

Refraction of near-inertial waves by submesoscale vorticity filaments

James P. Hilditch¹ , John R. Taylor²  and Leif N. Thomas¹ 

¹Department of Earth System Science, Stanford University, Stanford, CA 94305, USA

²Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK

Corresponding author: James P. Hilditch, hilditch@stanford.edu

(Received 13 August 2024; revised 26 August 2025; accepted 26 August 2025)

The interaction of near-inertial waves (NIWs) with submesoscale vorticity filaments is explored using theory and simulations. We study three idealised set-ups representative of submesoscale flows allowing for $O(1)$ or greater Rossby numbers. First, we consider the radiation of NIWs away from a cyclonic filament and develop scalings for the decay of wave energy in the filament. Second, we introduce broad anticyclonic regions that separate the cyclonic filaments mimicking submesoscale eddy fields and analyse the normal modes of this system. Third, we extend this set-up to consider the vertical propagation and the radiation of NIW energy. We identify a key length scale L_m , dependent on the strength of the filament, stratification and vertical scale of the waves, that when compared with the horizontal scales of the background flow determines the NIW behaviour. A generic expression for the vertical group velocity is derived that highlights the importance of horizontal gradients for vertical wave propagation. An overarching theme of the results is that NIW radiation, both horizontally and vertically, is most efficient when L_m is comparable to the length scales of the background flow.

Key words: waves in rotating fluids, internal waves, ocean processes

1. Introduction

Near-inertial waves (NIWs), unbalanced motions with frequencies close to the local Coriolis frequency f , are an important feature of upper ocean dynamics. However, their dynamics are non-trivial as interactions with the balanced flow modify their spatial structure and propagation behaviour. Perhaps the most important of these interactions is ζ -refraction arising from the modification of the local minimum frequency by geostrophic

vorticity $\zeta := \partial V/\partial x - \partial U/\partial y$, where U and V are the horizontal velocities of the geostrophically balanced background flow. Early theoretical work utilising a Wentzel–Kramers–Brillouin–Jeffreys (WKBJ) approximation (Kunze 1985) identified an increase in the local minimum frequency in regions of cyclonic vorticity ($\zeta > 0$, assuming $f > 0$ as we do throughout the paper) and conversely a decrease in regions of anticyclonic vorticity ($\zeta < 0$). As a result of these shifts, there is a tendency for NIW energy to accumulate in anticyclonic regions and it is possible to have trapped subinertial waves confined to anticyclones. Both effects are now well documented by observations (e.g. Martínez-Marrero *et al.* 2019; Thomas *et al.* 2020). However, NIWs often have large horizontal scales, particularly if forced by winds that typically have footprints 1000s of kilometres wide and orders of magnitude larger than typical mesoscale eddies (10–100 s of kilometres). This can place the WKBJ approximation, which requires the waves to vary rapidly on the scale of the background flow, on very weak theoretical footing.

A more robust model was proposed by Young & Ben Jelloul (1997), hereafter the YBJ model, which makes no spatial scale assumptions but rather models the evolution of the NIWs via a multiple scales expansion in time. In the YBJ model the fast dynamics capture the inertial oscillations with the effects of wave dispersion, geostrophic advection and ζ -refraction modifying the NIW amplitude on the slow time scale. The validity of the model only requires that the wave Burger number (see § 2.2) and the Rossby number, ζ/f , of the background flow be small. For many flows these assumptions are met, including the important case of a wind-forced inertial oscillation interacting with a mesoscale eddy. In these situations the YBJ model has had great success in predicting the evolution of the NIWs and in interpreting observational data (Asselin *et al.* 2020; Thomas *et al.* 2020; Conn, Fitzgerald & Callies 2024; Thomas *et al.* 2024).

However, in recent years growing attention has been paid to submesoscale flows, smaller-scale flows characterised by $O(1)$ or greater Rossby numbers (Thomas, Tandon & Mahadevan 2008; McWilliams 2016; Taylor & Thompson 2023). Frontogenetic processes, prevalent at submesoscales, tend to sharpen dense, cyclonic filaments leading to strongly skewed vorticity distributions. For example, Shcherbina *et al.* (2013) studied the statistical distribution of vertical vorticity in the North Atlantic Mode Water region south of the Gulf Stream using parallel transects from two ships and a regional model. They found that the distribution of vertical vorticity was asymmetric with the mode of the distribution near $\zeta = -0.5f$ and a long tail of cyclonic vorticity that extended well past $\zeta = f$. In the most extreme submesoscale environments, such as the northern Gulf of Mexico, a similarly skewed distribution can be found but with vorticity maxima orders of magnitude larger than f (Schlichting *et al.* 2023). Physically, the strong cyclonic vorticity corresponds with thin filaments and submesoscale eddies (see figure 1*b* in Shcherbina *et al.* 2013 or figure 2*f* in Schlichting *et al.* 2023). These highly localised vorticity structures imply a white enstrophy spectrum and a kinetic energy spectrum with a k^{-2} slope as has been observed in the upper ocean (Shcherbina *et al.* 2013; Callies *et al.* 2015). In the northern Gulf of Mexico, the large freshwater influx from the Mississippi–Atchafalaya river system not only sets up the lateral buoyancy gradients driving the submesoscale flows but also results in a very strong vertical density stratification (Zhang, Hetland & Zhang 2014). While many of the examples we use in this paper are motivated by the conditions in the northern Gulf of Mexico, the theory we develop is general and may be applied across a broad range of realistic oceanic conditions.

The interactions of NIWs with sharp vorticity filaments with a large Rossby number warrant theoretical consideration since this combination falls outside the purview of existing theory. In particular, we are interested in how the results of YBJ theory generalise to large Rossby numbers. Furthermore, the spatial distribution of submesoscale vorticity is

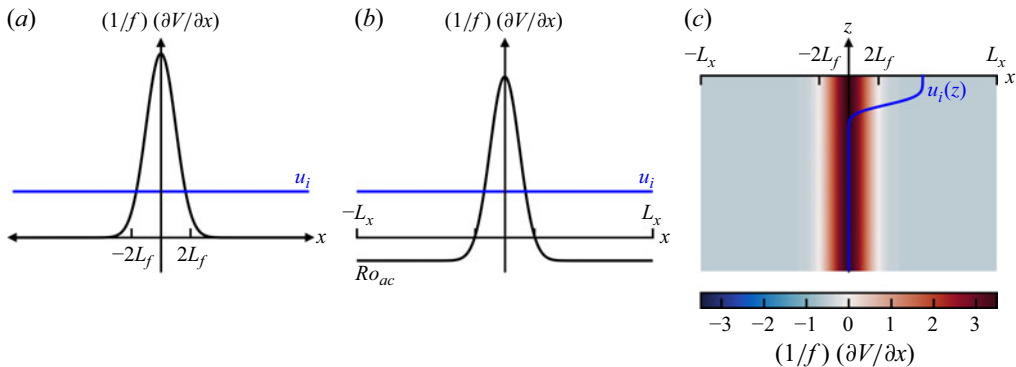


Figure 1. Schematic summarising the three set-ups we consider. (a) A cyclonic vorticity filament (with Gaussian shape) in an unbounded domain with an initially uniform across-filament velocity u_i . (b) The same cyclonic vorticity filament in an otherwise anticyclonic ($Ro_{ac} < 0$) flow in a periodic domain. Here we illustrate the case $\xi := L_x/L_f = 10$. (c) Same background flow as (b), now contoured in blue (anticyclonic) and red (cyclonic), but in two dimensions. The initial across-filament velocity $u_i(z)$ is a horizontally uniform near surface slab.

Physical parameter	Description
f	Coriolis frequency
c_m	Mode speed
ΔV	Filament strength
L_f	Filament width
L_x	Half-width of the anticyclonic region
L_m	Length scale of filament influence, $L_m := c_m^2/f \Delta V$

Table 1. Physical parameters describing the problems we consider. Here L_m is a very important length scale derived from the other parameters.

not well described by a single length scale and this begs the question of which length scales are most important for determining the behaviour of NIWs? To answer these questions, we consider a highly idealised set-up retaining only the key physics necessary to induce ζ -refraction. Although sharp filaments generally form through frontogenesis, we ignore frontal dynamics and consider a barotropic filament. We work in two dimensions x, z neglecting along-filament variations, which not only eliminates geostrophic advection but also precludes barotropic instability. Despite these simplifications the problem remains extremely rich and admits a range of phenomena across a large parameter space.

Our approach is to slowly introduce these phenomena by considering three problems of increasing complexity, summarised by the schematic in figure 1. In this paper we will meet five independent physical parameters that, for reference, are listed in table 1, along with a crucial derived length scale L_m . From these physical parameters we construct many non-dimensional parameters that we also list for reference in table 2. The paper is organised as follows: in § 2 we introduce the central equation of this study, a generalised Klein–Gordon equation; then in § 3 we consider the interaction of a single vertical mode with a cyclonic vorticity filament in an unbounded domain; in § 4 we introduce an additional length scale to the background flow, namely the width of the domain, to model the anticyclonic regions that separate cyclonic filaments; in § 5 we extend the set-up of § 4 to a two-dimensional (2-D) problem and consider multiple vertical modes, vertical propagation and the radiation

Non-dimensional parameter	Physical parameter dependence	Name
Bu_m	$c_m^2/f^2 L_f^2$	Filament Burger number
Ro_f	$\Delta V/f L_f$	Filament Rossby number
γ_m	$c_m^2/f L_f \Delta V \equiv L_m/L_f$	Tunnelling parameter
α_m	$\Delta V/c_m$	
Ro_{ac}	$-\frac{1}{2} \Delta V/f L_x$	Rossby number of the anticyclonic region
ξ	L_x/L_f	
Γ_m	$2c_m^2/f L_x \Delta V \equiv 2L_m/L_x$	

Table 2. Non-dimensional parameters expressed in terms of the physical parameters defining the problems.

of surface intensified NIW energy; finally, we offer conclusions and points of discussion in § 6.

2. Klein–Gordon equation

2.1. Derivation

We define a barotropic background geostrophic velocity $V(x)$ and stratification $N^2(z) := -(g/\rho_0)(\partial \bar{\rho}/\partial z)$, where $\bar{\rho}(z)$ is the background density, ρ_0 is the Boussinesq reference density and g the gravitational acceleration. Focusing on NIWs with small aspect ratios, we linearise the hydrostatic, inviscid, adiabatic Boussinesq equations about this background state and make the simplifying assumption that the dynamics are independent of y . The linearised equations are

$$\frac{\partial u}{\partial t} - f v + \frac{1}{\rho_0} \frac{\partial p}{\partial x} = 0, \quad (2.1a)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial V}{\partial x} + f u = 0, \quad (2.1b)$$

$$\frac{1}{\rho_0} \frac{\partial p}{\partial z} - b = 0, \quad (2.1c)$$

$$\frac{\partial b}{\partial t} + w N^2 = 0, \quad (2.1d)$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \quad (2.1e)$$

where u, v, w are the perturbation velocities, $b := g(\bar{\rho} - \rho)/\rho_0$ is the perturbation buoyancy defined from the density ρ , p is the perturbation pressure and f is the Coriolis frequency, taken to be constant under a traditional f -plane approximation.

By defining a streamfunction ψ such that $u = -\partial \psi/\partial z$, $w = \partial \psi/\partial x$ and systematically eliminating variables, (2.1) can be reduced to a single equation, i.e.

$$\left(\frac{\partial^2}{\partial t^2} + f_{\text{eff}}^2 \right) \frac{\partial^2 \psi}{\partial z^2} + N^2 \frac{\partial^2 \psi}{\partial x^2} = 0, \quad (2.2a)$$

where

$$f_{\text{eff}}^2(x) := f \left(f + \frac{\partial V}{\partial x} \right) \quad (2.2b)$$

is the square of the effective Coriolis frequency. Note that with variations in y neglected, the along-filament momentum equation is a conservation equation for the absolute

momentum $M := fx + V + v$. Furthermore, $f_{\text{eff}}^2(x)$ is proportional to the gradient of the background absolute momentum $V + fx$. Equation (2.2a) is separable and so we write

$$\psi = \sum_m \mathcal{X}_m(x, t) \mathcal{Z}_m(z), \quad (2.3a)$$

with the vertical modes defined by

$$c_m^2 \frac{\partial^2}{\partial z^2} \mathcal{Z}_m = -N^2 \mathcal{Z}_m, \quad (2.3b)$$

where the mode speed c_m is a constant of separation. Taking N^2 to be constant and projecting ψ onto sine modes over a domain of depth L_z , $\mathcal{Z}_m = \sin m\pi z/L_z$, thereby satisfying no-penetration boundary conditions, we have

$$c_m = \frac{NL_z}{m\pi}. \quad (2.4)$$

Alternatively, if the vertical structure is described by a plane wave with wavenumber k_z then $c_m = Nk_z^{-1}$. Before §5, we are not particularly interested in the structure of the vertical modes, only in the value of c_m , which increases with increasing stratification or vertical scale. Note that in these vertical mode decompositions we have filtered out the barotropic mode.

With the vertical structure set, the equation for the horizontal structure is

$$\left[\frac{\partial^2}{\partial t^2} - c_m^2 \frac{\partial^2}{\partial x^2} + f_{\text{eff}}^2(x) \right] \mathcal{X} = 0, \quad (2.5)$$

where we have dropped the subscript m on $\mathcal{X}$ to lighten the notation. This is the Klein–Gordon equation. First derived as a relativistic wave equation, the Klein–Gordon equation can also be interpreted as describing classical waves in an elastic medium where, in this case, the term involving f_{eff} represents elasticity. Both interpretations provide insight into the roles of the background flow and vertical structure of the waves in determining the behaviour of the solutions. For waves with frequencies much greater than f_{eff} , the elastic term is negligible and the Klein–Gordon equation reduces to the wave equation with wave speed c_m . Furthermore, the characteristics of (2.5) have slope c_m (in x, t space) and so c_m is the rate at which the Klein–Gordon equation propagates information.

We solve the Klein–Gordon equation subject to the initial conditions

$$\mathcal{X} = 1, \quad \frac{\partial \mathcal{X}}{\partial t} = 0. \quad (2.6a,b)$$

Such uniform initial conditions are a popular theoretical device (e.g. Balmforth *et al.* 1998; Asselin & Young 2020; Asselin *et al.* 2020; Kafiabad, Vanneste & Young 2021) motivated by the scenario in which a large-scale wind event impulsively excites ageostrophic motions in the upper ocean. Crucially, the spatial structure of the waves develops through interactions with the background flow rather than being prescribed by the initial conditions.

2.2. The YBJ approximation

Young & Ben Jelloul (1997; YBJ) model the evolution of NIWs via a multiple time-scale expansion. They formally justify the expansion by assuming that the Rossby number of the background flow and the wave Burger number $Bu_w := c_m^2/f^2 L_w^2$, where L_w is the horizontal scale of the waves, are small. For the barotropic case with a projection onto

vertical modes, the leading-order horizontal wave velocities are given by

$$u_m + iv_m = -f^2 c_m^2 \mathcal{A} e^{-ift}, \quad (2.7)$$

where the complex amplitude $\mathcal{A}$ varies on a slow time scale (Balmforth *et al.* 1998). Furthermore, the leading-order vertical velocity, buoyancy and pressure may also be expressed in terms of $\mathcal{A}$. The evolution equation for $\mathcal{A}$ is

$$\frac{\partial \mathcal{A}}{\partial t} + \frac{\partial(\Psi, \mathcal{A})}{\partial(x, y)} + i \frac{1}{2} \zeta \mathcal{A} = i \frac{c_m^2}{2f} \nabla^2 \mathcal{A}, \quad (2.8)$$

where Ψ is the geostrophic streamfunction, $\partial(\cdot, \cdot)/\partial(x, y)$ is the Jacobian, $\zeta := \nabla^2 \Psi \equiv \partial V/\partial x - \partial U/\partial y$ is the vertical vorticity of the background flow and $\nabla^2 := \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the horizontal Laplacian. Imposing $\partial/\partial y \equiv 0$, (2.8) reduces to

$$2fi \frac{\partial \mathcal{A}}{\partial t} = -c_m^2 \frac{\partial^2 \mathcal{A}}{\partial x^2} + f \frac{\partial V}{\partial x} \mathcal{A}, \quad (2.9)$$

which is the time-dependent Schrödinger equation in one spatial dimension.

To reduce the Klein–Gordon equation (2.5) to the time-dependent Schrödinger equation (2.9), we simply let

$$\mathcal{X} = \mathcal{A} e^{-ift} + \text{c.c.} \quad (2.10a)$$

such that

$$\frac{\partial^2}{\partial t^2} \mathcal{X} = \left[\left(-f^2 - 2if \frac{\partial}{\partial t} + \frac{\partial^2}{\partial t^2} \right) \mathcal{A} \right] e^{-ift} + \text{c.c.} \quad (2.10b)$$

and neglect the $\partial^2 \mathcal{A}/\partial t^2$ term, consistent with the multiple scales approximation underpinning the YBJ expansion. Thus, the validity of the YBJ approximation in this problem hinges on whether or not we are justified in dropping the $\partial^2 \mathcal{A}/\partial t^2$ term. This approximation has a direct analogue in quantum field theory where it is used to recover the Schrödinger equation from the Klein–Gordon equation in the non-relativistic limit (e.g. Stermann 1993). Throughout the paper we comment on the validity of the YBJ approximation as a function of the parameters describing the different set-ups we consider.

3. Radiation by a filament in an unbounded domain

3.1. Problem set-up

To proceed further we must now specify the background flow $V(x)$. In our first set-up, we consider a background state defined by a velocity scale and a single length scale. That is, a cyclonic filament in an unbounded domain (figure 1a) with geostrophic velocity gradient,

$$\frac{\partial V}{\partial x} = \Delta V \frac{1}{L_f} \mathcal{F} \left(\frac{x}{L_f} \right), \quad (3.1)$$

where L_f is the width of the filament, $\Delta V > 0$ is the change in geostrophic velocity over the filament and $\mathcal{F}(\eta)$ is a positive function of total integral 1. Since the filament is cyclonic, we expect to observe the radiation of NIWs out of the filament analogous to the results of Kafiabad *et al.* (2021) who considered axisymmetric vortices in the low-Rossby-number limit.

The boundary conditions are $\mathcal{X} \rightarrow \cos ft$ as $|x| \rightarrow \infty$. Clearly, the solution is not normalisable, i.e. it has infinite energy, but in this aspect the set-up is similar to a scattering problem in quantum mechanics. Furthermore, we always have the option of reframing our

problem in terms of $\mathcal{Y} := \mathcal{X} - \cos ft$. This converts the set-up into a forced problem with friendlier boundary conditions, $\mathcal{Y} \rightarrow 0$ as $|x| \rightarrow \infty$, and homogeneous initial conditions, $\mathcal{Y} = \partial \mathcal{Y} / \partial t = 0$.

Since we are considering a radiation problem, a quantity of considerable interest is the time scale, T , over which this radiation occurs. To quantify this, we define

$$T_{10} := \min \left\{ t \mid \mathcal{X}(0, t)^2 + f^{-2} \frac{\partial \mathcal{X}}{\partial t}(0, t)^2 < \frac{1}{10} \right\} \quad (3.2a)$$

and the YBJ equivalent

$$T_{10}^{(YBJ)} := \min \left\{ t \mid 4|\mathcal{A}(0, t)|^2 < \frac{1}{10} \right\}. \quad (3.2b)$$

Here T_{10} is the time at which this particular norm of the solution at the centre of the filament first drops to one tenth of its initial value. The factor of 4 in $T_{10}^{(YBJ)}$ comes from the fact that the initial value of $\mathcal{A}$ is $1/2$.

3.2. The parameter space

With background flow (3.1), the problem is defined by four dimensional parameters: the Coriolis frequency f , the mode speed c_m , the filament width L_f and the filament strength ΔV . Therefore, after choosing time and length scales to non-dimensionalise the problem, we are left with a 2-D parameter space. Non-dimensionalising using the available time and length scales,

$$\tilde{t} := ft, \quad \tilde{x} := \frac{x}{L_f}, \quad (3.3a,b)$$

the Klein–Gordon equation (2.5) becomes

$$\left[\frac{\partial^2}{\partial \tilde{t}^2} - Bu_m \frac{\partial^2}{\partial \tilde{x}^2} + 1 + Ro_f \mathcal{F}(\tilde{x}) \right] \mathcal{X} = 0. \quad (3.4)$$

The two parameters appearing here are the filament Burger and Rossby numbers:

$$Bu_m := \frac{c_m^2}{f^2 L_f^2}, \quad Ro_f := \frac{\Delta V}{f L_f}. \quad (3.5a,b)$$

However, this pair of non-dimensional numbers is not necessarily the best choice to span the parameter space. Indeed, we find that the ratio of the filament Burger and Rossby numbers,

$$\gamma_m := \frac{Bu_m}{Ro_f} \equiv \frac{c_m^2}{f L_f \Delta V}, \quad (3.6)$$

can be used to distinguish qualitatively different dynamical regimes. A heuristic explanation of why γ_m is an important parameter is that the spatial structure of the waves is set by the dispersive, $Bu_m \partial^2 / \partial \tilde{x}^2$, and refractive, $Ro_f \mathcal{F}(\tilde{x})$, terms. As a result, $\gamma_m := Bu_m / Ro_f$ determines whether the spatial scale of the waves is short or long compared with the width of the filament. In particular, we find that, for $\gamma_m \ll 1$, the spatial scale of the waves is $\sqrt{\gamma_m} L_f \ll L_f$ whereas, for $\gamma_m \gg 1$, the spatial scale of the waves is $\gamma_m L_f \gg L_f$. When γ_m is large, the waves may penetrate across the cyclonic filament in a manner analogous to quantum tunnelling. Therefore, we call γ_m the ‘tunnelling parameter’.

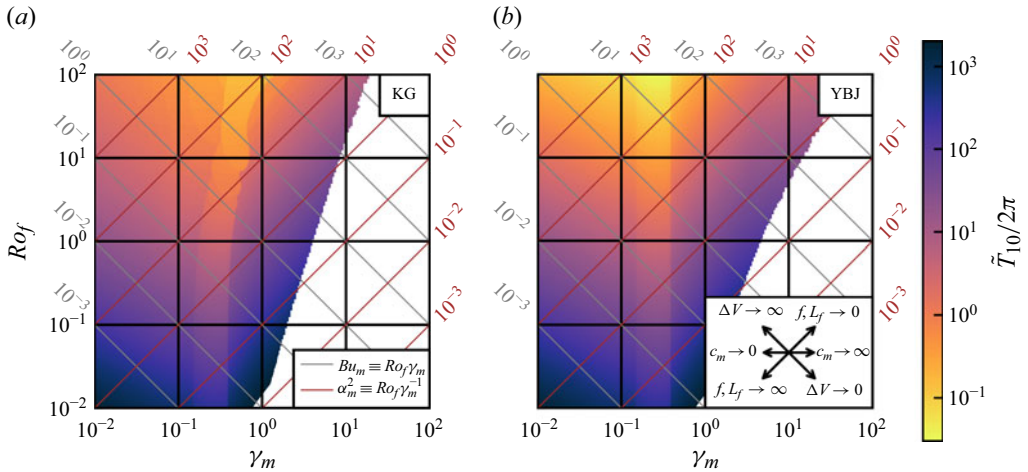


Figure 2. Radiation time $\tilde{T}_{10}/2\pi := fT_{10}/2\pi$ (3.2) in inertial periods computed from numerical simulations of a Gaussian filament (§ 3.5) as a function of γ_m and Ro_f for both the Klein–Gordon (a) and YBJ (b) problems. Lines of constant Bu_m and α_m^2 are overlaid in grey and brown, respectively. The diagram inset in (b) indicates the distinguished limits summarised in table 3. White regions are excluded as, for these parameter values, waves radiated from the filament loop around the finite numerical domain and return to the filament before the T_{10} criterion is met.

Critically, these length scale considerations determine which asymptotic techniques we may employ to investigate the problem. Namely, when $\gamma_m \ll 1$, we use ray-tracing results derived from a WKB approximation whereas when $\gamma_m \gg 1$, we approximate the filament as a delta-function.

To make the dependence on γ_m explicit, and to aid the asymptotic analysis by expressing the filament as an order 1 function of an order 1 parameter, we divide (3.4) by Ro_f to give

$$\left[Ro_f^{-1} \left(1 + \frac{\partial^2}{\partial \tilde{t}^2} \right) - \gamma_m \frac{\partial^2}{\partial \tilde{x}^2} + \mathcal{F}(\tilde{x}) \right] \mathcal{X} = 0. \quad (3.7a)$$

Applying the same manipulations to the time-dependent Schrödinger equation (2.9) gives

$$\left[-2iRo_f^{-1} \frac{\partial}{\partial \tilde{t}} - \gamma_m \frac{\partial^2}{\partial \tilde{x}^2} + \mathcal{F}(\tilde{x}) \right] \mathcal{A} = 0 \quad (3.7b)$$

where the filament Rossby number can be absorbed into a rescaled time. Thus, in YBJ theory, γ_m is the only dynamically interesting parameter (e.g. Young & Ben Jelloul 1997; Danioux, Vanneste & Bühler 2015; Asselin & Young 2019; Conn, Callies & Lawrence 2025). This parameter dependence is neatly illustrated by the radiation time scale T_{10} , that is, the time scale over which the solution decays at the centre of the filament. We plot the radiation time scale, computed from numerical simulations of a Gaussian filament (see § 3.5), in units of inertial periods (figure 2b). There is a simple monotonic structure along lines of $\gamma_m = \text{constant}$. Indeed, (3.7b) implies a linear dependence on Ro_f in log space. Whereas, on other lines with γ_m varying, we observe a more complex non-monotonic structure. The same plot for the full Klein–Gordon equation (figure 2a) displays similar parameter dependence but with more complicated behaviour when $Ro_f > 1$.

The critical importance of the parameter γ_m motivates using it as one of the two non-dimensional numbers spanning the parameter space. Writing the Klein–Gordon equation

Vary	Fixed non-dimensional number	$\gamma_m \rightarrow 0$ limit	$\gamma_m \rightarrow \infty$ limit
c_m	Ro_f	High mode, weakly stratified	Low mode, strongly stratified
f or L_f	α_m	Rapidly rotating, broad filament	Non-rotating, sharp filament
ΔV	Bu_m	Strong filament	Weak filament

Table 3. Distinguished limits achieved by fixing three of the four dimensional parameters and sending the fourth to 0 or ∞ .

in the form (3.7a) suggests using the filament Rossby number as the second non-dimensional number and by default this is what we do (see, e.g. figures 2 and 4). However, this is not the only choice. Indeed, there are at least two other interesting choices that warrant comment. Of the four dimensional parameters defining this problem two are velocity scales, c_m and ΔV , and the length and time scales may be combined to form a third velocity scale, fL_f . We can define three non-dimensional parameters that are independent of one of these velocity scales. The filament Rossby number (which is independent of c_m) is the first of these, the second is the filament Burger number (which is independent of ΔV) and the third is

$$\alpha_m := \frac{\Delta V}{c_m} \quad (3.8)$$

that is independent of f and L_f . These three parameters are related through γ_m :

$$Bu_m \gamma_m^{-1} \equiv Ro_f \equiv \alpha_m^2 \gamma_m. \quad (3.9)$$

By fixing one of these three parameters and then sending $\gamma_m \rightarrow 0$ or $\gamma_m \rightarrow \infty$, six physically interesting distinguished limits can be reached. For example, suppose that we wish to study the strongly stratified limit, then we fix the values of f , L_f and ΔV before sending $c_m \rightarrow \infty$. However, this is simply the distinguished limit $\gamma_m \rightarrow \infty$ with Ro_f fixed. Alternatively, the sharp filament limit $L_f \rightarrow 0$ with f , c_m and ΔV fixed is the distinguished limit $\gamma_m \rightarrow \infty$ with α_m fixed. The six distinct distinguished limits are summarised in table 3 and figure 2(b).

3.3. The WKB approximation – $\gamma_m \ll 1$

We begin by considering the regime in which the WKB approximation is valid, namely $\gamma_m \ll 1$. Here, we recall some standard ray-tracing results. However, in Appendix A we give a formal WKB derivation in the limit $\gamma_m \rightarrow 0$ considering the three distinguished limits summarised in table 3. The solution is expressed in terms of a slowly varying amplitude and rapidly varying phase

$$\mathcal{X} = Ae^{i\theta} + \text{c.c.} \quad (3.10)$$

Then, with non-dimensional frequency $\tilde{\omega} := -\partial\theta/\partial\tilde{t}$, which given the steady background flow is conserved along rays, and local wavenumber $\tilde{k} := \partial\theta/\partial\tilde{x}$, (3.7a) admits the dispersion relation

$$\tilde{\omega}^2 - 1 = Ro_f [\mathcal{F}(\tilde{x}) + \gamma_m \tilde{k}^2] \quad (3.11a)$$

and the group velocity is

$$\tilde{c}_g := \frac{\partial\tilde{\omega}}{\partial\tilde{k}} = Ro_f \gamma_m \frac{\tilde{k}}{\tilde{\omega}}. \quad (3.11b)$$

Notably, as the waves radiate away from the centre of the filament, to conserve $\tilde{\omega}$, $\tilde{k}$ becomes $O(\gamma_m^{-1/2})$, i.e. the waves have spatial scale $\sqrt{\gamma_m}L_f$. The YBJ versions of (3.11) are

$$\tilde{\omega}^{(YBJ)} = 1 + \frac{1}{2}Ro_f[\mathcal{F}(\tilde{x}) + \gamma_m\tilde{k}^2], \quad \tilde{c}_g^{(YBJ)} = Ro_f\gamma_m\tilde{k}. \quad (3.12a,b)$$

In both cases the transport equation is

$$\frac{\partial}{\partial \tilde{t}}(A^2) + \frac{\partial}{\partial \tilde{x}}(\tilde{c}_g A^2) = 0. \quad (3.13)$$

We note that $\tilde{\omega}^{(YBJ)}$ is the expansion of $\tilde{\omega}$ to $O(Ro_f)$ and that $\tilde{c}_g^{(YBJ)}$ is (3.11b) but using the leading-order, i.e. $\tilde{\omega} = 1$, expression for the frequency. Consequently, if we apply the YBJ approximation when high frequency waves are being radiated then the group velocity will be overpredicted and the wave energy radiated too rapidly.

We now compute expressions for T_{10} and $T_{10}^{(YBJ)}$ under the WKBJ approximation. We start by computing the travel time of a ray originating at $\tilde{x} = \tilde{x}_0$. Using the dispersion relation (3.11a) to eliminate $\tilde{k}$ from the group velocity (3.11b), we have

$$\tilde{c}_g^2 = \left(\frac{d\tilde{x}}{d\tilde{t}}\right)^2 = Ro_f\gamma_m \frac{\tilde{\omega}^2 - (1 + Ro_f\mathcal{F}(\tilde{x}))}{\tilde{\omega}^2}. \quad (3.14a)$$

As the initial conditions are uniform, we have $\tilde{k} = 0$ at $\tilde{t} = 0$ and, thus, the frequency satisfies

$$\tilde{\omega}^2 = 1 + Ro_f\mathcal{F}(\tilde{x}_0). \quad (3.14b)$$

It then follows that the travel time $\tilde{\tau}$ of a ray from $\tilde{x}_0 > 0$ to $\tilde{x} > \tilde{x}_0$ is

$$\tilde{\tau}(\tilde{x}; \tilde{x}_0) = Ro_f^{-\frac{1}{2}}\gamma_m^{-\frac{1}{2}} \int_{\tilde{x}_0}^{\tilde{x}} \sqrt{\frac{Ro_f^{-1} + \mathcal{F}(\tilde{x}_0)}{\mathcal{F}(\tilde{x}_0) - \mathcal{F}(\tilde{x}')}} d\tilde{x}'. \quad (3.14c)$$

The equivalent YBJ calculation gives

$$\tilde{\tau}^{(YBJ)}(\tilde{x}; \tilde{x}_0) = Ro_f^{-1}\gamma_m^{-\frac{1}{2}} \int_{\tilde{x}_0}^{\tilde{x}} \frac{1}{\sqrt{\mathcal{F}(\tilde{x}_0) - \mathcal{F}(\tilde{x}')}} d\tilde{x}', \quad (3.15)$$

which is the leading-order term of (3.14c) in the small Ro_f limit.

To compute T_{10} , we must consider rays close to the centre of the filament where $\mathcal{F}(\tilde{x}) = \mathcal{F}(0) + (1/2)\mathcal{F}''(0)\tilde{x}^2 + \text{h.o.t.}$. Here, $\mathcal{F}'' := \partial^2\mathcal{F}/\partial\tilde{x}^2$. Inserting this into (3.14c) we get

$$Ro_f^{\frac{1}{2}}\gamma_m^{\frac{1}{2}}\tilde{\tau}(\tilde{x}; \tilde{x}_0) \approx \int_{\tilde{x}_0}^{\tilde{x}} \sqrt{\frac{Ro_f^{-1} + \mathcal{F}(0)}{\frac{1}{2}\mathcal{F}''(0)(\tilde{x}_0^2 - \tilde{x}'^2)}} d\tilde{x}' = \sqrt{\frac{Ro_f^{-1} + \mathcal{F}(0)}{\frac{1}{2}\mathcal{F}''(0)}} \text{arccosh}\left(\frac{\tilde{x}}{\tilde{x}_0}\right). \quad (3.16)$$

From this expression we see that rays originating near the centre of the filament fan out uniformly in the sense that $\tau(\tilde{x}; \tilde{x}_0) = \tau(\tilde{x}/\tilde{x}_0)$. The transport equation (3.13) implies that the total wave amplitude squared between any two rays is conserved. That is, if $a(\tilde{t})$ and $b(\tilde{t})$ are two rays with initial positions a_0 and b_0 then

$$\frac{d}{d\tilde{t}} \int_{a(\tilde{t})}^{b(\tilde{t})} A^2 d\tilde{x} = 0 \implies \frac{d}{d\tilde{t}} \int_{a_0}^{b_0} A^2 \frac{d\tilde{x}}{d\tilde{x}_0} d\tilde{x}_0 = 0 \implies \int_{a_0}^{b_0} A^2 \frac{d\tilde{x}}{d\tilde{x}_0} d\tilde{x}_0 = \text{const.} \quad (3.17)$$

Not only does this hold for all rays but these rays are fanning out uniformly and, hence, $d\tilde{x}/d\tilde{x}_0$ is independent of $\tilde{x}_0$. It therefore follows that, for the amplitude to drop by a factor of 10, we require $\tilde{x}/\tilde{x}_0 = 10^2 = 100$. Thus, we have

$$\tilde{T}_{10} := fT_{10} = Ro_f^{-\frac{1}{2}}\gamma_m^{-\frac{1}{2}} \sqrt{\frac{Ro_f^{-1} + \mathcal{F}(0)}{\frac{1}{2}\mathcal{F}''(0)}} \operatorname{arccosh}(100) \quad (3.18a)$$

and

$$\tilde{T}_{10}^{(YBJ)} := fT_{10}^{(YBJ)} = Ro_f^{-1}\gamma_m^{-\frac{1}{2}} \frac{1}{\sqrt{\frac{1}{2}\mathcal{F}''(0)}} \operatorname{arccosh}(100). \quad (3.18b)$$

Comparing these ray-tracing results for the full Klein–Gordon and YBJ versions of the problem, we conclude that we may divide the $\gamma_m \ll 1$ regime into two subregimes (IA and IB) dependent on the Rossby number. Regime IA is the regime $\gamma_m \ll 1, Ro_f \ll 1$ where the YBJ approximation is valid. On the other hand, regime IB is the regime $\gamma_m \ll 1, Ro_f \gg 1$ and the YBJ approximation is not valid. These conclusions regarding the validity of the YBJ approximation are corroborated by the formal asymptotic analysis in [Appendix A](#). In the two subregimes the radiation time scale, $\tilde{T} := fT$, scales as

$$\tilde{T} \sim \begin{cases} \gamma_m^{-\frac{1}{2}} Ro_f^{-1}, & Ro_f \ll 1, \\ \gamma_m^{-\frac{1}{2}} Ro_f^{-\frac{1}{2}}, & Ro_f \gg 1. \end{cases} \quad (3.19)$$

In both subregimes the radiation time scale is a decreasing function of both the Rossby number and tunnelling parameter.

3.4. The delta-function limit – $\gamma_m \gg 1$

We now consider $\gamma_m \gg 1$. From (3.7a), we see that, to leading order, NIWs will not vary on the filament scale or the dispersive term, $\gamma_m \partial^2/\partial \tilde{x}^2$, would be unbalanced. Thus, our previous choice of spatial non-dimensionalisation, by L_f , is no longer appropriate. We choose a new non-dimensionalisation by defining

$$\tilde{x} := \frac{x}{L_m}, \quad L_m := \gamma_m L_f \equiv \frac{c_m^2}{f \Delta V}. \quad (3.20a,b)$$

Multiplying (3.7a) through by γ_m we have

$$\left[\alpha_m^{-2} \left(1 + \frac{\partial^2}{\partial \tilde{t}^2} \right) - \frac{\partial^2}{\partial \tilde{x}^2} + \gamma_m \mathcal{F}(\gamma_m \tilde{x}) \right] \mathcal{X} = 0. \quad (3.21a)$$

This choice of non-dimensionalisation has achieved two things. Firstly, we have cleared the coefficient of the spatial derivatives and, secondly, the filament is now in the form of a nascent delta-function. As $\gamma_m \rightarrow \infty$,

$$\gamma_m \mathcal{F}(\gamma_m \tilde{x}) \rightarrow \delta(\tilde{x}). \quad (3.21b)$$

We first study the distinguished limit in which α_m is held constant. This is the most natural interpretation of the delta-function limit as it corresponds to sending $L_f \rightarrow 0$ with c_m , f and ΔV held fixed. We consider the other distinguished limits summarised in [table 3](#) at the end of the section.

Integrating over the delta-function and requiring that the solution is even, we have the jump condition $\partial \mathcal{X} / \partial \tilde{x} = (1/2)\mathcal{X}$ at $\tilde{x} = 0^+$. Redimensionalising, the jump condition is

$L_m \partial \mathcal{X} / \partial x = (1/2) \mathcal{X}$. Thus, we may interpret L_m as the length scale imposed on the problem by the delta-function filament through the jump condition.

In [Appendix B](#) we derive a solution in the delta-function limit. For $\tilde{t} < \alpha_m^{-1} |\tilde{x}|$, the solution is simply $\mathcal{X}(\tilde{x}, \tilde{t}) = \cos \tilde{t}$. Note that in $(\tilde{x}, \tilde{t})$ coordinates, the characteristics of (3.21a) have slope $\pm \alpha_m^{-1}$ and so the points $|\tilde{x}| > \alpha_m \tilde{t}$ are beyond the influence of the filament. For $\tilde{t} \geq \alpha_m^{-1} |\tilde{x}|$, the solution is given by

$$\mathcal{X}(\tilde{x}, \tilde{t}) = \cos \tilde{t} - \frac{1}{2} \alpha_m \int_0^{\tilde{t} - \alpha_m^{-1} |\tilde{x}|} J_0 \left(\sqrt{(\tilde{t} - \tilde{t}')^2 - \alpha_m^{-2} \tilde{x}^2} \right) \mathcal{X}(0, \tilde{t}') d\tilde{t}', \quad (3.22a)$$

where J_0 is the zeroth-order Bessel function of the first kind. The solution at any point may be computed from the history of the solution at the centre of the filament. The solution at the centre of the filament is

$$\mathcal{X}(0, \tilde{t}) = \frac{2}{\pi} \int_0^\infty \frac{1}{1+u^2} \cos \left(\tilde{t} \sqrt{1 + \frac{1}{4} \alpha_m^2 u^2} \right) du. \quad (3.22b)$$

For fixed $\tilde{t}$, as $\alpha_m \rightarrow 0$, this reduces to the elementary integral $(2/\pi) \cos \tilde{t} \int_0^\infty (1+u^2)^{-1} du = \cos \tilde{t}$. Evaluating (3.22b) numerically, we find that, for larger α_m , i.e. stronger filaments, the decay at the centre of the filament is more rapid ([figure 3](#)). Furthermore, we find very good agreement between the analytic delta-function solution and numerical simulations (see § 3.5) with $\gamma_m = 5$.

For large $\tilde{t}$, (3.22b) is amenable to the stationary phase approximation and we find that

$$\mathcal{X}(0, \tilde{t}) \sim \sqrt{\frac{8}{\pi \alpha_m^2 \tilde{t}}} \cos \left(\tilde{t} + \frac{\pi}{4} \right). \quad (3.23)$$

Here, large $\tilde{t}$ means $\tilde{t} \gg \max(1, \alpha_m^{-2})$. For $O(1)$ values of α_m , the stationary phase approximation is very good on an inertial time scale (e.g. [figure 3b](#) where $\alpha_m = 2$). Furthermore, when $\tilde{t}$ is large, the solution is an inertial oscillation with a slowly decaying amplitude. This is a situation in which the YBJ approximation should be expected to perform well. However, this provides no guarantee that the YBJ approximation will correctly capture the small $\tilde{t}$ behaviour.

To compare to § 3.3, we again compute the radiation time scale. Here we have $\tilde{T} \sim \alpha_m^{-2} \equiv Ro_f^{-1} \gamma_m$, which is once again a decreasing function of the Rossby number but now an increasing function of the tunnelling parameter. More precisely,

$$\tilde{T}_{10} = 10^2 \times \frac{8}{\pi} \alpha_m^{-2} = \frac{800}{\pi} Ro_f^{-1} \gamma_m. \quad (3.24)$$

However, if α_m^2 is large then this result may be inaccurate as $\tilde{T}_{10}$ may be too small for the stationary phase approximation to be valid.

Finally, we consider the other distinguished limits. In particular, we consider $\gamma_m \rightarrow \infty$ with Ro_f fixed. The case with Bu_m fixed can be handled in exactly the same way. The complicating factor is that $\alpha_m^{-2} \equiv Ro_f^{-1} \gamma_m \rightarrow \infty$ as $\gamma_m \rightarrow \infty$. The solution is to utilise substitution (2.10), $\mathcal{X} = Ae^{-i\tilde{t}} + \text{c.c.}$,

$$\left[Ro_f^{-1} \gamma_m \left(\frac{\partial^2}{\partial \tilde{t}^2} - 2i \frac{\partial}{\partial \tilde{t}} \right) - \frac{\partial^2}{\partial \tilde{x}^2} + \gamma_m \mathcal{F}(\gamma_m \tilde{x}) \right] \mathcal{A} = 0, \quad (3.25a)$$

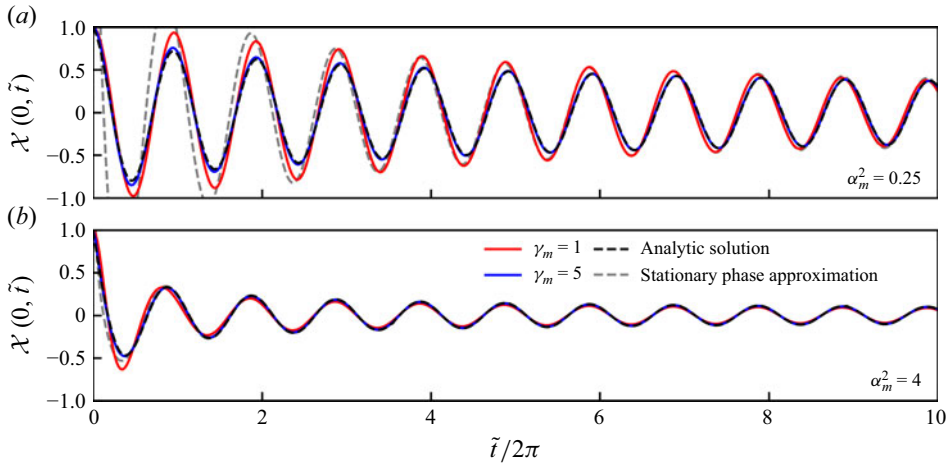


Figure 3. Solutions $\mathcal{X}(0, \tilde{t})$ to the Klein–Gordon equation at $x=0$ for (a) $\alpha_m^2 = 0.25$ and (b) $\alpha_m^2 = 4$. Solid lines are from the numerical solutions. The dashed black line shows the analytic solution (3.22b) and the dashed grey line shows the stationary phase approximation (3.23). The time axis is in inertial periods.

and then rescale time. We define $\check{t} := Ro_f \gamma_m^{-1} \tilde{t} \equiv \alpha_m^2 \tilde{t}$ such that

$$\left[Ro_f \gamma_m^{-1} \frac{\partial^2}{\partial \check{t}^2} - 2i \frac{\partial}{\partial \check{t}} - \frac{\partial^2}{\partial \check{x}^2} + \gamma_m \mathcal{F}(\gamma_m \check{x}) \right] \mathcal{A} = 0. \quad (3.25b)$$

The leading-order solution is thus given by the delta-function solution to the YBJ equation, i.e. the $\partial^2/\partial \check{t}^2$ term does not appear until $O(\gamma_m^{-1})$. That being said, the full Klein–Gordon delta-function solution is still informative in this case as it just corresponds to the inclusion of a higher-order term. Finally, we note that the temporal scaling $\check{t} \equiv \alpha_m^2 \tilde{t}$ introduced here in the YBJ solution is the same scaling that arose in the stationary phase approximation (3.23).

Importantly, this scaling analysis means that the delta-function approximation not only applies in the distinguished limit $L_f \rightarrow 0$ with c_m , f and ΔV fixed but also in the limit $c_m \rightarrow \infty$ with f , L_f and ΔV fixed. In both cases, the radiation time scale is $\tilde{T} \sim \alpha_m^2 = Ro_f \gamma_m^{-1}$, although in the latter case this time scale is guaranteed to be large. In other words, when the stratification is sufficiently strong or the vertical wavelength sufficiently large, the waves, to leading order, have no structure on the filament scale and it takes a very long time for the presence of the filament to be felt regardless of the strength of the filament.

3.5. Numerical solutions

To explore the intermediate regime $\gamma_m = O(1)$, validate our scalings for the decay at the centre of the filament and further assess the validity of the YBJ approximation, we numerically solve the Klein–Gordon and time-dependent Schrödinger equations for a Gaussian filament $\mathcal{F}(\eta) = \exp(-\eta^2/2)/\sqrt{2\pi}$ (figure 1a). Details of the numerical schemes can be found in Appendix C. We present the results using the non-dimensionalisation of § 3.3, $\tilde{t} = ft$, $\tilde{x} = x/L_f$, and use γ_m and Ro_f as coordinates for the parameter space. For the numerical solutions, the ‘unbounded’ domain has periodic boundary conditions at $\tilde{x} = \pm 5000$ and we use 2^{18} grid points.

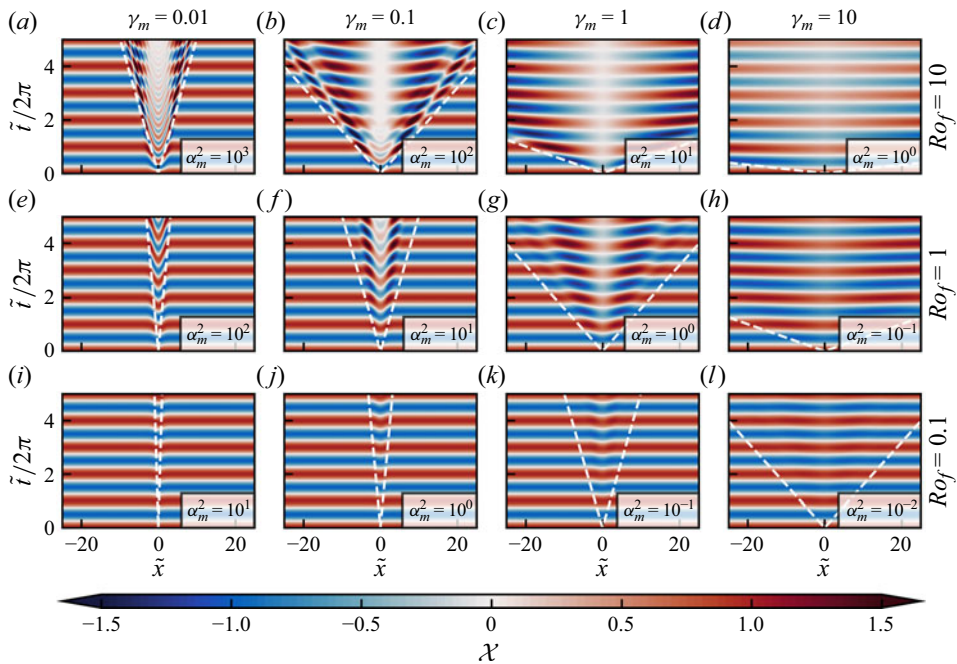


Figure 4. Hovmöller plots of $\mathcal{X}$ showing the radiation by an unbound cyclonic filament for varying Ro_f and γ_m . Dashed white lines indicate rays travelling at the mode speed $\tilde{t} = \pm \tilde{x} / \sqrt{Ro_f \gamma_m}$. The time axis is in inertial periods.

Regime	Limits	Spatial non-dimensionalisation	Approximation	Radiation time scale
IA	$\gamma_m \ll 1, Ro_f \ll 1$	$\tilde{x} = x/L_f$	WKBJ	$\tilde{T} \sim Ro_f^{-1} \gamma_m^{-\frac{1}{2}}$
IB	$\gamma_m \ll 1, Ro_f \gg 1$	$\tilde{x} = x/L_f$	WKBJ	$\tilde{T} \sim Ro_f^{-\frac{1}{2}} \gamma_m^{-\frac{1}{2}}$
II	$\gamma_m \gg 1$	$\check{x} = x/L_m = \tilde{x}/\gamma_m$	Delta-function	$\tilde{T} \sim \alpha_m^{-2} \equiv Ro_f^{-1} \gamma_m$

Table 4. Spatial non-dimensionalisations, approximations and scalings for the decay time scale at the centre of the filament in the different regimes of the unbounded radiation problem.

From a suite of 12 Klein–Gordon simulations (figure 4), we observe, for fixed Ro_f , a transition in behaviour from wave radiation for small γ_m to a more spatially coherent decaying response for large γ_m . We also observe that, for fixed γ_m , increasing Ro_f leads to more rapid decay in the filament. However, the more interesting dependence is on γ_m . The scalings derived earlier and summarised in table 4 predict that the radiation time scale is a decreasing function of γ_m , $\tilde{T} \sim \gamma_m^{-1/2}$, for small γ_m and an increasing function $\tilde{T} \sim \gamma_m$ for large γ_m . We should therefore expect the most rapid radiation for some intermediate value of γ_m . This is exactly what we observe in the simulations with the most rapid radiation occurring for the $\gamma_m = 0.1$ simulations. For a large Rossby number, this radiation can be extremely rapid. For example, for $\gamma_m = 0.1$, $Ro_f = 10$ (figure 4b), the solution at the centre of the filament becomes negligibly small in less than two inertial periods. It should be noted that these values are quite reasonable in some regions of the world’s oceans. With $Ro_f = 10$ the geostrophic velocity gradient at the centre of the filament is

$\partial V/\partial x = (10/\sqrt{2\pi})f \approx 4.0f$, well within the range of values simulated by Schlichting *et al.* (2023) for example.

Plotting $\tilde{T}_{10}$, computed by running numerical simulations until the stopping criterion (3.2) is reached, as a function of γ_m for various Ro_f (figures 2 and 5a) once again highlights the monotonic dependence on Rossby number but non-monotonic dependence on the tunnelling parameter. However, there is a significant and sharp drop in $\tilde{T}_{10}$ as we transition to intermediate values of γ_m . In particular, the transition into regime II, which, for $Ro_f \ll 1$, occurs at $\gamma_m \approx 0.4$, is abrupt and $\tilde{T}_{10}$ is very sensitive to the value of γ_m around this point. To validate our predictions for the radiation time scale, we should be able to make the curves collapse by appropriately rescaling the axes. Therefore, we let $Y = Ro_f^a \tilde{T}_{10}$ and $X = Ro_f^b \gamma_m$, where a and b are exponents to be determined. In each of the three subregimes we have scalings for $\tilde{T}$ in terms of γ_m and Ro_f (table 4). Eliminating γ_m gives the following relationships between Y and X for regimes IA, IB and II respectively:

$$Y^2 \sim Ro_f^{(2a+b-2)} X^{-1}, \quad Y^2 \sim Ro_f^{(2a+b-1)} X^{-1}, \quad Y \sim Ro_f^{(a-b-1)} X. \quad (3.26a-c)$$

To make the curves collapse, we must eliminate Ro_f . For small Ro_f , subregimes IA and II apply and, thus, we require $a = 1$ and $b = 0$. Whereas for large Ro_f , regimes IB and II apply and we require $a = 2/3$ and $b = -1/3$. Applying these rescalings to the axes (figures 5b and 5c) we find that the curves do indeed collapse. Furthermore, by plotting the asymptotic predictions for $\tilde{T}_{10}$ from (3.18) and (3.24) on top of the numerical results we find excellent agreement.

To further assess the validity of the YBJ approximation as a function of γ_m and Ro_f , we plot the difference between the YBJ approximation and Klein–Gordon solutions in figure 6. For small Ro_f , the YBJ approximation performs well for all values of γ_m . For larger Ro_f , when γ_m is small, we observe differences. Notably, the YBJ solution radiates waves outside of the region bounded by the mode speed, $\tilde{t} = \pm \tilde{x}/\sqrt{Ro_f \gamma_m}$ (white dashed lines in figure 6). This is a manifestation not just of the overprediction of the group velocity highlighted in § 3.3 but also of the fact that the Klein–Gordon equation is hyperbolic and propagates information at a finite speed (the mode speed) whereas the time-dependent Schrödinger equation is only first-order in time and can propagate information at any speed. However, when $\gamma_m = 10$, the YBJ approximation performs well even for $Ro_f = 10$. This is consistent with our analysis in § 3.4 as for $\gamma_m = Ro_f = 10$, $\alpha_m^2 = 1$ is not large. Furthermore, the YBJ prediction for the radiation time scale (figure 2) shows good agreement with the Klein–Gordon prediction when $\gamma_m \gg 1$ and $\alpha_m < 1$. These results and the earlier analysis suggest that the YBJ approximation is valid when $\gamma_m \ll 1$ and $Ro_f \ll 1$ or $\gamma_m \gg 1$ and $\alpha_m \ll 1$. If we assume that the validity criterion smoothly transitions across $\gamma_m = O(1)$, as figure 2 suggests it should, then we can say that the YBJ approximation should be valid whenever

$$Ro_f \ll 1 + \gamma_m. \quad (3.27)$$

3.6. Summary

In this section we considered the lateral radiation of a single vertical mode with no initial horizontal structure out of a cyclonic filament into an unbounded domain. The problem is described by a 2-D parameter space. There are two regimes distinguished by the tunnelling parameter γ_m . For small γ_m , a WKBJ approximation is permissible and we identify the filament Rossby number Ro_f as the most useful choice for the second non-dimensional parameter. In particular, for small Ro_f , the YBJ approximation is valid. For large γ_m , the filament may be treated as a delta-function and we derive an analytic solution. In this

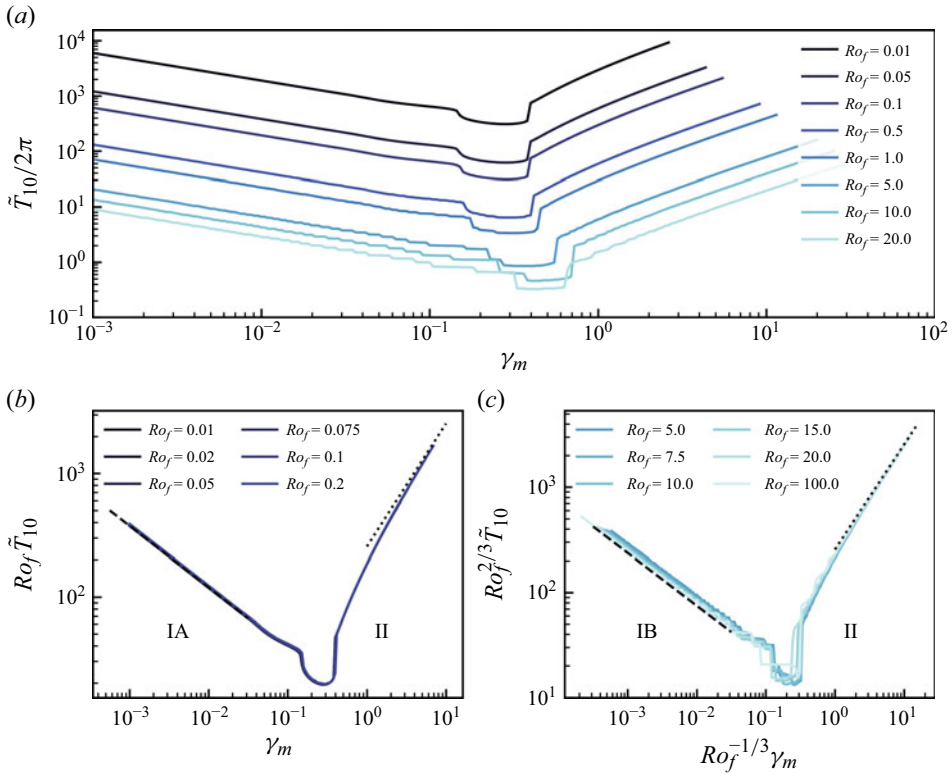


Figure 5. Time scale for the decay of velocity at the centre of the filament as a function of γ_m for various Ro_f . In (a) the time axis is in inertial periods. By appropriately rescaling the axes for small (b) and large (c) Ro_f the curves may be made to collapse. Dashed black lines indicate the radiation time scales (3.18) computed for regimes IA and IB. Dotted black lines indicate the radiation time scale (3.24) predicted by the stationary phase approximation for regime II.

regime the most useful choice for the second parameter is $\alpha_m^2 = Ro_f/\gamma_m$, which determines the validity of the YBJ approximation. A stationary phase approximation reveals that at large times the solution in the filament decays as $\alpha_m^{-1} \tilde{t}^{-(1/2)}$. A particularly important result comes from the distinguished limit $\gamma_m \rightarrow \infty$, $\alpha_m^2 \rightarrow 0$ with Ro_f held constant. In this limit, which may be achieved by sending the mode speed to infinity, we find that the filament, however large the Rossby number, is only felt by the inertial oscillations on very long time scales. Finally, we find that the decay at the centre of the filament is fastest for intermediate values (0.1 – 1) of γ_m and, for large but realistic values of the filament Rossby number, this decay can occur on inertial time scales. This ‘Goldilocks’ effect where the most efficient radiation for a given background flow (i.e. fixed values of ΔV , L_f , f and N), occurs for waves with just-the-right vertical scale is a recurring theme both of this paper and of previous studies of NIW-mean flow interactions (Balmforth *et al.* 1998; Klein & Llewellyn Smith 2001; Danioux *et al.* 2015).

4. Cyclonic filament and anticyclonic eddies

Now we move towards a more realistic set-up and introduce an additional length scale associated with the spacing between cyclonic filaments, i.e. the width of the anticyclonic regions. We take the same cyclonic filament and place it in an otherwise anticyclonic flow

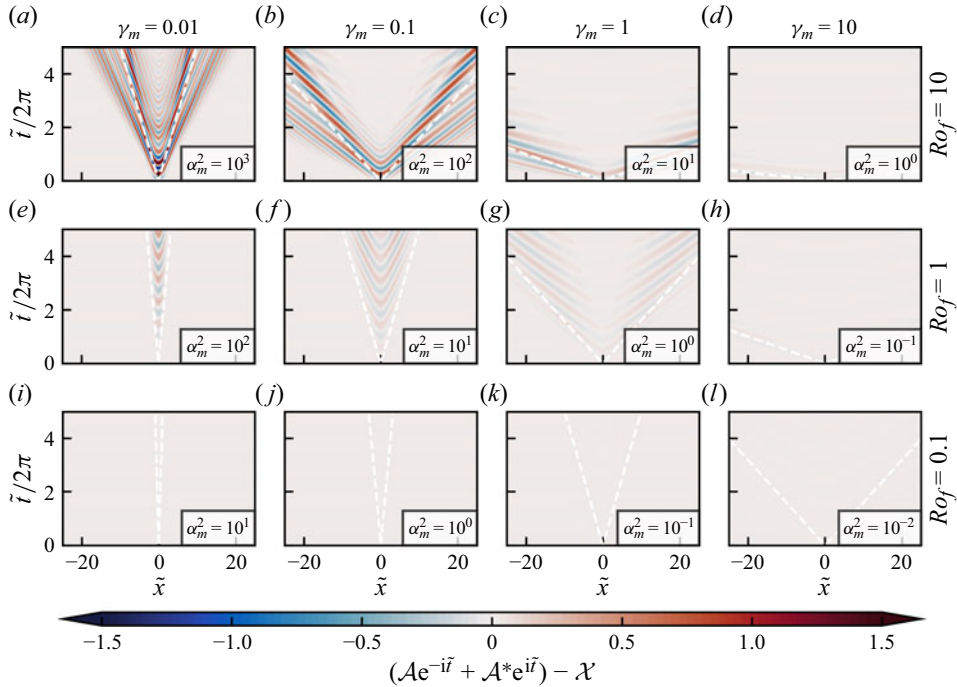


Figure 6. Hovmöller plots of the difference between the YBJ and Klein-Gordon solutions for varying Ro_f and γ_m . Dashed white lines indicate rays travelling at the mode speed $\tilde{t} = \pm \tilde{x} / \sqrt{Ro_f \gamma_m}$. The time axis is in inertial periods.

(figure 1b). The background shear is given by

$$\frac{1}{f} \frac{\partial V}{\partial x} = Ro_{ac} + \underbrace{\frac{\Delta V}{f L_f}}_{Ro_f} \mathcal{F} \left(\frac{x}{L_f} \right), \quad (4.1)$$

where $-1 < Ro_{ac} < 0$ is the Rossby number of the anticyclonic region. Then we impose periodic boundary conditions at $x = \pm L_x$ and require that the mean vorticity is zero implying that $Ro_{ac} = -(1/2)\Delta V/fL_x$. This constraint means that we have only added one additional degree of freedom to the problem, which is now defined by five dimensional parameters: c_m , f , L_f , ΔV and L_x . We require three non-dimensional parameters. A particularly useful one is

$$\xi := \frac{L_x}{L_f} \gg 1, \quad (4.2a)$$

which we can use to express the relationship between the two Rossby numbers

$$Ro_f \equiv 2\xi |Ro_{ac}|. \quad (4.2b)$$

4.1. Reduction to the radiation problem

If we may ignore the boundary conditions then this problem is equivalent to the unbounded radiation problem considered in § 3. However, we must account for the change in the background flow away from the filament. Separating out the constant anticyclonic part

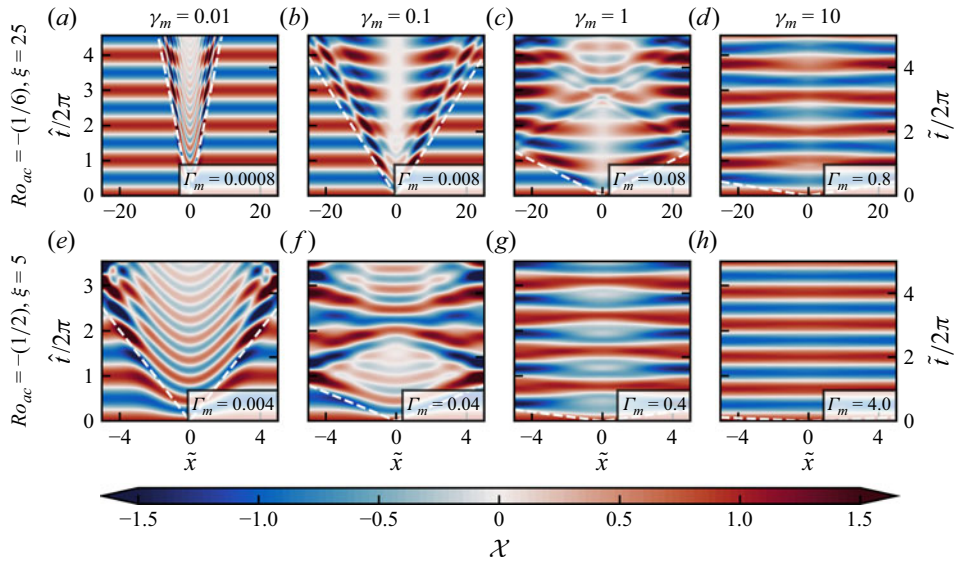


Figure 7. Hovmöller plots showing the evolution of NIWs interacting with a cyclonic filament in an otherwise anticyclonic flow. The parameters have been chosen to be dynamically equivalent to figure 4(a–d). Dashed white lines indicate rays travelling at the free wave speed. The left time axis is normalised by the period of a wave with frequency equal to the effective Coriolis frequency of the anticyclonic region. The right time axis is in inertial periods.

of the background flow, the Klein–Gordon equation (2.5) can be written as

$$\left[\left(1 + Ro_{ac} + \frac{\partial^2}{\partial \tilde{t}^2} \right) - \frac{c_m^2}{f^2} \frac{\partial^2}{\partial x^2} + Ro_f \mathcal{F} \left(\frac{x}{L_f} \right) \right] \mathcal{X} = 0. \quad (4.3a)$$

Now we rescale time by defining $\hat{t} := \sqrt{1 + Ro_{ac}} \tilde{t} \equiv f_{eff}^{(ac)} t$, where

$$f_{eff}^{(ac)} := f \sqrt{1 + Ro_{ac}} < f \quad (4.3b)$$

is the effective Coriolis frequency (2.2b) of the anticyclonic region. Finally, we proceed as in § 3.3 by non-dimensionalising space by L_f , $\tilde{x} := x/L_f$, and dividing (4.3a) by Ro_f to get

$$\left[\frac{1 + Ro_{ac}}{Ro_f} \left(1 + \frac{\partial^2}{\partial \hat{t}^2} \right) - \gamma_m \frac{\partial^2}{\partial \tilde{x}^2} + \mathcal{F}(\tilde{x}) \right] \mathcal{X} = 0. \quad (4.3c)$$

This is identical to (3.7a) up to a redefinition of the Rossby number. Crucially, the key dimensionless number, the tunnelling parameter γ_m , appears unchanged.

Consider the top row of figure 4 in which we looked at the unbounded radiation problem with $Ro_f = 10$ for various γ_m . We can reproduce those results in the current set-up, assuming the boundary conditions are not important, by using the same γ_m values and choosing Ro_{ac} and Ro_f such that $Ro_f/(1 + Ro_{ac}) = 10$. We do this for $Ro_{ac} = -1/6$, $Ro_f = 50/6 \Rightarrow \xi = 25$ (figures 7a–7d) and $Ro_{ac} = -(1/2)$, $Ro_f = 5 \Rightarrow \xi = 5$ (figures 7e–7h). We observe that figures 7(a) and 7(b) are essentially identical to figures 4(a) and 4(b). Note that this comparison is aided by the fact that in figure 4 we plot the solutions for $|\tilde{x}| \leq 25$ and that in figures 7(a)–7(d) we impose periodic boundary conditions at $\tilde{x} = \pm 25$. Furthermore, figure 7(e) is the same as figures 4(a) and 7(a) but zoomed in along the $\tilde{x}$ axis as the periodic boundary conditions are enforced at $\tilde{x} = \pm 5$.

This similitude is a manifestation of the dynamical equivalence of these two problems when the boundaries are not influencing the flow.

In order for the boundary conditions to not affect the flow we require that L_x be much larger than the important length scales of the radiation problem. First, we have $L_x \gg L_f$ by construction. Another important length scale, derived from the mode speed, is the Rossby radius, c_m/f , which determines the distance that the Klein–Gordon equation can propagate on an inertial time scale. If $L_x \gg c_m/f$ then it takes a long time for filament to feel the influence of the boundary conditions. Moving from figures 7(a) to 7(d) or from figures 7(e) to 7(h), corresponds to an increase in c_m with the other parameters held constant. As c_m is increased, the solutions diverge from figures 4(a) to 4(d) and tend towards uniform inertial oscillations. Notably, these oscillations are inertial, i.e. they have dimensional period $2\pi/f$ not $2\pi/f_{\text{eff}}^{(ac)}$, and do not decay in time. The other key length scale in the radiation problem was L_m and we find that the ratio of L_m to L_x is the key non-dimensional number in the following sections.

4.2. Horizontal modes

We now consider the cases for which the boundary conditions are important. Here and throughout the rest of the paper, it is most convenient to non-dimensionalise space by the half-width of the domain L_x . Therefore, we define

$$\hat{x} := \frac{x}{L_x} \quad (4.4a)$$

and we non-dimensionalise the lateral geostrophic shear with a factor of $|Ro_{ac}|$, i.e.

$$\frac{\partial \hat{V}}{\partial \hat{x}} := |Ro_{ac}|^{-1} \frac{1}{f} \frac{\partial V}{\partial x} = -1 + 2\xi \mathcal{F}(\xi \hat{x}), \quad (4.4b)$$

where again $\xi := L_x/L_f$. This implies that we have non-dimensionalised the geostrophic velocity using the velocity scale $(1/2)\Delta V$. Once more using the temporal non-dimensionalisation $\tilde{t} := ft$, we write the Klein–Gordon equation as

$$\left[|Ro_{ac}|^{-1} \left(1 + \frac{\partial^2}{\partial \tilde{t}^2} \right) - \Gamma_m \frac{\partial^2}{\partial \hat{x}^2} + \frac{\partial \hat{V}}{\partial \hat{x}} \right] \mathcal{X}. \quad (4.5)$$

The three dimensionless parameters describing the problem are

$$|Ro_{ac}| := \frac{1}{2} \frac{\Delta V}{f L_x}, \quad \Gamma_m := 2 \frac{c_m}{\Delta V} \frac{c_m}{f L_x} \equiv 2 \frac{L_m}{L_x}, \quad \xi := \frac{L_x}{L_f} \gg 1. \quad (4.6a-c)$$

The tunnelling parameter γ_m , filament Rossby number Ro_f and α_m may be expressed in terms of these parameters as

$$\gamma_m := \frac{L_m}{L_f} \equiv \frac{1}{2} \xi \Gamma_m, \quad Ro_f \equiv 2\xi |Ro_{ac}|, \quad \alpha_m^2 := \frac{Ro_f}{\gamma_m} \equiv \frac{\Delta V^2}{c_m^2} \equiv 4 \frac{|Ro_{ac}|}{\Gamma_m}. \quad (4.7a-c)$$

The imposition of periodic boundary conditions allows us to expand $\mathcal{X}$ into discrete normal modes

$$\mathcal{X} = \sum_{n=0}^{\infty} \mathcal{X}_n e^{i\tilde{\omega}_n \tilde{t}} + \text{c.c.}, \quad (4.8a)$$

where each mode satisfies

$$\left[|Ro_{ac}|^{-1} (1 - \tilde{\omega}_n^2) - \Gamma_m \frac{\partial^2}{\partial \hat{x}^2} + \frac{\partial \hat{V}}{\partial \hat{x}} \right] \mathcal{X}_n. \quad (4.8b)$$

We have a periodic Sturm–Liouville eigenvalue problem, more precisely the one-dimensional time-independent Schrödinger equation, which we may rearrange into the standard form

$$\left[-\Gamma_m \frac{\partial^2}{\partial \hat{x}^2} + \frac{\partial \hat{V}}{\partial \hat{x}} \right] \mathcal{X}_n = \lambda_n \mathcal{X}_n, \quad (4.9)$$

where λ_n is the n th eigenvalue. The problem is mathematically equivalent to the well-studied quantum mechanics problem of a particle in a one-dimensional crystal lattice. The eigenvalue problem is independent of $|Ro_{ac}|$. However, the relationship between the frequency and the eigenvalues

$$\tilde{\omega}_n^2 = 1 + |Ro_{ac}| \lambda_n \quad (4.10a)$$

does depend on $|Ro_{ac}|$. If we were to make the YBJ approximation (2.10) then we would arrive at the exact same eigenvalue problem, but the frequency, including the carrier inertial oscillation, would be given by

$$\tilde{\omega}_n^{(YBJ)} = 1 + \frac{1}{2} |Ro_{ac}| \lambda_n. \quad (4.10b)$$

Furthermore, consider the role of the parameter ξ . It only appears in (4.9) through $\partial \hat{V} / \partial \hat{x}$. It is a parameter that determines the shape, but not strength (which is determined by $|Ro_{ac}|$), of the background lateral shear. There are some results that hold for any $\partial \hat{V} / \partial \hat{x}$. In these cases the only remaining parameter is Γ_m .

One such result is a useful expression for the derivative of the eigenvalues with respect to Γ_m . Let $\langle \cdot \rangle := (1/2) \int_{-1}^1 d\hat{x}$ be the domain average. Consider $\langle \mathcal{X}_n (\partial(4.9) / \partial \Gamma_m) \rangle$:

$$-\left\langle \mathcal{X}_n \frac{\partial^2 \mathcal{X}_n}{\partial \hat{x}^2} \right\rangle + \left\langle \mathcal{X}_n \left[-\Gamma_m \frac{\partial^2}{\partial \hat{x}^2} + \frac{\partial \hat{V}}{\partial \hat{x}} \right] \frac{\partial \mathcal{X}_n}{\partial \Gamma_m} \right\rangle = \left\langle \lambda_n \mathcal{X}_n \frac{\partial \mathcal{X}_n}{\partial \Gamma_m} \right\rangle + \frac{\partial \lambda_n}{\partial \Gamma_m} \langle \mathcal{X}_n^2 \rangle. \quad (4.11)$$

The linear operator is self-adjoint, with respect to the inner product defined by the domain average, and so the second term on the left cancels the first term on the right. Manipulating what is left gives

$$\frac{\partial \lambda_n}{\partial \Gamma_m} = \frac{\langle \mathcal{X}_n'^2 \rangle}{\langle \mathcal{X}_n^2 \rangle}, \quad (4.12)$$

where $\mathcal{X}_n' := \partial \mathcal{X}_n / \partial \hat{x}$. As a corollary, we note that all the eigenvalues, and hence frequencies, are increasing functions of Γ_m .

4.3. The minimum frequency mode

4.3.1. Bounds

The minimum frequency mode, i.e. the eigenmode with the smallest eigenvalue, is the zeroth mode and is special in that the solution has no zero crossings. This allows us to derive bounds on the minimum frequency. Since $\mathcal{X}_0'' := \partial^2 \mathcal{X}_0 / \partial \hat{x}^2$ must

take both signs somewhere in the domain. It then follows from (4.9) that $\partial \hat{V} / \partial \hat{x} - \lambda_0$ must also take both signs. This implies the lower bound $\lambda_0 > \min_{\hat{x}} \partial \hat{V} / \partial \hat{x} = -1$. To derive an upper bound, divide (4.9) by $\mathcal{X}_0$ and average to get

$$-\Gamma_m \left\langle \frac{1}{\mathcal{X}_0} \frac{\partial^2 \mathcal{X}_0}{\partial \hat{x}^2} \right\rangle + \left\langle \frac{\partial \hat{V}}{\partial \hat{x}} \right\rangle = \lambda_0 \implies \lambda_0 = -\Gamma_m \left\langle \left(\frac{\mathcal{X}'_0}{\mathcal{X}_0} \right)^2 \right\rangle + \left\langle \frac{\partial \hat{V}}{\partial \hat{x}} \right\rangle < \left\langle \frac{\partial \hat{V}}{\partial \hat{x}} \right\rangle = 0, \quad (4.13)$$

where the second equality follows from integration by parts and the periodic boundary conditions. Together these bounds are

$$-1 < \lambda_0 < 0 \implies 1 + Ro_{ac} < \tilde{\omega}_0^2 < 1. \quad (4.14)$$

These are again examples of results that hold for arbitrary periodic background states. Redimensionalising, we can in general say that the minimum frequency ω_0 must satisfy

$$\min_x f_{eff}^2(x) < \omega_0^2 < \text{mean}_x f_{eff}^2(x). \quad (4.15)$$

We now consider a couple of cases where we can make analytic progress with a focus on the minimum frequency mode.

4.3.2. The WKBJ regime – $\gamma_m \ll 1$

In the regime $\gamma_m \equiv (1/2)\xi \Gamma_m \ll 1$ we may again utilise a WKBJ approximation. In the interests of brevity, we assume that the filament is symmetric and so we need only solve over $0 < \hat{x} < 1$. If we further assume that the filament has a single maximum, like the Gaussian filament, then we have a classic two turning point eigenvalue problem as found in Bender & Orszag (1999). We define the slowly varying wavenumber for mode n ,

$$\hat{k}_n(\hat{x}; \lambda_n) := \sqrt{\frac{|-1 + 2\xi \mathcal{F}(\xi \hat{x}) - \lambda_n|}{\Gamma_m}}, \quad (4.16)$$

that is determined by the eigenvalue and the local (non-dimensional) vorticity of the background flow. Enforcing even boundary conditions the solution is

$$\mathcal{X}_{2n} = \begin{cases} C_1 \hat{k}_{2n}^{-1/2} \cosh \left(\int_0^{\hat{x}} \hat{k}_{2n}(\hat{x}') d\hat{x}' \right), & 0 \leq \hat{x} < \hat{x}_* \\ C_2 \hat{k}_{2n}^{-1/2} \cos \left(\int_{\hat{x}}^1 \hat{k}_{2n}(\hat{x}') d\hat{x}' \right), & \hat{x}_* < \hat{x} \leq 1, \end{cases} \quad (4.17)$$

where $\hat{x}_*$ is the turning point defined by $\hat{k}_{2n}(\hat{x}_*; \lambda_{2n}) = 0$ assuming one exists. The eigenvalues are determined by the connection formula across the turning point,

$$\frac{1}{2} e^{-2A_{2n}} = \tan \left(\frac{\pi}{4} - B_{2n} \right), \quad (4.18a)$$

where

$$A_{2n} := \int_0^{\hat{x}_*} \hat{k}_{2n}(x') dx', \quad B_{2n} := \int_{\hat{x}_*}^1 \hat{k}_n(x') dx'. \quad (4.18b,c)$$

A brief derivation of these formulae is given in Appendix D. We also show that B_{2n} is bounded. In particular, $B_{2n} < n\pi + \pi/4$. Rewriting (4.18b,c), we have

$$\sqrt{\Gamma_m} B_{2n} = \int_{\hat{x}_*}^1 \sqrt{\lambda_{2n} + 1 - 2\xi \mathcal{F}(\xi x')} dx'. \quad (4.19)$$

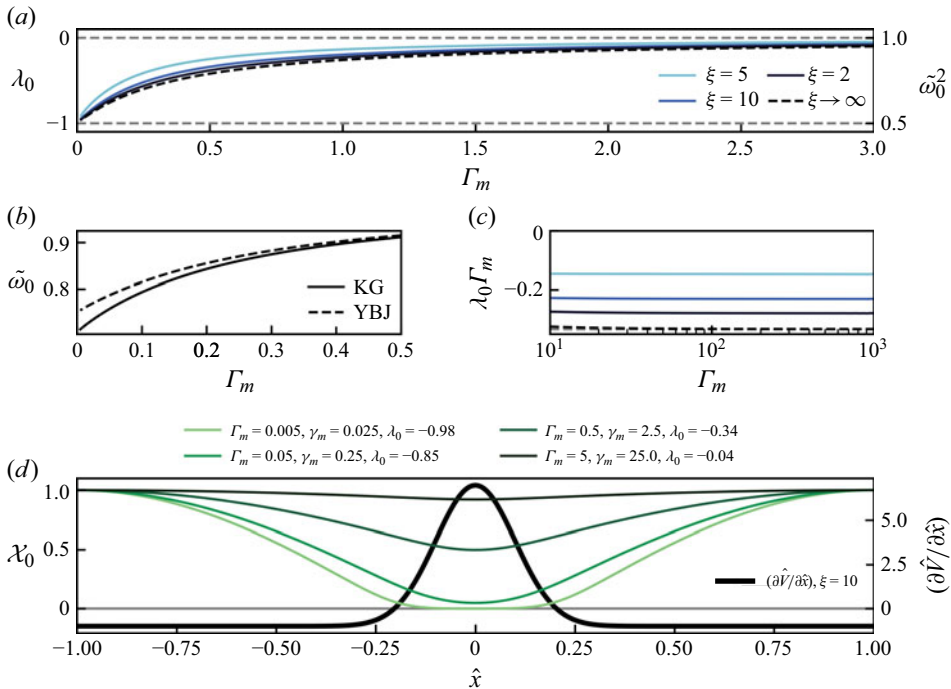


Figure 8. Minimum frequency eigenvalues and eigenmodes for a Gaussian vorticity filament. (a) Eigenvalue as a function of Γ_m for different ξ including theoretical results as $\xi \rightarrow \infty$. The right axis is the frequency squared (4.10a) for $|Ro_{ac}| = 0.5$. Dashed grey lines indicate the upper and lower bounds (4.14). (b) Minimum frequency computed using the Klein–Gordon equation (solid line, 4.10a) and YBJ approximation (dashed line, 4.10b) for $\xi = 10$ and $|Ro_{ac}| = 0.5$. (c) Large Γ_m behaviour of the eigenvalues. Using the delta-function solution $\lambda_0 \Gamma_m$ tends to a constant $-1/3$ (grey dashed line). (d) Structure of the lowest frequency modes for $\xi = 10$ and four values of Γ_m .

As $\Gamma_m \rightarrow 0$, this integral must vanish and the only way this can occur is if $\lambda_{2n} \rightarrow -1$. Therefore, all the modes, including the zeroth mode, will, for vanishingly small Γ_m , achieve the lower bound for the minimum frequency derived earlier. For analytic filament structure functions $\mathcal{F}$, including the Gaussian filament, this also requires $\hat{x}_* \rightarrow 1$ and the region in which the solution is wavelike becomes increasingly localised. Furthermore, in this limit the A_{2n} integral is very large and the zeroth mode has $B_0 \rightarrow \pi/4$ and in the anticyclonic region $\mathcal{X}_0(\hat{x}) = \cos(\pi/4 - \int_{\hat{x}_*}^{\hat{x}} \hat{k} dx')$. One can check that the low horizontal modes become exponentially small in the centre of the filament (D7, D10 and figure 8d). We also note that, as $\Gamma_m \rightarrow 0$, the leading-order correction to the minimum frequency is $O(|Ro_{ac}|)$ (figure 8a) and that there will be $O(|Ro_{ac}|)$ differences between the YBJ approximation to the frequency and the exact value (figure 8b).

In this regime we have made no assumption about the value of ξ , indeed the WKBJ approach works for arbitrary $\partial \hat{V} / \partial \hat{x}$. We only require $\gamma_m \ll 1$. However, since $\xi \gg 1$, in this regime we must have $\Gamma_m \lll 1$. Unless we also have $|Ro_{ac}| \lll 1$, $c_m^2 / f^2 L_x^2 \equiv |Ro_{ac}| \Gamma_m \ll 1$ and, thus, we are in the regime discussed in § 4.1 in which L_x is large compared with both R_m and L_m and the boundary conditions are not important. This is reflected in the fact that the eigenvalues are densely packed near to the minimum and that the projection onto horizontal modes is a very inefficient method for representing the solution (figures 9a and 9d).

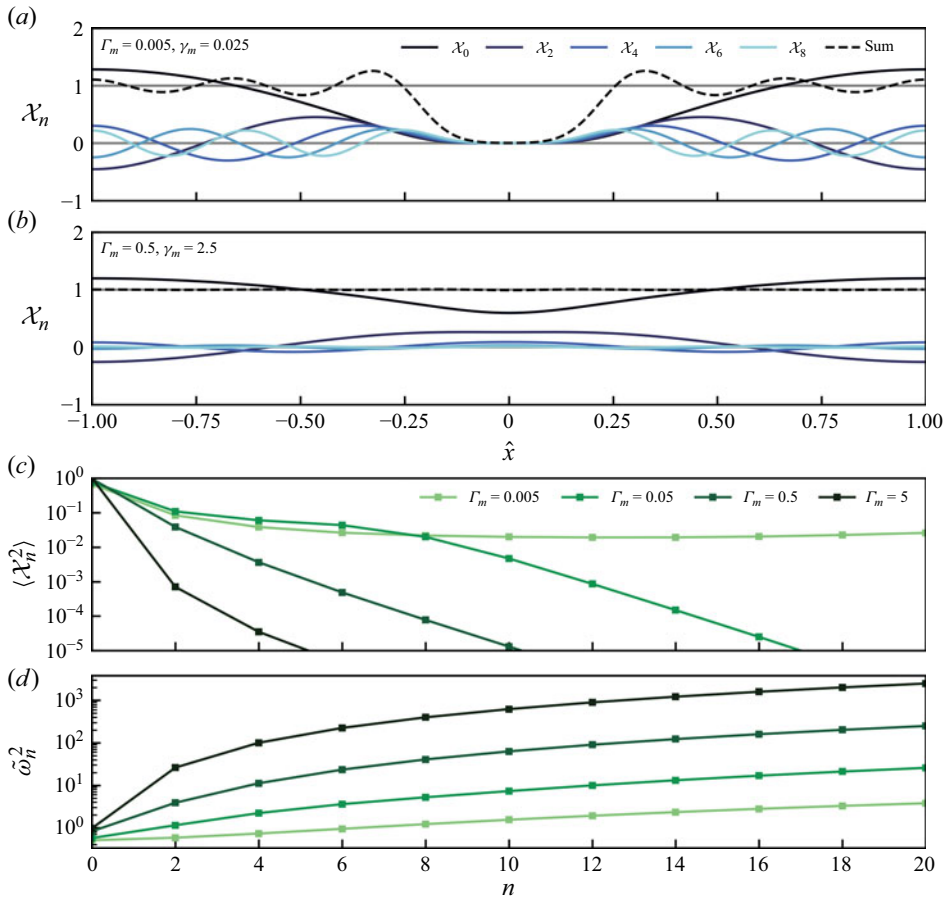


Figure 9. Projection of the uniform initial condition $\mathcal{X} = 1$ onto even horizontal modes for $\xi = 10$ and various Γ_m . The first five horizontal modes and their sum are plotted for points in (a) the WKBJ regime, $\Gamma_m = 0.005 \Rightarrow \gamma_m = 0.1$, and (b) the tunnelling regime, $\Gamma_m = 0.5 \Rightarrow \gamma_m = 10$. (c) Energy content of the even horizontal modes and (d) frequency squared for $|Ro_{ac}| = 0.5$ for four values of Γ_m .

4.3.3. The strong dispersion regime – $\Gamma_m \gg 1$

For large Γ_m , the spatial derivatives in (4.9) – the dispersive term – are multiplied by a large parameter that acts to suppress spatial variations in the zeroth mode. For the higher modes, the large contribution from the spatial derivatives can be balanced by a large eigenvalue. However, the eigenvalue of the zeroth mode is bounded (4.14) and as Γ_m gets larger the zeroth mode becomes more uniform. Young & Ben Jelloul (1997), and later Conn *et al.* (2025), derive a result, known as the ‘strong dispersion approximation’, that carries over to the zeroth mode of the Klein-Gordon equation for arbitrary $\partial \hat{V} / \partial \hat{x}$ in this regime. The idea is to look for a small correction to the uniform state, i.e. the first-order term in an expansion in Γ_m^{-1} .

Let $\mathcal{X}_0 = 1 + \Gamma_m^{-1} \mathcal{X}_0^{(1)} + O(\Gamma_m^{-2})$ and $\lambda_0 = 0 + \Gamma_m^{-1} \lambda_0^{(1)} + O(\Gamma_m^{-2})$. At $O(1)$, (4.9) gives

$$-\frac{\partial^2 \mathcal{X}_0^{(1)}}{\partial \hat{x}^2} + \frac{\partial \hat{V}}{\partial \hat{x}} \cdot 1 = 0. \quad (4.20)$$

Therefore,

$$\frac{\partial \mathcal{X}_0}{\partial \hat{x}} = \Gamma_m^{-1} \hat{V} + O(\Gamma_m^{-1}), \quad \mathcal{X}_0 = 1 - \Gamma_m^{-1} \int^{\hat{x}} \hat{V} d\hat{x}' + O(\Gamma_m^{-2}). \quad (4.21a,b)$$

The leading-order correction to the eigenvalue may be found by considering the secularity condition at $O(\Gamma_m^{-1})$, but also follows from (4.11),

$$\frac{\partial \lambda_0}{\partial \Gamma_m^{-1}} = -\Gamma_m^2 \frac{\partial \lambda_0}{\partial \Gamma_m} = -\Gamma_m^2 \frac{\langle \mathcal{X}_0'^2 \rangle}{\langle \mathcal{X}_0^2 \rangle} = -\langle \hat{V}^2 \rangle + O(\Gamma_m^{-1}). \quad (4.22a)$$

Thus,

$$\lambda_0 = -\langle \hat{V}^2 \rangle \Gamma_m^{-1} + O(\Gamma_m^{-2}). \quad (4.22b)$$

Unwrapping the non-dimensionalisation, the leading-order correction to the frequency is given by

$$\omega_0^2 = f^2 (1 + c_m^{-2} \langle \hat{V}^2 \rangle). \quad (4.23)$$

We have recovered Young & Ben Jelloul (1997)'s result that the leading-order correction is proportional to the average kinetic energy of the background flow.

The leading-order correction to the minimum frequency is $O(\Delta V^2/c_m^2 \equiv \alpha_m^2)$. Similar to the $\gamma_m \rightarrow \infty$ limit in § 3, we find that in the $\Gamma_m \rightarrow \infty$ limit the parameter determining the leading-order behaviour is α_m , which does not depend on the length scales of the background flow. However, unlike in § 3 we cannot look at the distinguished limit $\Gamma_m \rightarrow \infty$ with α_m held fixed as this would require $|Ro_{ac}| \rightarrow \infty$. This is unphysical as it results in an inertially unstable background flow (for $|Ro_{ac}| > 1$, the minimum frequency squared can be negative (4.10a)). This limitation means that the leading-order correction to the minimum frequency will always be small in the strong dispersion regime and, hence, the YBJ approximation will always be good.

4.4. Delta-function filament

So far we have considered the WKBJ regime, $\gamma_m \equiv (1/2)\xi \Gamma_m \ll 1$, and the minimum frequency mode in the strong dispersion regime, $\Gamma_m \gg 1$. By construction $\xi \gg 1$, but otherwise these regimes place no restriction on the value of ξ . Indeed these arguments hold with very few restrictions on $\partial \hat{V}/\partial \hat{x}$. We now consider a delta-function filament, which requires $L_f \ll L_m$ ($\gamma_m \gg 1$) and $L_f \ll L_x$ ($\xi \gg 1$) but places no restriction on the ratio of L_m to L_x . The delta-function filament is therefore reached through the distinguished limit $\xi \rightarrow \infty$ with Γ_m held fixed. The WKBJ, strong dispersion and delta-function regimes are summarised in table 5.

Away from the filament the even modes have the form

$$\mathcal{X}_{2n} = \cos K_{2n}(1 - \hat{x}) \quad (4.24a)$$

with

$$\lambda_{2n} = -1 + \Gamma_m K_{2n}^2. \quad (4.24b)$$

The jump condition is $\Gamma_m \mathcal{X}_n' = \mathcal{X}_n$ at $\hat{x} = 0^+$. Therefore, K_{2n} satisfies

$$K_{2n} \tan K_{2n} = \Gamma_m^{-1}, \quad (4.24c)$$

Regime	Limits	Comments
WKBJ regime	$L_m \ll L_f \ll L_x$	Wave behaviour determined by the local vorticity of the flow. Normal mode decomposition is a very inefficient representation of the solution.
Strong dispersion regime	$L_f \ll L_x \ll L_m$	The strong dispersion approximation may be applied to the zeroth mode. Frequency is determined by the kinetic energy of the background flow.
Delta-function regime	$L_f \ll L_x, L_f \ll L_m$	Admits an analytic solution with no restriction on $\Gamma_m := 2L_m/L_x$.

Table 5. Summary of the different regimes considered in the periodic problem.

where K_{2n} is the unique solution in $(n\pi, (n+1/2)\pi)$. For any given filament, the delta-function approximation will fail for sufficiently large n when the length scale of the waves becomes too short.

We now consider the limiting behaviours as $\Gamma_m \rightarrow 0$ and $\Gamma_m \rightarrow \infty$. For $\Gamma_m \ll 1$, let $K_{2n} = (n + (1/2))\pi - \Gamma_m k_{2n} + O(\Gamma_m^2)$. With this expansion, $K_{2n} \tan K_{2n} = (n + (1/2))\pi/(\Gamma_m k_{2n}) + O(1)$. Therefore, as $\Gamma_m \rightarrow 0$,

$$K_{2n} = \left(n + \frac{1}{2}\right)\pi - \Gamma_m \frac{1}{\left(n + \frac{1}{2}\right)\pi} + O(\Gamma_m^2), \quad \lambda_{2n} = -1 + \Gamma_m \left(n + \frac{1}{2}\right)^2 \pi^2 + O(\Gamma_m^2). \quad (4.25a,b)$$

In particular, $\lambda_0 \rightarrow -1$ as $\Gamma_m \rightarrow 0$ and, thus, even in the delta-function limit we can achieve the lower bound on the minimum frequency (figure 8a). However, to be in the delta-function limit $\gamma_m \gg 1$ when $\Gamma_m \ll 1$ we must have $\xi \gg 1$. Furthermore, there are differences in the structure of the eigenmodes in the WKBJ and delta-function regimes as $\Gamma_m \rightarrow 0$. For example, consider the phase of the zeroth mode as it enters the filament. In the delta-function solution, as $\hat{x} \rightarrow 0$, $\mathcal{X}_0 \rightarrow \cos \pi/2 = 0$. Whereas in the WKBJ solution, at $\hat{x} = \hat{x}_*$, $\mathcal{X}_0 \sim \cos B_0 = \cos \pi/4$, which represents a phase shift of $\pi/4$.

For large Γ_m , we need different expansions for $n=0$ and $n \geq 1$, highlighting that the zeroth mode is special. First, $n \geq 1$, where we let $K_{2n} = n\pi + \Gamma_m^{-1}k_{2n} + O(\Gamma_m^{-2})$. Working through the expansion, we find that, for $\Gamma_m \gg 1$, $n \geq 1$,

$$K_{2n} = n\pi + \Gamma_m^{-1} \frac{1}{n\pi} + O(\Gamma_m^{-2}), \quad \lambda_{2n} = \Gamma_m n^2 \pi^2 + 1 + O(\Gamma_m^{-1}). \quad (4.26a,b)$$

Note that the eigenvalues are $O(\Gamma_m)$ and thus large. To leading order, the frequency is given by

$$\omega_{2n}^2 = n^2 \pi^2 |Ro_{ac}| \Gamma_m = n^2 \pi^2 \frac{c_m^2}{f^2 L_x^2}, \quad (4.26c)$$

which is independent of the properties of the background flow with the exception of the length scale L_x . Here L_x is setting the length scale of the waves $L_w = L_x/n\pi$ and the frequency is determined by the wave Burger number $(c_m/fL_w)^2$.

For, $n=0$, we expand K_0^2 in powers of Γ_m^{-1} , $K_0^2 = \Gamma_m^{-1}a_1 + \Gamma_m^{-2}a_2 + O(\Gamma_m^{-3})$. We have $K_0 \tan K_0 = K_0^2 + K_0^4/3 + O(K_0^6) = \Gamma_m^{-1}a_1 + \Gamma_m^{-2}(a_2 + a_1^2/3) + O(\Gamma_m^{-3})$. Therefore, for $\Gamma_m \gg 1$,

$$K_0^2 = \Gamma_m^{-1} - \frac{1}{3}\Gamma_m^{-2} + O(\Gamma_m^{-3}), \quad \lambda_0 = -\frac{1}{3}\Gamma_m^{-1} + O(\Gamma_m^{-2}). \quad (4.27a,b)$$

This not only achieves the upper bound $\lambda_0 \rightarrow 0$ as $\Gamma_m \rightarrow \infty$ (figure 8a), but we observe that the wavenumber $K_0 \rightarrow 0$ as well. The solution is tending towards a uniform state. Moreover, the coefficient $-1/3$ in (4.27a,b) is consistent with the strong dispersion approximation, i.e. for the sawtooth geostrophic velocity $\hat{V} = \text{sign}(\hat{x})(1 - \hat{x})$ associated with the delta-function filament, we have $\langle \hat{V}^2 \rangle = 1/3$.

The utility of the delta-function filament is not only that it provides an analytical result for intermediate values of Γ_m but also that it is a good model for sharp submesoscale filaments. Solving the eigenvalue problem numerically for the Gaussian filament, we observe very good agreement between the eigenvalues computed for $\xi = 10, 20$ and the theoretical values assuming a delta-function (figure 8a). This is important as these values are realisable in submesoscale flows, for example, taking $|Ro_{ac}| = 0.5$ and $\xi = 10$, the peak vorticity is only $\partial V / \partial x = 3.5f$, and thus, the delta-function filament is a relevant model for oceanic applications.

4.5. Lateral refraction

In this problem we are considering a discrete spectrum of horizontal modes. Lateral refraction and the associated horizontal flux of NIW energy are determined by the time scale over which the different modes dephase, the structure of each mode and their relative energy content. To understand this process, we look at how the laterally uniform initial condition $\mathcal{X} = 1$ projects onto horizontal modes. Conveniently, the Sturm–Liouville problem has a trivial weight function and, thus, the projection onto horizontal modes satisfies a particularly simple form of Parseval’s theorem,

$$\langle \mathcal{X}^2 \rangle = 1 = \sum_n \langle \mathcal{X}_n^2 \rangle, \quad (4.28)$$

where $\langle \cdot \rangle$ again denotes a lateral average. For a given vertical mode, $\mathcal{X}$ is proportional to the across-filament velocity and we may interpret $\langle \mathcal{X}_n^2 \rangle$ as the fraction of the across-filament kinetic energy contained in the n th horizontal mode. In the strong dispersion regime, $\Gamma_m \gg 1$, the initial condition projects almost entirely onto the zeroth mode. The kinetic energy in the higher modes drops off exponentially (figures 9b and 9c) and, hence, the solution (e.g. figure 7h) remains nearly uniform even as the different modes dephase. Even for smaller Γ_m , the zeroth mode dominates the energy content. For example, in the case $\Gamma_m = 0.005$, $\xi = 10$, more than 65 % of the energy is in the zeroth mode. However, many horizontal modes are required to reproduce the initial condition (figures 9a and 9c).

The time scale over which the different horizontal modes dephase is determined by the differences in their frequencies (figure 9d). While these differences are strongly dependent on Γ_m and $|Ro_{ac}|$ (differences in $\tilde{\omega}_n^2$ scale linearly with $|Ro_{ac}|$), the refraction time scales can be of the order of an inertial period (e.g. figure 7b). This is much faster than the refraction time scales at small Rossby numbers (e.g. Balmforth *et al.* 1998; Danioux *et al.* 2015; Rocha, Wagner & Young 2018; Asselin *et al.* 2020), as one might expect given the large Rossby numbers in the filament we are considering. This is an important result because in the submesoscale regime the background flows also evolve on inertial time scales. However, we find that ζ -refraction can occur even faster. The rapid lateral refraction and their domination of the energy content means that the zeroth modes dictate the long time dynamics of the NIWs. With this in mind we now consider how the modes propagate in the vertical.

5. Vertical propagation

So far we have been solving a one-dimensional problem and have not needed to concern ourselves with the structure of the stratification or the vertical modes as the lateral dynamics of a particular vertical mode are only influenced by the mode speed c_m . However, the vertical propagation of NIW energy does require knowledge of the vertical structure. Here, we use uniform stratification and assume the NIWs have a plane wave structure with vertical wavenumber k_z for which the mode speed is $c_m = N/k_z$. The plane wave assumption ignores the role of the boundaries. In particular, the group velocity arguments we make assume that ω is a function of the continuous variable k_z , whereas the imposition of boundary conditions implies discrete vertical modes. These arguments are valid if the discrete modes sufficiently resolve $\omega(k_z)$. In practice, this generally means that the high mode dynamics are well described by plane wave theory but that some caution is required when considering the lowest modes.

5.1. Group velocity

5.1.1. General expressions

With these choices we can consider the vertical group velocity of a horizontal mode:

$$c_{g,n} := \frac{\partial \omega_n}{\partial k_z}. \quad (5.1)$$

However, we can use $\Gamma_m \equiv c_m^2/(f^2 L_x^2 |Ro_{ac}|)$, (4.10a) and (4.11) to express the group velocity as

$$c_{g,n} \equiv \frac{\partial \Gamma_m}{\partial k_z} \frac{\partial \omega_n}{\partial \Gamma_m} \equiv \frac{-2N^2 |Ro_{ac}|^{-1}}{k_z^3 L_x^2 f^2} \frac{1}{2\omega_n} \frac{\partial \omega_n^2}{\partial \Gamma_m} \equiv \frac{-N^2}{\omega_n k_z^3} \frac{1}{L_x^2} \frac{\partial \lambda_n}{\partial \Gamma_m} \equiv -\frac{N^2}{\omega_n k_z^3} \frac{1}{L_x^2} \frac{\langle \mathcal{X}_n'^2 \rangle}{\langle \mathcal{X}_n^2 \rangle}. \quad (5.2)$$

This final expression is reminiscent of the ray-tracing expression for the group velocity and would be the same if we replaced $\langle \mathcal{X}_n'^2 \rangle / \langle \mathcal{X}_n^2 \rangle$ by $(k_x L_x)^2$, where k_x is the horizontal wavenumber. Furthermore, this expression for the group velocity holds for arbitrary $\partial \hat{V} / \partial \hat{x}$. One of the key motivations for studying ζ -refraction is the intuition that it is necessary to shrink the horizontal scales of NIWs in order to allow them to propagate rapidly (Gill 1984; Balmforth *et al.* 1998). This intuition is based upon the plane wave dispersion relation and group velocity but we can see that it holds more generally: the generation of horizontal gradients is critical to the vertical propagation of NIWs.

This expression also holds under the YBJ approximation if we replace ω_n by f ,

$$c_{g,n}^{(YBJ)} \equiv -\frac{N^2}{f k_z^3} \frac{1}{L_x^2} \frac{\langle \mathcal{X}_n'^2 \rangle}{\langle \mathcal{X}_n^2 \rangle}. \quad (5.3)$$

As a result, YBJ theory always underpredicts the group velocity of the minimum frequency mode, since $\omega_0 < f$, but overpredicts the group velocity of the super-inertial high modes.

Here we ask: For a given background flow, what wavenumber maximises the magnitude of the vertical group velocity? It is natural to non-dimensionalise and express the group velocity in terms of Γ_m , $|Ro_{ac}|$ and ξ . For given $|Ro_{ac}|$ and ξ , we then find the optimal Γ_m . The non-dimensional group velocity is given by

$$\hat{c}_{g,n} := \frac{N}{f} \frac{c_{g,n}}{f L_x} = -\frac{c_m^3}{f^3 L_x^3} \frac{1}{\tilde{\omega}_n} \frac{\langle \mathcal{X}_n'^2 \rangle}{\langle \mathcal{X}_n^2 \rangle} = -|Ro_{ac}|^{3/2} \tilde{\omega}_n^{-1} \Gamma_m^{3/2} \frac{\langle \mathcal{X}_n'^2 \rangle}{\langle \mathcal{X}_n^2 \rangle}. \quad (5.4)$$

In this expression, there is explicit dependence on both $|Ro_{ac}|$ and Γ_m . The contribution from the horizontal structure

$$\mathcal{S}_n := \frac{\langle \mathcal{X}'_n{}^2 \rangle}{\langle \mathcal{X}_n{}^2 \rangle} \quad (5.5)$$

depends only on Γ_m and ξ , but not $|Ro_{ac}|$. The frequency term depends on all three dimensionless parameters. However, under the YBJ approximation the frequency dependence drops out,

$$\hat{c}_{g,n}^{(YBJ)} := -\frac{N}{f} \frac{c_{g,n}^{(YBJ)}}{f L_x} = -|Ro_{ac}|^{3/2} \Gamma_m^{3/2} \frac{\langle \mathcal{X}'_n{}^2 \rangle}{\langle \mathcal{X}_n{}^2 \rangle}. \quad (5.6)$$

Ignoring the frequency term for the moment, we see that Γ_m has two roles in determining the group velocity. First, there is the explicit dependence on $\Gamma_m^{3/2}$ and second, it determines the structural contribution $\mathcal{S}_n$.

5.1.2. Delta-function filament

For the delta-function filament, we can compute $\mathcal{S}_{2n}$ exactly using (4.24a),

$$\mathcal{S}_{2n} := \frac{\langle \mathcal{X}'_{2n}{}^2 \rangle}{\langle \mathcal{X}_{2n}{}^2 \rangle} = K_{2n}^2 \frac{\int_0^1 \sin^2 K_{2n}(1-\hat{x}) d\hat{x}}{\int_0^1 \cos^2 K_{2n}(1-\hat{x}) d\hat{x}} = K_{2n}^2 \frac{1 - \frac{1}{2K_{2n}} \sin 2K_{2n}}{1 + \frac{1}{2K_{2n}} \sin 2K_{2n}}. \quad (5.7)$$

Furthermore, we can use the asymptotic expansions for K_{2n} and $\tilde{\omega}_n$, in the limit of small and large Γ_m , derived in § 4.4 to compute $\mathcal{S}_{2n}$ and $\hat{c}_{g,2n}$ in these limits.

However, in all but one case the leading-order term in the expansion for K_{2n} is a constant. Consequently, $\mathcal{S}_{2n}$ also tends to a constant. For example, as $\Gamma_m \rightarrow 0$, $K_0 \rightarrow \pi/2$ and, thus, $\mathcal{S}_0 \rightarrow \pi^2/4$. In general, as $\Gamma_m \rightarrow 0$, $\mathcal{S}_{2n} \rightarrow (n + (1/2))^2 \pi^2$. Moreover, $\tilde{\omega}_n \rightarrow \sqrt{1 - |Ro_{ac}|}$ and, hence, $\hat{c}_{g,n} \sim \Gamma_m^{3/2}$ (figure 10a). In the limit of small Γ_m we have recovered the familiar k_z^{-3} vertical wavenumber dependence of ray-tracing.

The limit $\Gamma_m \rightarrow \infty$ is more interesting. For $n \geq 1$ (4.26a), $K_{2n} \rightarrow n\pi$ and, thus, $\mathcal{S}_{2n} \rightarrow n^2 \pi^2$. However, in this case the leading-order contribution to the frequency also depends on Γ_m , i.e.

$$\tilde{\omega}_n \rightarrow n\pi |Ro_{ac}|^{1/2} \Gamma_m^{1/2} \quad (5.8a)$$

and, thus, the group velocity is only linear in $\Gamma_m \sim k_z^{-2}$ (figure 10a),

$$\hat{c}_{g,n} \rightarrow -n\pi |Ro_{ac}| \Gamma_m \equiv -n\pi \frac{c_m^2}{f^2 L_x^2}. \quad (5.8b)$$

Finally, consider the zeroth mode in the strong dispersion regime $\Gamma_m \rightarrow \infty$. In this limit, $K_0 \rightarrow \Gamma_m^{-1/2} \Rightarrow \mathcal{S}_0 \rightarrow (1/3) K_0^4 \rightarrow (1/3) \Gamma_m^{-2}$ (figure 10b). Since $\tilde{\omega}_0 \rightarrow 1$, we have

$$\hat{c}_{g,0} \rightarrow -\frac{1}{3} |Ro_{ac}|^{3/2} \Gamma_m^{-1/2}. \quad (5.9)$$

In the strong dispersion regime the contribution from the horizontal structure is very sensitive to the vertical wavenumber to the extent that the group velocity is proportional to k_z .

This result for the zeroth mode in the strong dispersion regime generalises to arbitrary periodic background flows through the strong dispersion approximation. From (4.22a) we

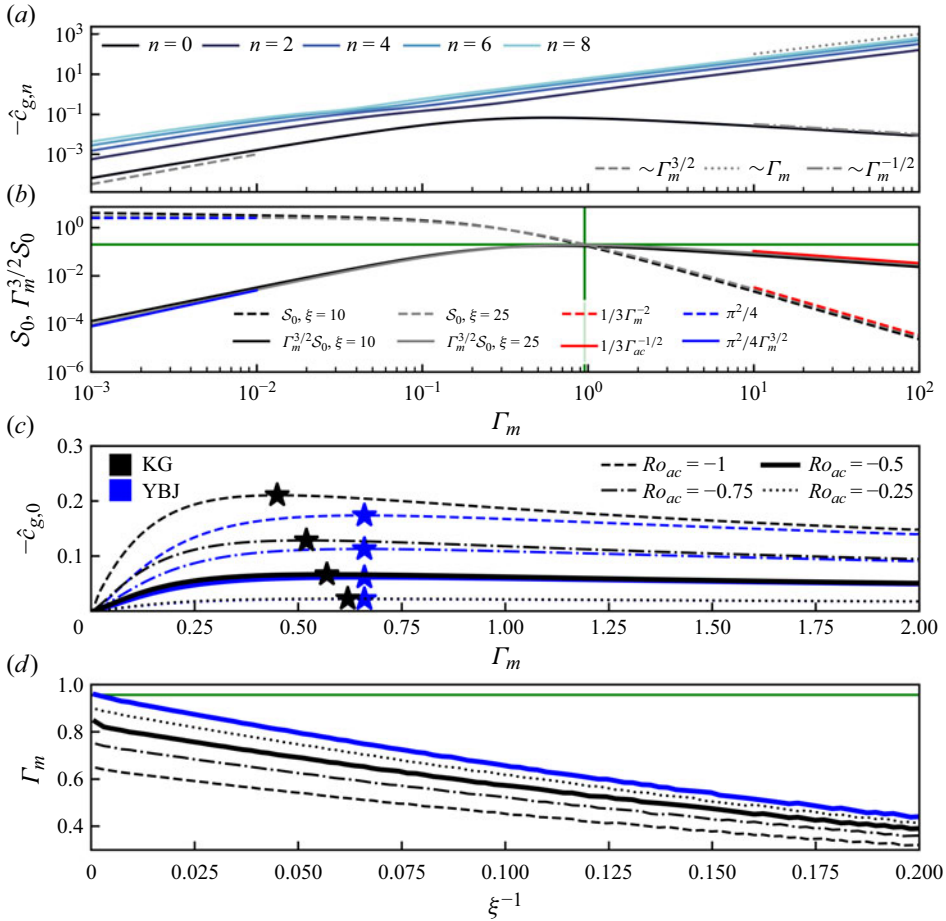


Figure 10. (a) Non-dimensionalised group velocity of the first five even modes for the Gaussian filament with $Ro_{ac} = -0.5$ and $\xi = 10$. (b) Horizontal structure contribution S_0 to the group velocity of the zeroth mode for $\xi = 10, 25$ with asymptotic scalings for the delta-function filament in red and blue. (c) Non-dimensionalised group velocity of the zeroth mode for $\xi = 10$ and four values of Ro_{ac} . Both the Klein–Gordon (5.2, black) group velocity and the YBJ approximation (5.3, blue) are plotted. The maxima of the group velocities are denoted by stars. (d) Optimal value of Γ_m, Γ_m^* , for radiating NIW energy as a function of ξ^{-1} for four values of Ro_{ac} . Under the YBJ approximation the optimal value is independent of Ro_{ac} . Green lines in (b) and (d) denote $\Gamma_m = 0.956$, the location of the maximum of $\Gamma_m^{3/2} S_0$ for the delta-function filament.

have $S_0 \rightarrow \Gamma_m^2 \langle \hat{V}^2 \rangle$ and, thus,

$$\hat{c}_{g,0} \rightarrow -|Ro_{ac}|^{3/2} \langle \hat{V}^2 \rangle \Gamma_m^{-1/2}. \quad (5.10a)$$

Redimensionalising,

$$c_{g,0} \rightarrow -\frac{f \langle \hat{V}^2 \rangle}{N c_m} \equiv -\frac{f \langle \hat{V}^2 \rangle}{N^2} k_z, \quad (5.10b)$$

consistent with Young & Ben Jelloul (1997). We again emphasise that under the strong dispersion approximation the dynamics are independent of the length scales of the background flow and that results derived using the YBJ approximation extend to the Klein–Gordon equation as the zeroth mode frequency tends to f .

5.1.3. Optimal radiation

The special behaviour of the zeroth mode, i.e. that the group velocity magnitude becomes a decreasing function of Γ_m for large Γ_m , means that there exists an optimal value of Γ_m for radiating energy, generalising the findings of Balmforth *et al.* (1998). We call this value Γ_m^* . Thus, for a given background flow, i.e. fixed values of ΔV , L_f , L_x , N and f , the optimal vertical wavenumber for radiating NIW energy is

$$k_z^* := \sqrt{\frac{2N^2}{\Delta V f L_x} \frac{1}{\Gamma_m^*}}. \quad (5.11)$$

Under the YBJ approximation, the group velocity (5.3) factorises into a term depending on $|Ro_{ac}|$ and a term depending on Γ_m (and ξ) but not $|Ro_{ac}|$. As a result, Γ_m^* is independent of $|Ro_{ac}|$ (figures 10c and 10d) and occurs at the maximum of $\Gamma_m^{3/2} S_0$. For the delta-function filament, this maximum occurs at $\Gamma_m^* = 0.956$ with value $\Gamma_m^{3/2} S_0 = 0.20$ (figure 10b).

Using the Klein-Gordon group velocity expression with frequency dependence (5.2) not only increases the group velocity of the zeroth mode but also shifts the location of the maximum to smaller Γ_m (figures 10c and 10d). Furthermore, at finite ξ , Γ_m^* is smaller than the delta-function value (figure 10d). For the Gaussian filament with $\xi = 10$ and $Ro_{ac} = -1$, Γ_m^* is approximately half the delta-function value of 0.956.

The non-monotonic Γ_m dependence of the zeroth mode group velocity and the subsequent existence of an optimal Γ_m for vertically radiating NIW energy mirrors the γ_m dependence of the decay time scale of the lateral radiation problem in § 3. Both problems highlight the importance of $O(1)$ parameter regimes that are not easily tackled by asymptotic approaches. Notably, this non-monotonic behaviour only occurs for the zeroth horizontal mode. However, as we argued in § 4.5, for uniform initial conditions, this mode dominates the energetics even for small Γ_m . Moreover, the higher horizontal modes have larger group velocities since S_n is larger and so they propagate away more quickly leaving the zeroth mode behind. We now demonstrate this explicitly with 2-D linear simulations.

5.2. Linear simulations

We run 2-D linear simulations studying the evolution of a slab initial condition in the across-filament velocity,

$$u_i := \frac{1}{2} \left(1 + \operatorname{erf} \left(3 + \frac{2z}{H} \right) \right). \quad (5.12)$$

We project u_i onto cosine modes and evolve each vertical mode, with the exception of the barotropic mode that we discard, according to the Klein–Gordon equation. For the background flow, we use the Gaussian filament with $Ro_{ac} = -0.5$ and $\xi = 10$. We take the depth of the domain to be $L_z = 0.1L_x$ and $H = 2 \times 10^{-3}L_x$. We use 512 vertical modes and 512 horizontal grid points. With u expressed in cosine modes, it is simple to compute w from continuity and then v and b from (2.1b,d).

In the initial condition the energy content of each vertical mode is a decreasing function of m , but the shear is maximal at $k_z \approx 1/H$ (figure 11a). We define bulk parameters Γ_H and γ_H , analogous to Γ_m and γ_m , using the shear scale H by

$$\Gamma_H := \frac{1}{|Ro_{ac}|} \frac{N^2 H^2}{f^2 L_x^2}, \quad \gamma_H := \frac{1}{2} \xi \Gamma_H. \quad (5.13a,b)$$

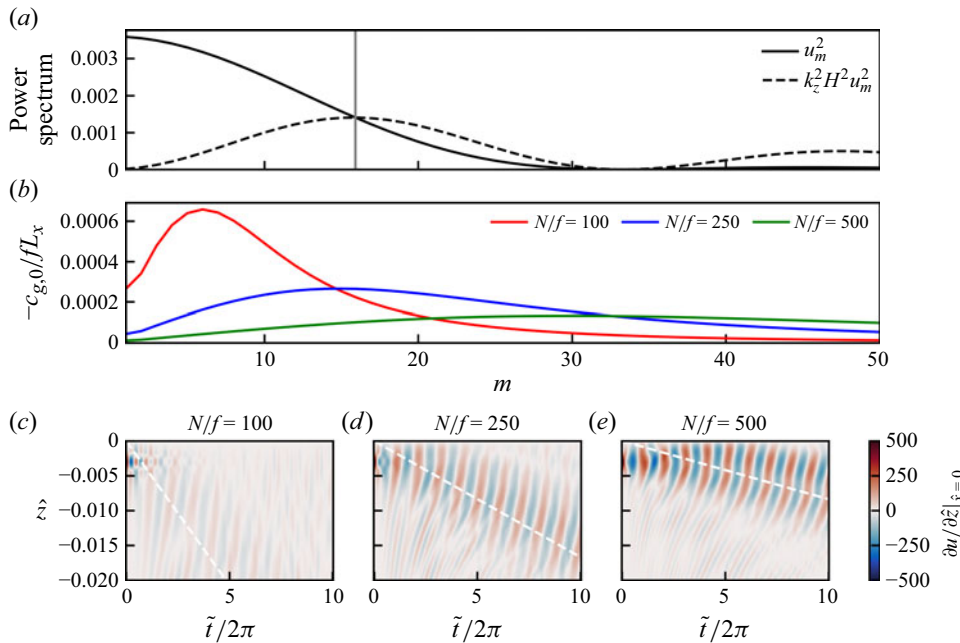


Figure 11. (a) Power spectrum of the initial across-filament velocity (5.12) and vertical shear as a function of the vertical mode number. The vertical grey line indicates $k_z = 1/H$. (b) Theoretical prediction of the group velocity (5.2) of the zeroth horizontal mode as a function of the vertical mode number for the three different values of N/f simulated. (c–e) Across-filament vertical shear at the centre of the filament. In (c–e) the x axes are in inertial periods and the y axes are $\hat{z} := z/L_x$. White lines indicate a ray moving at the theoretical maximum group velocity, i.e. the maxima of (b).

N/f	Γ_H	γ_H
100	0.08	0.4
250	0.5	2.5
500	2	10

Table 6. Two-dimensional linear simulation parameters.

These bulk parameters best describe the shear containing modes but most of the energy is contained in lower modes with $\Gamma_m > \Gamma_H$ and $\gamma_m > \gamma_H$ (figure 11a). Having fixed the value of H/L_x and $|Ro_{ac}|$, we vary Γ_H by changing the strength of the stratification, i.e. N/f . We focus on three cases, $N/f = 100$, 250 and 500, which respectively give $\Gamma_H = 0.08$, 0.5 and 2 (table 6). These particular choices of N/f are inspired by the conditions in the northern Gulf of Mexico and are thus towards the large end of the range of realistic oceanic values, at least away from the low latitudes. However, the ratio N/f has a limited independent value outside of the parameters Γ_H and γ_H . We could achieve the same Γ_H and γ_H values for different stratifications by adjusting the value of H/L_x . We discuss the range of realistic values of Γ_H and γ_H in § 6. That being said, N/f does have some independent value in that it is used to non-dimensionalise the group velocity (5.4). Consequently, if we consider two cases with the same Γ_H but different stratifications then the NIW energy will propagate more rapidly in the less stratified case.

An animation of the evolution of u in the three simulations (see supplementary movie 1 available at <https://doi.org/10.1017/jfm.2025.10637>) demonstrates the two major effects of varying the stratification. Firstly, in the $N/f = 100$ case we observe beams radiating down and out from the centre of the filament. This classical behaviour of inertia-gravity waves is indicative of high frequency waves radiating along rays and occurs because at low γ_m , where the zeroth horizontal mode is strongly evanescent in the filament, the initial condition projects broadly onto horizontal modes. After the beams have propagated away, we are left with the slower propagating low horizontal modes – primarily the zeroth mode. The beams are not observed in the higher stratification cases where the projection onto higher horizontal modes is much weaker, although a weak signal from the other low horizontal modes, mostly mode 2, can be observed. The stronger tunnelling in the higher stratification cases is also evidenced by the greater shear amplitude in the slowly propagating wavepacket at the centre of the filament (figure 11c–e). A higher frequency wavepacket associated with the second horizontal mode is also present in figures 11(d) and 11(e).

The second observed effect is the slower propagation of the zeroth mode as the stratification increases. Plotting the group velocity (5.2) as a function of vertical mode number, which is proportional to k_z , we see that the location of the maximum group velocity increases with N/f since the dispersivity is proportional to $(N/f k_z)^2$ (figure 11b). Furthermore, the magnitude of the maximum group velocity decreases, which follows from the explicit dependence of the group velocity on N/f . In other words, increasing N/f is equivalent to stretching the group velocity–mode number curve along the x axis while compressing it down the y axis. The result is that, for the energy containing low vertical modes, the magnitude of the group velocity decreases with increasing N/f . This counter-intuitive dependence on stratification in the strong dispersion regime (small Γ_m), first noted by Balmforth *et al.* (1998), is the opposite of the behaviour one would predict by naively applying ray-tracing. If the energy were contained in the high modes then we could apply ray-tracing and we would observe the opposite behaviour – at high vertical wavenumbers the group velocity is an increasing function of N/f (figure 11b). In addition, the slices of the across-filament shear at the filament centre (figure 11c–e) demonstrate vertically coherent wavepackets radiating at the maximum group velocity. The maximum is a stationary point of the group velocity and, hence, wavepackets propagating at the maximum group velocity are weakly dispersive and able to maintain a coherent spatial structure.

6. Discussion

In this paper we explored the interaction of NIWs with strongly cyclonic vorticity filaments to which YBJ theory does not necessarily apply. First, we considered the lateral radiation of a single vertical mode from a cyclonic filament and developed scaling for the decay of the NIWs in the filament as functions of the tunnelling parameter γ_m and filament Rossby number Ro_f . Then we considered the case of a cyclonic filament in an otherwise anticyclonic flow in a finite width domain. The problem was approached via a normal mode decomposition with a focus on the zeroth mode, which has unique behaviour and, for an initially uniform velocity field, dominates the solution. Finally, the vertical propagation of the waves was considered by deriving a generic expression for the group velocity of each normal mode that highlights the importance of gradients in the wave field. A fruitful approach throughout the paper is to model the sharp filaments as delta-functions. We find this to be a good model for large yet realistic values of the Rossby number.

A recurring theme of the paper is the vastly differing behaviour of the waves as the mode speed c_m , which depends on the stratification and vertical wavelength of the waves, is varied. The key consideration is how the length scale $L_m := c_m^2/f\Delta V$ compares to the length scales of the background flow. For waves with short vertical wavelengths such that the tunnelling parameter $\gamma_m := L_m/L_f$ is small, WKB and ray-tracing ideas may be applied. Only the local properties of the background flow determine the wave behaviour and increasing the vertical wavelength leads to more rapid propagation both horizontally and vertically. However, if the vertical wavelength is large such that the length scale L_m is larger than the length scales of the background flow then increasing the vertical wavelength further leads to less rapid radiation as dispersive effects act to smooth the response to the filament. In the periodic problem this smoothing effect only applies to the zeroth horizontal mode. But in the strong dispersion regime, $\Gamma_m := 2L_m/L_x \gg 1$, the zeroth mode increasingly dominates the solution assuming a uniform initial condition. When L_m is much larger than length scales of the background flow, i.e. in the limits $\gamma_m \rightarrow \infty$, $\Gamma_m \rightarrow \infty$, the dynamics become independent of the length scales of the background flow. Indeed, the strong dispersion approximation shows that the only property of the background flow that matters as $\Gamma_m \rightarrow \infty$ is the average kinetic energy.

The strongest radiation occurs when L_m is comparable to the length scales of the background flow. The most efficient lateral radiation out of the filament occurs for $O(1)$ values of γ_m . However, if we are more interested in the vertical propagation of NIW energy then the key parameter is Γ_m because the uniform initial condition projects mostly onto the zeroth mode even in the WKB regime. The differing dynamics for small and large Γ_m means that the group velocity of the zeroth horizontal is maximised at some intermediate $\Gamma_m^* = O(1)$ value.

The importance of both the vertical wavelength and stratification for the propagation of NIWs was recently highlighted in Thomas *et al.* (2024) and is a particularly important consideration in the submesoscale regime where the small horizontal length scales of the background flow compared with the mesoscale mean that Γ_m (and γ_m) will generally be larger and any intuition honed on ray-tracing may be found wanting. Since these parameters depend on the vertical wavelength of the waves in addition to the stratification, velocity and length scales of the background flow a vast range of values can be found in the ocean. In particular, by allowing the vertical wavelength to tend to zero we can make Γ_m arbitrarily small. Conversely, in very weak background flows Γ_m can be very large. It is therefore necessary for any given problem to compute for what vertical wavelengths Γ_m is large and for what vertical wavelengths it is small. For example, if we take some reasonable mid-latitude values for the background submesoscale flow $N^2 = 10^{-5} \text{ s}^{-2}$, $f = 10^{-4} \text{ s}^{-1}$, $\Delta V = 10^{-1} \text{ m s}^{-1}$ and $L_x = 10^4 \text{ m}$, then we find that $\Gamma_m = 1$ for $k_z = 10^{-2} \text{ m}^{-1}$. Therefore, the lowest modes with $k_z \approx 10^{-3} \text{ m}^{-1}$ will have $\Gamma_m \approx 10^2$ and will fall into the strong dispersion regime but the high modes will have small Γ_m . One should also expect to find that the fastest radiating waves have vertical wavelengths of a few hundred metres.

Finally, we comment on some of the physics not included in this study. The set-ups considered here were designed to isolate and focus upon the effects of vorticity. The 2-D set-up is convenient in that it eliminates advective effects and makes the problem tractable. However, even though we demonstrated that ζ -refraction can be a very fast process, advective effects are likely to play an important role at the submesoscale. Furthermore, while Asselin *et al.* (2020) found that strain was remarkably unimportant in the quasi-geostrophic barotropic weak dispersion regime ($\Gamma_m \ll 1$), their results do not extend to the submesoscale regime. The other advantage of the 2-D set-up is that barotropic instability is excluded. The vorticity structures that we considered are

unstable to barotropic instability according to the Rayleigh–Kuo criterion (Kuo 1949). Nevertheless, such vorticity structures are observed in the ocean and are sustained through frontogenetic processes. Thomas (2019) found that the secondary circulations associated with frontogenesis are themselves able to enhance the vertical radiation of NIWs through a differential vertical Doppler shift. Lastly, the effects of vertical variations in vorticity, and baroclinic background flows more generally, were not considered here but are the subject of ongoing work.

Supplementary movie. Supplementary movie is available at <https://doi.org/10.1017/jfm.2025.10637>.

Acknowledgements. The authors would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme Anti-diffusive dynamics: from sub-cellular to astrophysical scales, where work on this paper was undertaken. Some of the computing was performed on the Sherlock cluster. The authors would like to thank Stanford University and the Stanford Research Computing Center for providing these computational resources. Finally, the authors would like to thank William R. Young and two anonymous reviewers for their comments that greatly improved the paper.

Funding. J.P.H. and L.N.T. were supported by a grant from the National Science Foundation, award number OCE-1851450. J.R.T. was supported by a grant from the Natural Environment Research Council, grant number NE/T004223/1. Work at the Isaac Newton Institute was supported by EPSRC grant EP/R014604/1.

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

Appendix A. Derivation of the ray-tracing equations

A.1. Fixed Ro_f

Starting from (3.7a) with Ro_f fixed, we derive the ray-tracing equations by making a WKB approximation in the small parameter $\epsilon := \sqrt{\gamma_m}$. Introducing a rescaled time $\tilde{t} := \sqrt{Ro_f} \epsilon \tilde{t} \equiv (c_m/L_f)t$ we have

$$\left[\epsilon^2 \left(\frac{\partial^2}{\partial \tilde{t}^2} - \frac{\partial^2}{\partial \tilde{x}^2} \right) + Ro_f^{-1} + \mathcal{F}(\tilde{x}) \right] \mathcal{X} = 0. \quad (\text{A1})$$

The WKB ansatz is

$$\mathcal{X}(\tilde{x}, \tilde{t}) = \exp \left(i \epsilon^{-1} \sum_{j=0}^{\infty} \epsilon^j S_j(\tilde{x}, \tilde{t}) \right) + \text{c.c.} \quad (\text{A2})$$

At $O(\epsilon^0)$ we have the eikonal equation

$$- \left(\frac{\partial S_0}{\partial \tilde{t}} \right)^2 + \left(\frac{\partial S_0}{\partial \tilde{x}} \right)^2 + Ro_f^{-1} + \mathcal{F}(\tilde{x}) = 0. \quad (\text{A3})$$

Introducing $p_{\tilde{t}} := -\partial S_0 / \partial \tilde{t}$ and $p_{\tilde{x}} := \partial S_0 / \partial \tilde{x}$, (A3) can be solved by the method of characteristics. The dispersion relation and ray-tracing equations are

$$p_{\tilde{t}}^2 = p_{\tilde{x}}^2 + Ro_f^{-1} + \mathcal{F}(\tilde{x}), \quad (\text{A4a})$$

$$\frac{d\tilde{t}}{ds} = 2\lambda p_{\tilde{t}}, \quad \frac{d\tilde{x}}{ds} = 2\lambda p_{\tilde{x}}, \quad \frac{dp_{\tilde{t}}}{ds} = 0, \quad \frac{dp_{\tilde{x}}}{ds} = -\lambda \frac{\partial \mathcal{F}}{\partial \tilde{x}} p_{\tilde{t}}, \quad (\text{A4b})$$

where s parameterises the rays and λ is a constant. Equation (A4b) can be combined to express the conservation of frequency along the rays in $(\tilde{x}, \tilde{t})$ space,

$$\left(\frac{\partial}{\partial \tilde{t}} + \frac{p_{\tilde{x}}}{p_{\tilde{t}}} \frac{\partial}{\partial \tilde{x}} \right) p_{\tilde{t}} = 0, \quad (\text{A5})$$

where $p_{\tilde{x}}/p_i = (d\tilde{x}/ds)/(d\tilde{f}/ds)$ is the group velocity. At $O(\epsilon^1)$ we have the transport equation

$$i \frac{\partial^2 S_0}{\partial \tilde{t}^2} - 2 \frac{\partial S_0}{\partial \tilde{t}} \frac{\partial S_1}{\partial \tilde{t}} - i \frac{\partial^2 S_0}{\partial \tilde{x}^2} + 2 \frac{\partial S_0}{\partial \tilde{x}} \frac{\partial S_1}{\partial \tilde{x}} = 0. \quad (\text{A6})$$

Making the substitution $S_1 = -i \log A = -(1/2)i \log A^2$, this can be formed into a conservation law

$$\frac{\partial}{\partial \tilde{t}}(p_i A^2) + \frac{\partial}{\partial \tilde{x}}(p_{\tilde{x}} A^2) = 0. \quad (\text{A7})$$

Utilising (A5) we can manipulate the transport equation into its most useful form

$$\frac{\partial}{\partial \tilde{t}}(A^2) + \frac{\partial}{\partial \tilde{x}} \left(\frac{p_{\tilde{x}}}{p_i} A^2 \right) = 0. \quad (\text{A8})$$

Unwrapping the temporal rescaling to express the equations in terms of $\tilde{t}$, $\tilde{x}$, and phase $\theta := \epsilon^{-1} S_0$ gives the ray-tracing results stated in § 3.3.

A.2. Fixed Bu_m

We also consider the distinguished limits $\gamma_m \rightarrow 0$ with Bu_m or α_m fixed. In the former case, the derivation is exactly the same as above except that the term

$$Ro_f^{-1} \equiv \gamma_m Bu_m^{-1} \equiv \epsilon^2 Bu_m^{-1} \quad (\text{A9})$$

drops out of (A3) and (A4a). In practice however, it is useful to retain this lower-order term in the dispersion relation as it becomes relevant in the far field as $\mathcal{F} \rightarrow 0$.

A.3. Fixed α_m

The latter case requires more work as

$$Ro_f^{-1} \equiv \gamma_m^{-1} \alpha_m^{-2} \equiv \epsilon^{-2} \alpha_m^{-2} \quad (\text{A10})$$

is large and must be accounted for at leading order. However, this is precisely the case where YBJ theory applies. Making the substitution (2.10), i.e. $\mathcal{X} = \mathcal{A}e^{-i\tilde{t}} + \text{c.c.}$, (3.7a) becomes

$$\left[\alpha_m^{-2} \epsilon^{-2} \left(\frac{\partial^2}{\partial \tilde{t}^2} - 2i \frac{\partial}{\partial \tilde{t}} \right) - \epsilon^2 \frac{\partial^2}{\partial \tilde{x}^2} + \mathcal{F}(\tilde{x}) \right] \mathcal{A} = 0. \quad (\text{A11})$$

Here, the appropriate rescaling of time is $\tilde{t} := \alpha_m^2 \epsilon^3 \tilde{t}$ giving

$$\left[\alpha_m^2 \epsilon^4 \frac{\partial^2}{\partial \tilde{t}^2} - 2i \epsilon \frac{\partial}{\partial \tilde{t}} - \epsilon^2 \frac{\partial^2}{\partial \tilde{x}^2} + \mathcal{F}(\tilde{x}) \right] \mathcal{A} = 0. \quad (\text{A12})$$

The first term will play no role until $O(\epsilon^2)$ and so the YBJ approximation is valid to that order. The WKB ansatz is

$$\mathcal{A}(\tilde{x}, \tilde{t}) = \exp \left(i \epsilon^{-1} \sum_{j=0}^{\infty} \epsilon^j S_j(\tilde{x}, \tilde{t}) \right) \quad (\text{A13})$$

and the eikonal equation is

$$2 \frac{\partial S_0}{\partial \tilde{t}} + \left(\frac{\partial S_0}{\partial \tilde{x}} \right)^2 + \mathcal{F}(\tilde{x}) = 0. \quad (\text{A14})$$

Again applying the method of characteristics, we have the dispersion relation

$$p_{\tilde{t}} = p_{\tilde{x}}^2 + \mathcal{F}(\tilde{x}), \quad (\text{A15a})$$

and the ray-tracing equations

$$\frac{d\tilde{t}}{ds} = 1, \quad \frac{d\tilde{x}}{ds} = 2\lambda p_{\tilde{x}}, \quad \frac{dp_{\tilde{t}}}{ds} = 0, \quad \frac{dp_{\tilde{x}}}{ds} = -\lambda \frac{\partial \mathcal{F}}{\partial \tilde{x}} p_{\tilde{t}}, \quad (\text{A15b-e})$$

where $p_{\tilde{t}} = \partial S_0 / \partial \tilde{t}$ and $p_{\tilde{x}} = \partial S_0 / \partial \tilde{x}$. At $O(\epsilon^1)$ the transport equation is

$$2 \frac{\partial S_1}{\partial \tilde{t}} - i \frac{\partial^2 S_0}{\partial \tilde{x}^2} + 2 \frac{\partial S_0}{\partial \tilde{x}} \frac{\partial S_1}{\partial \tilde{x}} = 0. \quad (\text{A16})$$

Once again making the substitution $S_1 = -i \log A = -(1/2)i \log A^2$ we immediately get the conservation law

$$\frac{\partial}{\partial \tilde{t}}(A^2) + \frac{\partial}{\partial \tilde{x}}(p_{\tilde{x}} A^2) = 0. \quad (\text{A17})$$

Once more, unwrapping the temporal rescaling to express the equations in terms of $\tilde{t}$, $\tilde{x}$, and phase $\theta := \epsilon^{-1} S_0$ gives the YBJ ray-tracing results stated in § 3.3.

Appendix B. Solution for a delta-function filament

We convert (3.21a) to an integral equation by means of a Green's function. We write

$$\mathcal{L}_{KG} \mathcal{X} = \left[\alpha_m^{-2} \left(1 + \frac{\partial^2}{\partial \tilde{t}^2} \right) - \frac{\partial^2}{\partial \tilde{x}^2} \right] \mathcal{X} = -\gamma_m \mathcal{F}(\gamma_m \tilde{x}) \mathcal{X}. \quad (\text{B1a})$$

The Klein–Gordon operator $\mathcal{L}_{KG}$ has constant coefficients and causal Green's function

$$G(\tilde{x} - \tilde{x}', \tilde{t} - \tilde{t}') = \begin{cases} \frac{1}{2} \alpha_m J_0 \left(\sqrt{(\tilde{t} - \tilde{t}')^2 - \alpha_m^{-2} (\tilde{x} - \tilde{x}')^2} \right), & \tilde{t} - \tilde{t}' > \alpha_m^{-1} |\tilde{x} - \tilde{x}'|, \\ 0, & \text{otherwise,} \end{cases} \quad (\text{B1b})$$

where J_0 is the zeroth-order Bessel function of the first kind (Polyanin & Nazaikinskii 2016). Including a term $\cos \tilde{t}$ that satisfies the initial conditions and the homogeneous Klein–Gordon equation, the solution satisfies

$$\mathcal{X}(\tilde{x}, \tilde{t}) = \cos \tilde{t} - \int_0^{\tilde{t}} \int_{-\infty}^{\infty} G(\tilde{x} - \tilde{x}', \tilde{t} - \tilde{t}') \gamma_m \mathcal{F}(\gamma_m \tilde{x}') \mathcal{X}(\tilde{x}', \tilde{t}') d\tilde{x}' d\tilde{t}'. \quad (\text{B1c})$$

We now make the limit $\gamma_m \rightarrow \infty$ and treat the filament as a delta-function. Evaluating the spatial integral we have

$$\mathcal{X}(\tilde{x}, \tilde{t}) = \cos \tilde{t} - \int_0^{\tilde{t}} G(\tilde{x}, \tilde{t} - \tilde{t}') \mathcal{X}(0, \tilde{t}') d\tilde{t}'. \quad (\text{B2})$$

Substituting for the Green's function produces (3.22a).

Setting $\tilde{x} = 0$ we have an integral equation for the solution at the centre of the filament,

$$\mathcal{X}(0, \tilde{t}) = \cos \tilde{t} - \frac{1}{2} \alpha_m \int_0^{\tilde{t}} J_0(\tilde{t} - \tilde{t}') \mathcal{X}(0, \tilde{t}') d\tilde{t}'. \quad (\text{B3})$$

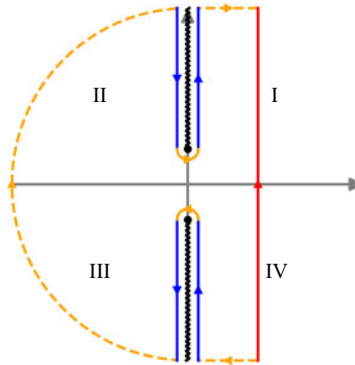


Figure 12. Deformed contour for the Bromwich integral. The original contour is in red but can be deformed into the orange and blue contour. The branch points are at $s = \pm i$, indicated by black circles and the branch cuts run along the imaginary axis through infinity. The dashed orange lines indicate contours at infinity. The four quadrants are indicated by roman numerals.

Taking a Laplace transform, $\hat{\mathcal{X}} = \int_0^\infty \mathcal{X} e^{s\tilde{t}} d\tilde{t}$, converts the convolution into a product

$$\hat{\mathcal{X}}(0, s) = \frac{s}{1+s^2} - \frac{1}{2} \alpha_m \frac{1}{\sqrt{1+s^2}} \hat{\mathcal{X}}(0, s). \quad (\text{B4})$$

The solution can then be expressed as a Bromwich integral

$$\mathcal{X}(0, \tilde{t}) = \frac{1}{2\pi i} \int_{\varsigma-i\infty}^{\varsigma+i\infty} \frac{s}{\sqrt{1+s^2}} \frac{1}{\frac{1}{2}\alpha_m + \sqrt{1+s^2}} e^{s\tilde{t}} ds, \quad (\text{B5})$$

where the contour is taken over a line $\text{Re}(s) = \varsigma > 0$ that lies to the right of the branch points at $s = \pm i$.

Connecting the branch points with a branch cut along the imaginary axis through infinity, we deform the contour as indicated in figure 12. The contour is deformed in this way so that the dashed orange contours at infinity do not contribute to the integral. This can be seen by applying Jordan's lemma to the semi-circular contour and the estimation lemma to the remaining parts. The small orange contours around the branch points also do not contribute to the integral. For example, putting $s = i + \epsilon e^{i\theta}$, $\theta \in (-\pi, 0)$, we find that the integrand is $O(\sqrt{\epsilon})$ and, hence, the integral vanishes as $\epsilon \rightarrow 0$.

On the blue contours we make a substitution. In the quadrants I and II we put $s = i\sqrt{1+\sigma^2}$, whereas in quadrants III and IV we put $s = -i\sqrt{1+\sigma^2}$. In all cases $\sigma \in (0, \infty)$. With this substitution we have $\sqrt{1+s^2} = i\sigma$ in quadrants I and III and $\sqrt{1+s^2} = -i\sigma$ in quadrants II and IV. Summing the four pieces the integral reduces to

$$\begin{aligned} \mathcal{X}(0, \tilde{t}) &= \frac{1}{2\pi} \int_0^\infty \left[\frac{e^{i\tilde{t}\sqrt{1+\sigma^2}}}{\frac{1}{2}\alpha_m + i\sigma} + \frac{e^{i\tilde{t}\sqrt{1+\sigma^2}}}{\frac{1}{2}\alpha_m - i\sigma} + \frac{e^{-i\tilde{t}\sqrt{1+\sigma^2}}}{\frac{1}{2}\alpha_m + i\sigma} + \frac{e^{-i\tilde{t}\sqrt{1+\sigma^2}}}{\frac{1}{2}\alpha_m - i\sigma} \right] d\sigma \quad (\text{B6}) \\ &= \frac{1}{\pi} \int_0^\infty \frac{\alpha_m \cos(\tilde{t}\sqrt{1+\sigma^2})}{\sigma^2 + \frac{1}{4}\alpha_m^2} d\sigma. \end{aligned}$$

The substitution $\sigma = (1/2)\alpha_m u$ gives (3.22b).

Appendix C. Numerical solution of the Klein–Gordon and time-dependent Schrödinger equations

Discretising the spatial derivatives with a fourth-order accurate centred five point finite difference scheme, the Klein–Gordon equation (3.4) is written as

$$\mathcal{X}_{\tilde{t}\tilde{t}} = -\mathbf{M}\mathcal{X}, \quad (\text{C1})$$

where subscripts denote time derivatives and $\mathbf{M}$ is a positive definite symmetric matrix (assuming the flow is not inertially unstable). Similarly, the time-dependent Schrödinger equation can be written as

$$\mathcal{A}_{\tilde{t}} = -\frac{1}{2}i(\mathbf{M} - \mathbf{I})\mathcal{A}, \quad (\text{C2})$$

where $\mathbf{I}$ is the identity matrix. We use a three-stage fourth-order diagonally implicit Runge–Kutta (Nyström) schemes to time step these equations.

For the Klein–Gordon equation, the state variables $\mathcal{X}^n$ and $\mathcal{X}_{\tilde{t}}^n$ are advanced with time step h according to

$$\mathcal{X}^{n+1} = \mathcal{X}^n + h\mathcal{X}_{\tilde{t}}^n + h^2 \sum_{j=1}^3 b_j \mathcal{X}_{\tilde{t}\tilde{t}}^{n+c_j}, \quad (\text{c3a})$$

$$\mathcal{X}_{\tilde{t}}^{n+1} = \mathcal{X}_{\tilde{t}}^n + h \sum_{j=1}^3 b'_j \mathcal{X}_{\tilde{t}\tilde{t}}^{n+c_j}, \quad (\text{c3b})$$

and the intermediate steps are given by the implicit equations

$$\mathcal{X}^{n+c_i} = \mathcal{X}^n + c_i h \mathcal{X}_{\tilde{t}}^n + h^2 \sum_{k=1}^i a_{ik} \mathcal{X}_{\tilde{t}\tilde{t}}^{n+c_k}. \quad (\text{C3c})$$

The coefficients are

$$\begin{array}{c|ccc|c|ccc} c_1 & a_{11} & & & \frac{3}{5} & \frac{9}{50} & & \\ c_2 & a_{21} & a_{22} & & \frac{9}{10} & \frac{9}{40} & \frac{9}{50} & \\ c_3 & a_{31} & a_{32} & a_{33} & = \frac{6}{37} & \frac{234}{1\ 266\ 325} & \frac{-891}{2\ 532\ 650} & \frac{9}{50} \\ & b_1 & b_2 & b_3 & & \frac{115}{729} & \frac{55}{2457} & \frac{42\ 439}{132\ 678} \\ & b'_1 & b'_2 & b'_3 & & \frac{575}{1458} & \frac{550}{2457} & \frac{50\ 653}{132\ 678} \end{array} \quad (\text{C4})$$

and the resulting scheme is unconditionally stable (Sharp, Fine & Burrage 1990).

For the time-dependent Schrödinger equation, the state variable $\mathcal{A}$ is advanced according to

$$\mathcal{A}^{n+1} = \mathcal{A}^n + h \sum_{j=1}^3 \bar{b}_j \mathcal{A}_{\tilde{t}\tilde{t}}^{n+\bar{c}_j}, \quad (\text{C5})$$

with the intermediate steps given by

$$\mathcal{A}^{n+\bar{c}_i} = \mathcal{A}^n + h \sum_{k=1}^i \bar{a}_{ik} \mathcal{A}_{\tilde{t}\tilde{t}}^{n+\bar{c}_k}. \quad (\text{C6})$$

The coefficients are

$$\begin{array}{c|ccc} \bar{c}_1 & \bar{a}_{11} & & \\ \bar{c}_2 & \bar{a}_{21} & \bar{a}_{22} & \\ \bar{c}_3 & \bar{a}_{31} & \bar{a}_{32} & \bar{a}_{33} \end{array} = \begin{array}{c} \gamma \\ \frac{1}{2} \\ 1 - \gamma \end{array} \begin{array}{ccc} \gamma & & \\ \frac{1}{2} - \gamma & \gamma & \\ 2\gamma & 1 - 4\gamma & \gamma \end{array} \quad (C7)$$

$$\begin{array}{c|ccc} & \bar{b}_1 & \bar{b}_2 & \bar{b}_3 \end{array} \begin{array}{ccc} \frac{2\gamma(\gamma-1)}{2\gamma-1} & \frac{6\gamma-1}{6\gamma(2\gamma-1)} & \frac{2\gamma(\gamma-1)}{2\gamma-1} \end{array}$$

with $\gamma = (3 + 2\sqrt{3} \cos(\pi/18))/6$ (Kennedy & Carpenter 2016). This scheme is also unconditionally stable.

Appendix D. The WKBJ solution

Here we derive the connection formula for the WKBJ solution. The derivation, by and large, closely follows Bender & Orszag (1999) and uses some intermediate results given therein. Define $Q_n(\hat{x}; \lambda_n) := (\partial V / \partial \hat{x} - \lambda_n) / \Gamma_m$ such that the Schrödinger equation may be written as

$$\frac{\partial^2 \mathcal{X}_n}{\partial \hat{x}^2} = Q_n(\hat{x}; \lambda_n) \mathcal{X}_n. \quad (D1)$$

In § 4.3.2 we defined $k_n := \sqrt{|Q_n|}$. Applying even boundary conditions the solutions away from the turning point are

$$\mathcal{X}_{2n} = C_1 |Q_{2n}|^{-1/4} \cosh \left(\int_0^{\hat{x}} \sqrt{Q_{2n}} d\hat{x}' \right), \quad \mathcal{X}_{2n} = C_2 |Q_{2n}|^{-1/4} \cos \left(\int_{\hat{x}}^1 \sqrt{-Q_{2n}} d\hat{x}' \right) \quad (D2)$$

for $\hat{x} < \hat{x}_*$ and $\hat{x} > \hat{x}_*$, respectively. Near the turning point $\hat{x}_*$ defined by $Q_{2n}(\hat{x}_*) = 0$ the Schrödinger equation is the Airy equation

$$\frac{\partial^2 \mathcal{X}_{2n}}{\partial \hat{x}^2} = a_{2n}^3 (\hat{x} - \hat{x}_*) \mathcal{X}_{2n}, \quad (D3)$$

where $a_{2n}^3 := -(\partial Q_{2n} / \partial \hat{x})|_{\hat{x}_*} > 0$. Defining $t := a_{2n}(\hat{x} - \hat{x}_*)$ the solution near the turning point is

$$\mathcal{X}_{2n} = C_A \text{Ai}(-t) + C_B \text{Bi}(-t). \quad (D4)$$

Near the turning point we have $Q_{2n}(x) = -a_{2n}^2 t$ and the solution entering the filament, $t < 0$, is

$$\mathcal{X}_{2n} = C_1 a_{2n}^{-1/2} (-t)^{-1/4} \cosh \left(A_{2n} - \frac{2}{3} (-t)^{3/2} \right), \quad A_{2n} := \int_0^{\hat{x}_*} \sqrt{Q_{2n}} d\hat{x}'. \quad (D5a,b)$$

Comparing to the asymptotic forms of the Airy functions as $t \rightarrow -\infty$,

$$\text{Ai}(-t) \sim \frac{1}{2\sqrt{\pi}} (-t)^{-1/4} \exp \left(-\frac{2}{3} (-t)^{3/2} \right), \quad \text{Bi}(-t) \sim \frac{1}{\sqrt{\pi}} (-t)^{-1/4} \exp \left(\frac{2}{3} (-t)^{3/2} \right), \quad (D6a,b)$$

we find the matching conditions.

$$C_A = a_{2n}^{-1/2} \sqrt{\pi} e^{A_{2n}} C_1, \quad C_B = \frac{1}{2} a_{2n}^{-1/2} \sqrt{\pi} e^{-A_{2n}} C_1. \quad (D7a,b)$$

Similarly, for $t > 0$, we have

$$\mathcal{X}_{2n} = C_2 a_{2n}^{-1/2} t^{-1/4} \cos \left(B_{2n} - \frac{2}{3} t^{3/2} \right), \quad B_{2n} := \int_{\hat{x}_*}^1 \sqrt{-Q_{2n}} \, d\hat{x}' \quad (\text{D8a},b)$$

and, as $t \rightarrow \infty$, the asymptotic forms of the Airy functions are

$$\text{Ai}(-t) \sim \frac{1}{\sqrt{\pi}} t^{-1/4} \sin \left(\frac{2}{3} t^{3/2} + \frac{\pi}{4} \right), \quad \text{Bi}(-t) \sim \frac{1}{\sqrt{\pi}} t^{-1/4} \cos \left(\frac{2}{3} t^{3/2} + \frac{\pi}{4} \right). \quad (\text{D9a},b)$$

Therefore, the matching conditions are

$$C_A = a_{2n}^{-1/2} \sqrt{\pi} \sin \left(B_{2n} + \frac{\pi}{4} \right) C_2, \quad C_B = a_{2n}^{-1/2} \sqrt{\pi} \cos \left(B_{2n} + \frac{\pi}{4} \right) C_2. \quad (\text{D10a},b)$$

Finally, we have the connection formula

$$\frac{C_B}{C_A} = \frac{1}{2} e^{-2A_{2n}} = \cot \left(B_{2n} + \frac{\pi}{4} \right) = \tan \left(\frac{\pi}{4} - B_{2n} \right). \quad (\text{D11})$$

The left-hand side of this connection formula is bounded $0 < (1/2)e^{-2A_{2n}} < (1/2)$ since A_{2n} is positive. Furthermore, we know that the $2n$ th eigenmode of the symmetric periodic Sturm–Liouville problem has n zero crossing in the half-domain $0 < \hat{x} < 1$. Therefore, we can bound B_{2n} by

$$n\pi + \frac{\pi}{4} - \arctan \frac{1}{2} < B_{2n} < n\pi + \frac{\pi}{4}. \quad (\text{D12})$$

REFERENCES

- 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, S.G. & YOUNG, W.R. 1998 Enhanced dispersion of near-inertial waves in an idealized geostrophic flow. *J. Mar. Res.* **56**, 1–40.
- BENDER, C.M. & ORSZAG, S.A. 1999 WKB Theory. In *Advanced mathematical methods for scientists and engineers I: asymptotic methods and perturbation theory*, chap. 10, pp. 484–543. Springer, New York.
- CALLIES, JÖRN, FERRARI, R., KLYMAK, J.M. & GULA, J. 2015 Seasonality in submesoscale turbulence. *Nat. Commun.* **6** (1), 6862.
- CONN, S., CALLIES, J. & LAWRENCE, A. 2025 Regimes of near-inertial wave dynamics. *J. Fluid Mech.* **1002**, A22.
- 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.
- GILL, A.E. 1984 On the behavior of internal waves in the wakes of storms. *J. Phys. Oceanogr.* **14** (7), 1129–1151.
- 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.
- KENNEDY, C.A. & CARPENTER, M.H. 2016 Diagonally implicit Runge–Kutta methods for ordinary differential equations. A review. *Tech. Rep.* NF1676L-19716, NASA Langley Research Center, NTRS Document ID: 20160005923.
- KLEIN, P. & SMITH, S.L. 2001 Horizontal dispersion of near-inertial oscillations in a turbulent mesoscale eddy field. *J. Mar. Res.* **59** (5), 697–723.
- KUNZE, E. 1985 Near-inertial wave propagation in geostrophic shear. *J. Phys. Oceanogr.* **15** (5), 544–565.

- KUO, H.-L. 1949 Dynamic instability of two-dimensional nondivergent flow in a barotropic atmosphere. *J. Meteorol.* **6** (2), 105–122.
- MARTÍNEZ-MARRERO, A., BARCELÓ-LLULL, B., PALLÀS-SANZ, E., AGUIAR-GONZÁLEZ, B., ESTRADA-ALLIS, S.N., GORDO, C., GRISOLÍA, D., RODRÍGUEZ-SANTANA, A. & ARÍSTEGUI, J. 2019 Near-inertial wave trapping near the base of an anticyclonic mesoscale eddy under normal atmospheric conditions. *J. Geophys. Res.: Oceans* **124** (11), 8455–8467.
- MCWILLIAMS, J.C. 2016 Submesoscale currents in the ocean. *Proc. R. Soc. A: Math. Phys. Engng Sci.* **472**(2189), 20160117.
- POLYANIN, A.D. & NAZAIKINSKII, V.E. 2016 *Handbook of Linear Partial Differential Equations for Engineers and Scientists*. 2nd edn. Chapman and Hall / CRC.
- 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.
- SCHLICHTING, D., QU, L., KOBASHI, D. & HETLAND, R. 2023 Quantification of physical and numerical mixing in a coastal ocean model using salinity variance budgets. *J. Adv. Model. Earth Syst.* **15** (4), e2022MS003380.
- SHARP, P.W., FINE, J.M. & BURRAGE, K. 1990 Two-stage and three-stage diagonally implicit Runge–Kutta Nyström methods of orders three and four. *IMA J. Numer. Anal.* **10** (4), 489–504.
- SHCHERBINA, A.Y., D’ASARO, E.A., LEE, C.M., KLYMAK, J.M., MOLEMAKER, M.J. & MCWILLIAMS, J.C. 2013 Statistics of vertical vorticity, divergence, and strain in a developed submesoscale turbulence field. *Geophys. Res. Lett.* **40** (17), 4706–4711.
- STERMAN, G. 1993 *An Introduction to Quantum Field Theory, Chap. 1: Classical Fields and Symmetries*. Cambridge University Press, 3–28.
- TAYLOR, J.R. & THOMPSON, A.F. 2023 Submesoscale dynamics in the upper ocean. *Annu. Rev. Fluid Mech.* **55** (1), 103–127.
- THOMAS, L.N. 2019 Enhanced radiation of near-inertial energy by frontal vertical circulations. *J. Phys. Oceanogr.* **49** (9), 2407–2421.
- THOMAS, L., KELLY, S., KLENZ, T., YOUNG, W., RAINVILLE, L., SIMMONS, H., HORMANN, V. & STOKES, I. 2024 Why near-inertial waves are less affected by vorticity in the Northeast Pacific than in the North Atlantic. *Oceanography* **37** (4), 10–21.
- 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.
- THOMAS, L.N., TANDON, A. & MAHADEVAN, A. 2008 *Submesoscale Processes and Dynamics*. American Geophysical Union (AGU).
- YOUNG, W.R. & JELLOUL, M.B. 1997 Propagation of near-inertial oscillations through a geostrophic flow. *J. Mar. Res.* **55** (4), 735–766.
- ZHANG, Z., HETLAND, R. & ZHANG, X. 2014 Wind-modulated buoyancy circulation over the Texas–Louisiana shelf. *J. Geophys. Res.: Oceans* **119** (9), 5705–5723.