{\boldsymbol{\mu}}$. If we make the assumption that the $r \ll c$, meaning that the time scale of NIW dissipation is much longer than the memory of the winds, this expression reduces to

$$Q = \sum_{\boldsymbol{\mu}} |F_{\boldsymbol{\mu}}|^2 \frac{c^2 + f^2}{c^2 + (f + \omega_{\boldsymbol{\mu}})^2}. \tag{6.12}$$

We use this reduced expression in the following analysis because $r \ll c$ appears reasonable and because r would be difficult to estimate.

Modulation of the NIW wind work by mesoscale eddies occurs only for $\varepsilon \lesssim 1$. Using the dipole flow as an example, we calculate Q from (6.12) as a function of c and ε (figure 11). For large ε, Q quickly approaches unity, regardless of the value of c. For small ε, the contours of Q become horizontal and there is little dependence of Q on ε. The dependence is primarily on c with a lower value of c resulting in a higher value of Q, i.e. a more substantial enhancement of the wind work.

Our framework provides physical motivation for why mesoscale eddies can modulate the wind work in the weak-dispersion case. Assuming $c \ll f$, which is generally the case for the wind stress over the ocean, we see that the inertial frequency f is in the ω^{-2} part of the wind power spectrum. Any process that shifts the frequency of NIWs will modulate the wind power felt by the waves. Because the wind power spectrum falls off like ω^{-2}, a shift to lower frequencies will raise the wind power felt by the waves and a shift to higher frequencies will lower it. This is the essence of (6.12). As we have shown above, frequency shifts are small in the strong-dispersion limit, and so the waves should feel similar wind power regardless of the presence of mesoscale eddies. As such, Q is close to unity in the strong-dispersion limit. In the weak-dispersion limit, in contrast, there can be significant frequency shifts. A uniform forcing will project onto many modes with a range of frequency shifts. Due to the curvature of the wind power spectrum, going like ω^{-2}, the fractional increase in power for negative frequency shifts will be greater than the fractional decrease in power for positive frequency shifts of the same magnitude. As a result, there will be a net increase in NIW wind work when summing over all modes (see figure 11(*b*) for a schematic). The question remains whether this will be an appreciable effect in the ocean.

(*a*)

$$Q$$

1.000   1.025   1.050

(*b*)

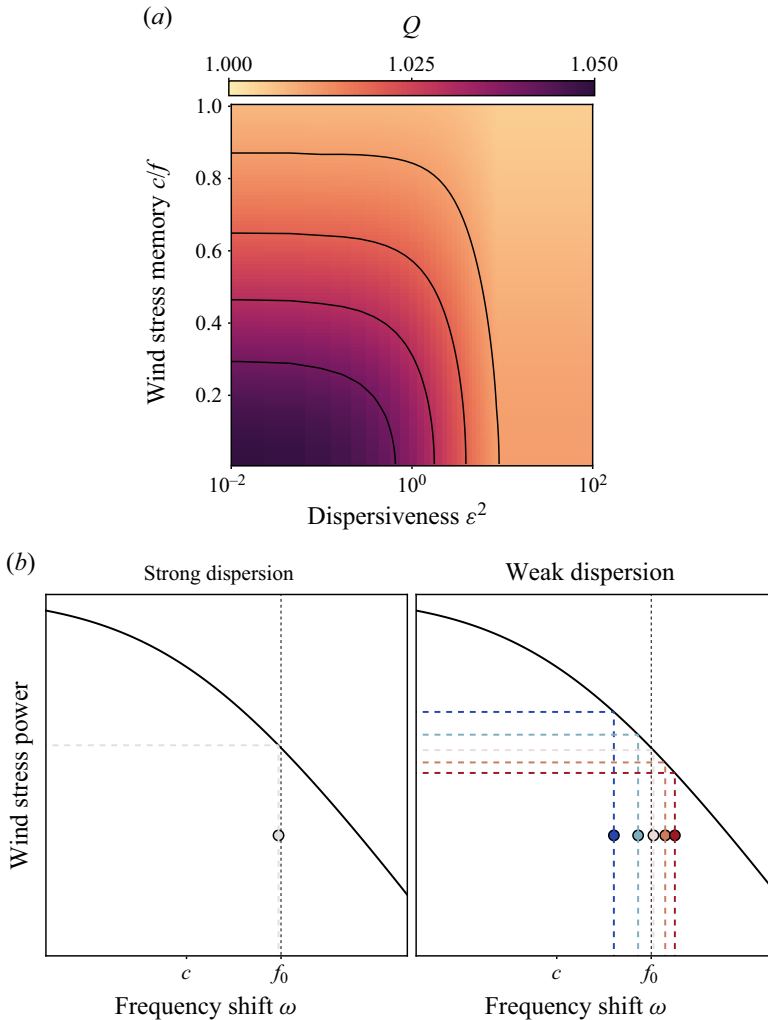Strong dispersion   Weak dispersion

Figure 11. (*a*) Ratio Q of NIW wind work in the presence of mesoscale eddies to that without as a function of the wave dispersiveness ε^2 and the wind stress memory parameter c/f. Modulation of the NIW wind work by the mesoscale eddy field appears only for low ε and c/f. These values correspond to a re-dimensionalisation of ω by $Ro = 0.5$. (*b*) Schematic illustrating the enhancement of NIW wind work in the weak-dispersion regime. The solid black curve illustrates the wind stress power as a function of frequency on a log-log plot. At the inertial frequency, the power is rapidly falling off. The circles indicate the modes that a uniform initial condition projects onto. In the strong-dispersion case, the forcing projects onto a single mode with a small frequency shift. In the weak-dispersion case, the forcing projects onto a wide variety of modes with large frequency shifts.

We estimate Q from observations. For each location in the ocean, we estimate ε from the deformation radius and satellite altimetry observations of the eddy field (see Appendix A), and we estimate c from atmospheric reanalysis (see Appendix E). We calculate the modes of the dipole flow for a range of ε, which gives us ω_{μ} and $|F_{\mu}|^2$, and we re-dimensionalise ω_{μ} using the Rossby number $Ro = \zeta/f$ calculated from satellite altimetry. We use the spatial structure of the vortex dipole as a stand-in for the real eddy structure. This provides an (admittedly crude) estimate of the combined effect of an anticyclone and a cyclone. We calculate Q by using (6.12) and then interpolating onto the correct ε.

Our estimate reveals that deviations of Q from unity are weak, at most 5%. This effect is entirely concentrated in the western boundary current regions. This is because the dimensional frequency shift scales with Ro. Over most of the ocean Ro is far too weak to produce any modulation of the NIW wind work. While this mechanism may be important for individual NIW events (Conn *et al.* 2024), it is clear that on average there is not a significant modulation of the NIW wind work by mesoscale eddies. The maximum modulation of 5% is significantly smaller than current uncertainties in the NIW wind work (Alford 2020). That being said, our approximation of the wind stress as an Ornstein–Uhlenbeck process is highly simplified. The real forcing is dominated by intermittent atmospheric cyclones. The linear drag is also an extremely crude representation of NIW dissipation and should be thought of as nothing more than a stand-in for a more realistic representation.

7. Limitations of the model

All the limitations of the YBJ model are inherent in our analysis above. Specifically, there is no feedback of the waves onto the background flow. Xie & Vanneste (2015) extended the YBJ model by coupling it to a quasi-geostrophic model for the background flow that included wave feedbacks. These wave feedbacks can significantly alter the characteristics of the background flow (Xie & Vanneste 2015; Wagner & Young 2016; Rocha *et al.* 2018). Kafiabad *et al.* (2021) showed that this wave feedback can cause frequency shifts in the NIWs. Furthermore, Kafiabad *et al.* (2021) also noticed that strong wave feedbacks can generate instabilities that cause small-scale structure in the vorticity field.

The scaling assumptions of the original YBJ equation should be kept in mind in the context of the asymptotic expansions performed above. The Rossby number is $Ro = \Psi/fL^2$ and the Burger number is $Bu = \lambda^2/L^2$, such that $\varepsilon^2 = Bu/Ro$. The YBJ equation arises asymptotically in the limit where $Bu \to 0$ while Bu/Ro is kept fixed (see Asselin & Young 2019), so $\varepsilon \ll 1$ and $\varepsilon \gg 1$ do not violate the YBJ scaling.

Thomas *et al.* (2017) conducted a detailed study of the evolution of NIWs in different scaling regimes. They considered a 'very weak-dispersion regime' where $Bu \sim Ro^2$ that is equivalent to $\varepsilon^2 \sim Ro$. An additional term arises compared with the YBJ equation, but they found the YBJ equation to still work well in simulations. They also considered a 'strong-dispersion regime' where $Bu \sim 1$. In this regime they found a leading-order uniform NIW solution, but also with the excitation of super-inertial frequencies that are not captured by YBJ. The frequency shift of the uniform mode is as predicted by YBJ. Another way to improve on the YBJ model is the YBJ$^+$ scheme of Asselin & Young (2019). It has a dispersion relation that remains accurate over a wider range of Bu, and it has desirable numerical properties. The YBJ$^+$ equation is not a Schrödinger equation any more, however, and we do not pursue its analysis here.

Throughout this paper we have dealt with the case in which the background flow does not evolve. In the ray tracing framework, the background flow could be allowed to evolve. The Hamiltonian operator would be time dependent, but the equations can still be integrated along rays. For our analysis of eigenmodes to be applicable to the time-dependent case, the evolution of the background flow should be adiabatic, i.e. it should be slow compared with the wave evolution. The time for eigenmodes to dephase depends on the difference between their frequencies. In the strong-dispersion case the frequency difference between the leading-order eigenmode and the higher eigenmodes is $O(\varepsilon^2)$, meaning that the time to dephase should be small relative to the time scale for evolution of the background flow. In the weak-dispersion limit the eigenvalues become

ever-closely packed, meaning the time scale for dephasing can become large. For the adiabatic assumption to hold, an invariant torus should deform much more slowly than the time it takes a particle to traverse the torus. If the time taken to traverse the torus is given by the advective time scale, then these two time scales are formally the same order. The adiabatic assumption will only hold if there is a symmetry that causes the torus to persist for a longer time scale. The dipole vortex is an extreme example of this where the tori never deform, yet the advective time scale is finite. In the ocean, eddies often persist as coherent features for times much longer than the advective time scale. As such, we expect that the weak-dispersion results to continue to provide insight even in the time-dependent case.

We have also assumed that the background flow is barotropic. This allows the YBJ equation to be expanded into the baroclinic normal modes. If the background flow is baroclinic, such a decomposition is not possible and the modes become coupled. This coupling of the modes means that the YBJ equation no longer reduces to the Schrödinger equation. This does not necessarily destroy the quantum analogy, as the techniques employed here may still be applicable with some modifications. We leave an exploration of these issues to future work.

Finally, we note that in the real ocean, vorticity variance increases at smaller scales. One may worry that the frequency shifts would diverge with increasing resolution. While a detailed discussion of this issue is beyond the scope of this paper, dispersion should have a regularising effect on small-scale vorticity, which leaves the problem well-posed.

8. Conclusions

In the YBJ framework the evolution of NIWs in the presence of a mesoscale eddy field is governed by the wave dispersiveness $\varepsilon^2 = f\lambda^2/\Psi$. The limit of $\varepsilon \gg 1$ corresponds to the strong-dispersion limit and $\varepsilon \ll 1$ corresponds to the weak-dispersion limit. Both of these limits are relevant for the ocean, as the wave dispersiveness decreases with vertical mode number and the strength of mesoscale eddies.

The YBJ equation is a Schrödinger equation, with the YBJ operator playing the role of the Hamiltonian operator in quantum mechanics. As is conventional in quantum mechanics, the evolution of NIWs can be described using the eigenmodes of the YBJ operator and their eigenvalues, which determine the frequency shift away from the inertial frequency. Perturbation methods from quantum mechanics yield insight into YBJ dynamics and its relationship to the ray tracing equations of Kunze (1985).

In the strong-dispersion regime $\varepsilon \gg 1$, perturbation theory yields closed-form expressions for the NIW modes. To leading order, a spatially uniform forcing excites a spatially uniform NIW mode. This mode is modulated by an order ε^{-2} perturbation proportional to the streamfunction of the eddy field. The frequency shift is also of order ε^{-2} and proportional to the average kinetic energy of the eddies. Both of these results recover predictions from Young & Ben Jelloul (1997) through an alternative approach. The same approach also yields expressions for the modes that are not spatially uniform to leading order. The degeneracy of these modes at leading order is lifted at higher order, and the frequency shifts and spatial structures can be determined. Wind patterns associated with sharp atmospheric fronts may excite these modes more strongly than the uniform forcing assumed throughout this work (e.g. Thomas 2017).

In the weak-dispersion regime $\varepsilon \ll 1$, the YBJ equation is amenable to WKB analysis. In simple (separable) background flow geometries, this allows the straightforward calculation of eigenmodes and their frequency shifts, which are excellent approximations

of the exact frequency shifts even for modestly small ε. More generally, the weak-dispersion limit of the YBJ equation corresponds to the classical limit of quantum mechanics. The YBJ equation reduces to the ray equations of Kunze (1985), the equivalent to the corresponding classical Hamiltonian dynamics. The semi-classical EBK analysis allows the calculation of frequency shift for non-separable background flows for the regular part of the spectrum, which again are in excellent agreement with the full shifts. The emergence of the ray equations in the classical limit furthermore suggests that they can be applied if dispersion is weak, whether or not the forcing has a large horizontal scale. The spatial scale separation underlying the ray equations emerges because a uniform initial condition projects onto many modes, and these modes exhibit small-scale structure.

The frequency shift of NIW modes away from the inertial frequency implies that the NIW wind work can be modulated by mesoscale eddies. We quantify this using Q that measures the ratio of the NIW wind work in the presence of mesoscale eddies to that without mesoscale eddies. This modulation arises due to the curvature of the wind power spectrum, which enhances the power input into modes with a shift to lower frequencies more than it suppresses the power input into modes with a shift to higher frequencies. On average, this effect is weak in the ocean, however, with the modulations always being less than 5 %.

Author ORCIDs.

Scott Conn https://orcid.org/0000-0003-1404-0196;

Jörn Callies https://orcid.org/0000-0002-2278-2811;

Albion Lawrence https://orcid.org/0000-0003-4116-045X.

Appendix A. Calculating the wave dispersiveness

Here, we describe the calculations used to estimate the wave dispersiveness $\varepsilon^2 = f\lambda^2/\Psi$ from observations. At each location, we estimate the set of deformation radii λ from hydrography and the characteristic strength of the streamfunction Ψ from altimetry.

Following Smith (2007), we calculate λ by solving the baroclinic eigenvalue equation using finite differences. We perform this calculation using the climatology from the estimating the circulation and climate of the ocean (ECCO) state estimate version 4 release 4 (Forget *et al.* 2015; Fukumori *et al.* 2020). We solve the baroclinic eigenvalue equation at each horizontal grid cell on the ECCO grid to obtain maps of the deformation radii. We display ε^2 for the lowest four baroclinic modes only, for which the numerical approximation has a minimal effect.

To calculate Ψ, we use sea surface height (SSH) observations from the data unification and altimeter combination system's delayed-time 2018 release Taburet *et al.* (2019). The SSH is provided at a (nominal) $(1/4)^\circ$ and daily resolution. We calculate a geostrophic streamfunction using $\psi = g\eta/f$, where η is the SSH and f is the (now latitude-dependent) Coriolis frequency. We take observations from 2007 to 2022 and estimate Ψ as the

root mean square ψ over that period. Again, we are assuming that the streamfunction is barotropic.

Appendix B. Numerical solutions to the eigenvalue problem

To numerically solve the eigenvalue problem (2.4), we discretise the Hamiltonian operator H using finite differences. We use a fourth-order central difference scheme for the Laplacian operator in the dispersion term. For the advection term, we employ the fourth-order enstrophy-conserving scheme of Arakawa (1966), which preserves the Hermitian nature of the operator and translates into energy conservation in this context. In the notation of Arakawa (1966), we employ $2J_1 - J_2$, where $J_1 = \frac{1}{3}(J^{++} + J^{+\times} + J^{\times+})$ and $J_2 = \frac{1}{3}(J^{\times\times} + J^{\times+} + J^{+\times})$. For the refraction term, we evaluate ζ analytically at each point, although in general it could be calculated from the streamfunction using finite differences as well.

We use a spatial resolution of up to 1024×1024 points and solve for the lowest eigenvalues using Lanczos iteration. The resolution is chosen by checking the convergence of the eigenvalues. The number of eigenvalues solved for depends on the value of ε, which controls how densely packed the eigenvalues are and, thus, how many must be computed to find all eigenmodes that a uniform forcing projects onto substantially. We ensure a large enough number of eigenvalues are computed by summing the square of the projection coefficients.

Appendix C. Analytical solutions to shear flow WKB integrals

Here, we provide analytical solutions to the WKB problem with the sinusoidal shear flow. First, we rewrite the potential as

$$V(x) = A_m \cos{(x + \delta_m)} + \frac{\varepsilon^2 m^2}{2}, \quad \text{where} \quad A_m = \sqrt{m^2 + \frac{1}{4}} \quad \text{and} \quad \tan \delta_m = -2m.$$
$$(C1a,b)$$

If $m > 0$ and arctan corresponds to the principal value, then it follows that $\delta_m = \pi + \arctan{(-2m)}$. Because the domain is periodic, we can consider any interval of length 2π. For convenience, we choose $[-\pi - \arctan(-2m), \pi - \arctan(-2m)]$. We can now make the change of variable $x' = x + \arctan(-2m)$. The transformed potential is

$$V(x') = \frac{\varepsilon^2 m^2}{2} - A_m \cos(x').$$
$$(C2)$$

With the potential in this form, the WKB integral (4.6) can be evaluated in terms of the elliptic integral of the second kind $E(\varphi|k^2)$:

$$S_0 = \pm 2\sqrt{2} \mathrm{i} \varepsilon \sqrt{\omega - \frac{\varepsilon^2 m^2}{2} + A_m} \, E\left(\frac{x'}{2} \, \middle| \, \frac{2A_m}{\omega - \frac{\varepsilon^2 m^2}{2} + A_m}\right).$$
$$(C3)$$

We can obtain an equation for the eigenvalues from (4.13). Letting x_1' denote the positive turning point given by

$$x_1' = \pi - \arccos\left(\frac{\omega - \dfrac{\varepsilon^2 m^2}{2}}{A_m}\right), \tag{C4}$$

we obtain

$$E(\varphi|k^2) = \frac{\varepsilon\pi\left(n + \dfrac{1}{2}\right)}{4\sqrt{2\left(\omega - \dfrac{\varepsilon^2 m^2}{2} + A_m\right)}}, \quad \text{where } \varphi = \frac{x_1'}{2} \quad \text{and} \quad k^2 = \frac{2A_m}{\omega - \varepsilon^2 m^2/2 + A_m}. \tag{C5a,b}$$

This is a transcendental equation that can be solved numerically for the eigenvalues ω.

The eigenvectors can be normalised by requiring

$$\int_{-\pi}^{\pi} [\hat{\phi}_n(x)]^2 \, dx = 2\pi. \tag{C6}$$

Letting C be the normalisation constant, we obtain

$$\frac{8C^2}{\sqrt{\omega - \dfrac{\varepsilon^2 m^2}{2} + A_m}} F(\varphi|k^2) = 2\pi, \tag{C7}$$

where $F(\varphi|k^2)$ is the elliptic integral of the first kind (see Bender & Orszag 1999).

The projection of a uniform forcing onto a given mode with even symmetry about the bottom of the potential is

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\phi}_n(x) \, dx. \tag{C8}$$

This integral can be evaluated (again see Bender & Orszag 1999) as

$$|a_n|^2 = \frac{1}{\sqrt{A_m - \omega + \dfrac{\varepsilon^2 m^2}{2}}} \frac{\varepsilon}{\sqrt{2}F(\varphi|k^2)}. \tag{C9}$$

Appendix D. Further details about the EBK method

We principally follow Percival & Pomphrey (1976) to calculate the invariant tori satisfying quantisation conditions and the associated EBK predictions for the eigenvalues ω. We

write the angle Hamilton equations, which are partial differential equations that describe the invariant torus:

$$\nu \frac{\partial x}{\partial \theta} + \mu \frac{\partial x}{\partial \varphi} = \varepsilon^2 k + u, \tag{D1}$$

$$\nu \frac{\partial y}{\partial \theta} + \mu \frac{\partial y}{\partial \varphi} = \varepsilon^2 l + v, \tag{D2}$$

$$\nu \frac{\partial k}{\partial \theta} + \mu \frac{\partial k}{\partial \varphi} = -\left(k \frac{\partial u}{\partial x} + l \frac{\partial v}{\partial x} + \frac{1}{2} \frac{\partial \zeta}{\partial x} \right), \tag{D3}$$

$$\nu \frac{\partial l}{\partial \theta} + \mu \frac{\partial l}{\partial \varphi} = -\left(k \frac{\partial u}{\partial y} + l \frac{\partial v}{\partial y} + \frac{1}{2} \frac{\partial \zeta}{\partial y} \right). \tag{D4}$$

The quantisation conditions can then be written as integrals over the angles θ and φ:

$$\int \left(k \frac{\partial x}{\partial \theta} + l \frac{\partial y}{\partial \theta} \right) \mathrm{d}\theta = 2\pi m, \quad \int \left(k \frac{\partial x}{\partial \varphi} + l \frac{\partial y}{\partial \varphi} \right) \mathrm{d}\varphi = 2\pi (2n + 1 + m). \tag{D5a,b}$$

The integration along θ passes around the hole of the torus (like contour $\mathcal{C}_2$ in figure 9). The integration along φ passes through the hole twice and also around the hole once, so we double the radial phase increment $2\pi(n + \frac{1}{2})$ and add the azimuthal phase increment $2\pi m$ in the second quantisation condition. We average these numerical integrals over the respective other coordinate to increase the accuracy.

We discretise the above equations by dividing the $[0, 2\pi]$ intervals that the angles θ and φ vary over using 64 points and approximate derivatives using an eighth-order finite difference scheme. We initialise the calculation with $\theta = -\frac{1}{2}$, $\varphi = \frac{1}{10}$,

$$\left. \begin{aligned} x &= \tfrac{1}{2}(1 + \cos \varphi) \cos \theta, \quad y = \tfrac{1}{2}(1 + \cos \varphi) \sin \theta, \quad k = -\varepsilon^{-1} \sin \varphi \cos \theta, \\ l &= -\varepsilon^{-1} \sin \varphi \sin \theta, \end{aligned} \right\} \tag{D6a–d}$$

for the $(n, m) = (0, 0)$ torus and apply Newton iteration to satisfy the above equations. We determine ω by applying the dispersion relation at each point of the torus and averaging over all grid points. We then change the quantum numbers to other values and start the Newton iteration from the previously found torus, using iterations at intermediate values if needed.

Appendix E. Estimating the decorrelation time of wind stress

Here, we describe the calculations used to estimate the decorrelation time c^{-1} of the wind stress. For the wind forcing, we use the European Centre for Medium-Range Weather Forecasting ERA-5 reanalysis (Hersbach *et al.* 2020). For the calculations below, we use data from 2015 to 2020. At each grid cell, we use the 10 m zonal (u_w) and meridional (v_w) winds with hourly resolution. Following Pollard & Millard (1970) we convert this to a wind stress using a bulk aerodynamic drag formulation. The time series at each location

is used to calculate a power spectrum of the wind stress. The decorrelation time scale is obtained by fitting the following model to the estimated spectrum:

$$S(\omega) = \frac{A}{1 + \left(\dfrac{\omega}{c}\right)^s}. \tag{E1}$$

Here A and s are additional fitted parameters that we do not use here. Over the ocean $s = 2$ is a reasonable approximation, which motivates our use of the Ornstein–Uhlenbeck process above.

<div align="center">REFERENCES</div>

ALFORD, M.H. 2020 Revisiting near-inertial wind work: slab models, relative stress, and mixed layer deepening. *J. Phys. Oceanogr.* **50** (11), 3141–3156.

ALFORD, M.H., MACKINNON, J.A., SIMMONS, H.L. & NASH, J.D. 2016 Near-inertial internal gravity waves in the ocean. *Annu. Rev. Mar. Sci.* **8**, 95–123.

ANDERSON, D.L.T. & GILL, A.E. 1979 Beta dispersion of inertial waves. *J. Geophys. Res.: Oceans* **84** (C4), 1836–1842.

ARAKAWA, A. 1966 Computational design for long term integration of the equations of motion: two-dimensional incompressible flow. Part I. *J. Comput. Phys* **1** (1), 9–143.

ASSELIN, O., THOMAS, L.N., YOUNG, W.R. & RAINVILLE, L. 2020 Refraction and straining of near-inertial waves by barotropic eddies. *J. Phys. Oceanogr.* **50** (12), 3439–3454.

ASSELIN, O. & YOUNG, W.R. 2019 An improved model of near-inertial wave dynamics. *J. Fluid Mech.* **876**, 428–448.

ASSELIN, O. & YOUNG, W.R. 2020 Penetration of wind-generated near-inertial waves into a turbulent ocean. *J. Phys. Oceanogr.* **50** (6), 1699–1716.

BALMFORTH, N.J., LLEWELLYN SMITH, S.G. & YOUNG, W.R. 1998 Enhanced dispersion of near-inertial waves in an idealized geostrophic flow. *J. Mar. Res.* **56** (1), 1–40.

BALMFORTH, N.J. & YOUNG, W.R. 1999 Radiative damping of near-inertial oscillations in the mixed layer. *J. Mar. Res.* **57** (4), 561–584.

BENDER, C.M. & ORSZAG, S.A. 1999 *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*, chap. 10, Springer Science & Business Media.

BERRY, M.V. & MOUNT, K.E. 1972 Semiclassical approximations in wave mechanics. *Rep. Prog. Phys.* **35** (1), 315.

BERRY, M.V. & OZORIO DE ALMEIDA, A.M. 1973 Semiclassical approximation of the radial equation with two-dimensional potentials. *J. Phys.* A **6** (10), 1451.

BRILLOUIN, L. 1926 Remarques sur la mécanique ondulatoire. *J. Phys.* **7** (12), 353–368.

BÜHLER, O. 2006 *A Brief Introduction to Classical, Statistical, and Quantum Mechanics*. American Mathematical Society.

BYUN, S.-S., PARK, J.J., CHANG, K.-I. & SCHMITT, R.W. 2010 Observation of near-inertial wave reflections within the thermostad layer of an anticyclonic mesoscale eddy. *Geophys. Res. Lett.* **37** (1).

CHEN, G., XUE, H., WANG, D. & XIE, Q. 2013 Observed near-inertial kinetic energy in the northwestern south China Sea. *J. Geophys. Res.: Oceans* **118** (10), 4965–4977.

CONN, S., FITZGERALD, J. & CALLIES, J. 2024 Interpreting observed interactions between near-inertial waves and mesoscale eddies. *J. Phys. Oceanogr.* **54** (2), 485–502.

DANIOUX, E., VANNESTE, J. & BÜHLER, O. 2015 On the concentration of near-inertial waves in anticyclones. *J. Fluid Mech.* **773**, R2.

D'ASARO, E.A. 1985 The energy flux from the wind to near-inertial motions in the surface mixed layer. *J. Phys. Oceanogr.* **15** (8), 1043–1059.

D'ASARO, E.A. 1995 Upper-ocean inertial currents forced by a strong storm. Part III: interaction of inertial currents and mesoscale eddies. *J. Phys. Oceanogr.* **25** (11), 2953–2958.

D'ASARO, E.A., ERIKSEN, C.C., LEVINE, M.D., PAULSON, C.A., NIILER, P. & VAN MEURS, P. 1995 Upper-ocean inertial currents forced by a strong storm. Part I: data and comparisons with linear theory. *J. Phys. Oceanogr.* **25** (11), 2909–2936.

EDELMAN, A. & RAO, N.R. 2005 Random matrix theory. *Acta Numerica* **14**, 233–297.

EDELMAN, A. & SUTTON, B.D. 2007 From random matrices to stochastic operators. *J. Stat. Phys.* **127**, 1121–1165.

EINSTEIN, A. 1917 Zum Quantensatz von Sommerfeld und Epstein. *Ver. Deut. Phys. Gesell.* **19**, 82–92.

ESSINK, S., KUNZE, E., LIEN, R.-C., INOUE, R. & ITO, S.-I. 2022 Near-inertial wave interactions and turbulence production in a Kuroshio anticyclonic eddy. *J. Phys. Oceanogr.* **52** (11), 2687–2704.

FERRARI, R. & WUNSCH, C. 2009 Ocean circulation kinetic energy: reservoirs, sources, and sinks. *Annu. Rev. Fluid Mech.* **41**, 253–282.

FORGET, G.A.E.L., CAMPIN, J.-M., HEIMBACH, P., HILL, C.N., PONTE, R.M. & WUNSCH, C. 2015 ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geosci. Model Develop.* **8** (10), 3071–3104.

FUKUMORI, I., WANG, O., FENTY, I., FORGET, G., HEIMBACH, P. & PONTE, R.M. 2020 Synopsis of the ECCO central production global ocean and sea-ice state estimate, version 4 release 4. Zenodo. https://doi.org/10.5281/zenodo.4533349.

GARRETT, C. 2001 What is the "near-inertial" band and why is it different from the rest of the internal wave spectrum? *J. Phys. Oceanogr.* **31** (4), 962–971.

GILL, A.E. 1984 On the behavior of internal waves in the wakes of storms. *J. Phys. Oceanogr.* **14** (7), 1129–1151.

GUTZWILLER, M.C. 1992 Quantum chaos. *Sci. Am.* **266** (1), 78–85.

HERSBACH, H., BELL, B., BERRISFORD, P., HIRAHARA, S., HORÁNYI, A., MUÑOZ SABATER, J., NICOLAS, J., PEUBEY, C., RADU, R. & SCHEPERS, D. 2020 The ERA5 global reanalysis. *Q. J. R. Meteorol Soc.* **146**, 1999–2049.

JAIMES, B. & SHAY, L.K. 2010 Near-inertial wave wake of hurricanes Katrina and Rita over mesoscale oceanic eddies. *J. Phys. Oceanogr.* **40** (6), 1320–1337.

JOCHUM, M., BRIEGLEB, B.P., DANABASOGLU, G., LARGE, W.G., NORTON, N.J., JAYNE, S.R., ALFORD, M.H. & BRYAN, F.O. 2013 The impact of oceanic near-inertial waves on climate. *J. Clim.* **26** (9), 2833–2844.

KAFIABAD, H.A., VANNESTE, J. & YOUNG, W.R. 2021 Interaction of near-inertial waves with an anticyclonic vortex. *J. Phys. Oceanogr.* **51** (6), 2035–2048.

KELLER, J.B. 1958 Corrected Bohr–Sommerfeld quantum conditions for nonseparable systems. *Ann. Phys.* **4** (2), 180–188.

KELLER, J.B. 1985 Semiclassical mechanics. *SIAM Rev.* **27** (4), 485–504.

KUNZE, E. 1985 Near-inertial wave propagation in geostrophic shear. *J. Phys. Oceanogr.* **15** (5), 544–565.

KUNZE, E. & SANFORD, T.B. 1984 Observations of near-inertial waves in a front. *J. Phys. Oceanogr.* **14** (3), 566–581.

KUNZE, E., SCHMITT, R.W. & TOOLE, J.M. 1995 The energy balance in a warm-core ring's near-inertial critical layer. *J. Phys. Oceanogr.* **25** (5), 942–957.

LLEWELLYN SMITH, S.G. 1999 Near-inertial oscillations of a barotropic vortex: trapped modes and time evolution. *J. Phys. Oceanogr.* **29** (4), 747–761.

MASLOV, V.P. 1972 *Théorie des Perturbations et Méthodes Asymptotiques*. Dunod.

MUNK, W. & WUNSCH, C. 1998 Abyssal recipes II: energetics of tidal and wind mixing. *Deep-Sea Res. (*I*)* **45** (12), 1977–2010.

OEY, L.-Y., INOUE, M., LAI, R., LIN, X.-H., WELSH, S.E. & ROUSE, L.J.J. 2008 Stalling of near-inertial waves in a cyclone. *Geophys. Res. Lett.* **35** (12).

PERCIVAL, I.C. 1977 Semiclassical theory of bound states. *Adv. Chem. Phys.* **36**, 1–61.

PERCIVAL, I.C. & POMPHREY, N. 1976 Vibrational quantization of polyatomic molecules. *Mol. Phys.* **31** (1), 97–114.

PERKINS, H. 1976 Observed effect of an eddy on inertial oscillations. *Deep-Sea Res. Oceanogr. Abstracts* **23**, 1037–1042.

POLLARD, R.T. & MILLARD, R.C. 1970 Comparison between observed and simulated wind-generated inertial oscillations. *Deep Sea Res. Oceanogr. Abstracts* **17**, 813–821.

QU, L., THOMAS, L.N. & HETLAND, R.D. 2021 Near-inertial-wave critical layers over sloping bathymetry. *J. Phys. Oceanogr.* **51** (6), 1737–1756.

ROCHA, C.B., WAGNER, G.L. & YOUNG, W.R. 2018 Stimulated generation: extraction of energy from balanced flow by near-inertial waves. *J. Fluid Mech.* **847**, 417–451.

SAKURAI, J.J. & NAPOLITANO, J. 2020 *Modern Quantum Mechanics*. Cambridge University Press.

SANFORD, T.B., MA, B.B. & ALFORD, M.H. 2021 Stalling and dissipation of a near-inertial wave (NIW) in an anticyclonic ocean eddy: direct determination of group velocity and comparison with theory. *J. Geophys. Res.: Oceans* **126** (5), e2020JC016742

SMITH, K.S. 2007 The geography of linear baroclinic instability in Earth's oceans. *J. Mar. Res.* **65** (5), 655–683.

STONE, A.D. 2005 Einstein's unknown insight and the problem of quantizing chaos. *Phys. Today* **58** (8), 37.

TABURET, G., SANCHEZ-ROMAN, A., BALLAROTTA, M., PUJOL, M.-I., LEGEAIS, J.-F., FOURNIER, F., FAUGERE, Y. & DIBARBOURE, G. 2019 DUACS DT2018: 25 years of reprocessed sea level altimetry products. *Ocean Sci.* **15** (5), 1207–1224.

THOMAS, J., SMITH, K.S. & BÜHLER, O. 2017 Near-inertial wave dispersion by geostrophic flows. *J. Fluid Mech.* **817**, 406–438.

THOMAS, L., KELLY, S., KLENZ, T., YOUNG, W., RAINVILLE, L., SIMMONS, H., HORMANN, V. & STOKES, I. 2024*a* Why near-inertial waves are less affected by vorticity in the Northeast Pacific than in the North Atlantic. *Oceanography* **37** (4).

THOMAS, L.N. 2017 On the modifications of near-inertial waves at fronts: implications for energy transfer across scales. *Ocean Dyn.* **67** (10), 1335–1350.

THOMAS, L.N., MOUM, J., QU, L., HILDITCH, J., KUNZE, E.L., RAINVILLE, L. & LEE, C. 2024*b* Blocked drainpipes and smoking chimneys: discovery of new near-inertial wave phenomena in anticyclones. *Oceanography* **37** (4).

THOMAS, L.N., RAINVILLE, L., ASSELIN, O., YOUNG, W.R., GIRTON, J., WHALEN, C.B., CENTURIONI, L. & HORMANN, V. 2020 Direct observations of near-inertial wave ζ-refraction in a dipole vortex. *Geophys. Res. Lett.* **47** (21), e2020GL090375.

VOET, G., *et al.* 2024 Near-inertial energy variability in a strong mesoscale eddy field in the Iceland Basin. *Oceanography* **37** (4).

WAGNER, G.L. & YOUNG, W.R. 2016 A three-component model for the coupled evolution of near-inertial waves, quasi-geostrophic flow and the near-inertial second harmonic. *J. Fluid Mech.* **802**, 806–837.

XIE, J.-H. & VANNESTE, J. 2015 A generalised-Lagrangian-mean model of the interactions between near-inertial waves and mean flow. *J. Fluid Mech.* **774**, 143–169.

YOUNG, W.R. & BEN JELLOUL, M. 1997 Propagation of near-inertial oscillations through a geostrophic flow. *J. Mar. Res.* **55**, 735–766.

YU, X., NAVEIRA GARABATO, A.C., VIC, C., GULA, J., SAVAGE, A.C, WANG, J., WATERHOUSE, A.F. & MACKINNON, J.A. 2022 Observed equatorward propagation and chimney effect of near-inertial waves in the midlatitude ocean. *Geophys. Res. Lett.* **49** (13), e2022GL098522.

ZHANG, F. & XIE, J.-H. 2023 Scale dependence of near-inertial wave's concentration in anticyclones. *Front. Mar. Sci.* **10**, 1085679.